



# **La investigación científica en el aula: de la transmisión a la creación del conocimiento**

Óscar Moreno Díaz

Catálogo de publicaciones del Ministerio: [sede.educacion.gob.es/publiventa](http://sede.educacion.gob.es/publiventa)  
Catálogo general de publicaciones oficiales: [publicacionesoficiales.boe.es](http://publicacionesoficiales.boe.es)

*La investigación científica en el aula: de la transmisión a la creación del conocimiento*

Autor  
Óscar Moreno Díaz

Esta publicación constituye un material de guía para participar en el Certamen Jóvenes Investigadores que convoca la Secretaría General de **Universidades en colaboración con el Instituto de la Juventud (INJUVE)**



MINISTERIO DE EDUCACIÓN Y FORMACIÓN PROFESIONAL  
INSTITUTO NACIONAL DE TECNOLOGÍAS EDUCATIVAS Y DE FORMACIÓN DEL PROFESORADO

MINISTERIO DE UNIVERSIDADES

Coeditan:

© SECRETARÍA GENERAL TÉCNICA DEL MINISTERIO DE EDUCACIÓN Y FORMACIÓN PROFESIONAL  
© SECRETARÍA GENERAL TÉCNICA DEL MINISTERIO DE UNIVERSIDADES

Edición: 2020  
NIPO (línea) MEFP: 847-20-108-X  
NIPO (línea) MUNI: 097-20-004-8

ISBN 978-84-369-5960-4

Maquetación: Cristina Rico

<b>Introducción</b> .....	8
<b>PARTE I</b>	
<b>Capítulo 1.- A ciencia cierta</b> .....	12
1.1. El oráculo científico: explicar, predecir, y otros objetivos de la ciencia .....	12
- Historias de la ciencia. Eddington y las dos mesas .....	14
1.2. <i>Arbor scientiae</i> : la clasificación de las ciencias.....	15
<b>Capítulo 2.- Cuestión de protocolo: los procedimientos generales de la ciencia</b> .....	19
2.1. Si no lo veo, no lo creo: observación y experimentación .....	20
- Ampliación. La ciencia según descartes y diderot.....	21
- Historias de la ciencia. Eddington y el eclipse que probó la relatividad .....	22
- Historias de la ciencia. Experimentación al límite en el LHC .....	23
2.2. Divide y vencerás. División y clasificación en la ciencia .....	24
- Ampliación. División física y división esencial.....	26
- Historias de la ciencia. Los átomos, invisibles e indivisibles? .....	27
2.3. Definir los límites: definición de conceptos y terminología científica .....	28
- Ampliación. Otras definiciones y el lenguaje natural .....	29
2.4. El color de los cisnes: la inducción.....	30
- Historias de la ciencia. Cisnes negros.....	30
2.5. Lógica en estado puro: la deducción.....	31
- Ampliación. Propositiones y clases, reglas y leyes .....	33
2.6. Nos merecemos una explicación: la explicación en ciencia.....	34
- Ampliación. La lógica de la explicación .....	35
<b>Capítulo 3.- No hay ciencia sin método: una visión actual del método científico</b> .....	36
3.1. Espiral de conocimiento: el ciclo del método científico.....	37
3.2. ¿Un modelo a seguir?: críticas al método científico .....	38
- Historias de la ciencia. La serendipia de Fleming y Becquerel .....	39
3.3. Al traste con el contraste: de la verificación a la falsación .....	40
- Ampliación. La lógica de la verificación y la falsación .....	41
3.4. Formalidad ante todo: el método de las ciencias formales.....	42
- Historias de la ciencia. Los postulados de Euclides .....	44
<b>Capítulo 4.- Rebelde con causa: la explicación causal</b> .....	45
4.1. Inventario de causas: tipos de relaciones causa-efecto.....	46
4.2. Oro parece, causa no es: falsas relaciones causales .....	47
- Historias de la ciencia. Cigüeñas con bebés, niños miopes y premios nobel golosos .....	48
<b>Capítulo 5.- Las tablas de la ley: la explicación nomológica</b> .....	50
5.1. El código legal: tipos de leyes científicas .....	51
- Ampliación. Cada ciencia tiene su explicación .....	51
5.2. La letra pequeña de la ley: la cláusula <i>ceteris paribus</i> .....	53
5.3. Tener razón a toda costa: las hipótesis <i>ad hoc</i> .....	53
- Historias de la ciencia. La errónea hipótesis <i>ad hoc</i> de Einstein que acabó siendo necesaria.....	54
<b>Capítulo 6.- Comienza la función: la explicación funcional</b> .....	57
6.1. El fin no justifica los medios: la función como efecto .....	58
- Ampliación. La lógica de la explicación funcional.....	58
6.2. La intención es lo que cuenta: explicación teleológica e intencional.....	59
6.3. Con vuelta atrás: la retroalimentación causal.....	59
- Ampliación. El mecanismo de retroalimentación en la ciencia .....	60
6.4. El cuello de las jirafas: la explicación funcional en biología .....	60
- Ampliación. La síntesis evolutiva .....	62
- Historias de la ciencia. Darwin y el pico de los pinzones .....	63
6.5. Fallo estructural: la falacia estructuralista .....	64
- Historias de la ciencia. Vacas sagradas.....	65

<b>Capítulo 7.- Fantasía y realidad: los límites de la ciencia</b> .....	67
- Ampliación. Ciencia y creencia.....	68
- Historias de la ciencia. Qué es la ciencia según un juez de Arkansas.....	70
<b>PARTE II</b>	
<b>Capítulo 8.- Metodologías de investigación</b> .....	73
8.1. Metodología cuantitativa.....	75
8.2. Metodología cualitativa.....	76
- Ampliación. Historia, comparación, crítica y etnografía.....	76
8.3. Metodologías y tipos de ciencias.....	78
8.4. Triangulación.....	78
- Esquema del capítulo.....	79
<b>Capítulo 9.- Variables de una investigación</b> .....	80
9.1. Escalas de medida.....	83
- Ampliación. Escalas de intervalo y de razón.....	85
Esquema del capítulo.....	85
<b>Capítulo 10.- Diseños de investigación</b> .....	87
- Ampliación. Diseños transversales y longitudinales.....	89
- Ampliación. Algunos tipos de diseños no experimentales.....	90
- Ampliación. Diseños factoriales y cuadrado latino.....	92
- Historias de la ciencia. Fisher y los cuadrados latinos en agricultura.....	93
- Esquema del capítulo.....	94
<b>Capítulo 11.- Control de variables</b> .....	95
- Ampliación. Métodos de aleatorización.....	97
- Ampliación. Simbolización del control y el diseño de Solomon.....	98
- Esquema del capítulo.....	99
<b>Capítulo 12.- Validez de una investigación</b> .....	100
12.1. Efectos sobre la validez interna.....	101
- Historias de la ciencia. De la regresión a la media a la regresión en la educación.....	103
12.2. Efectos sobre la validez externa.....	104
- Esquema del capítulo.....	105
<b>PARTE III</b>	
<b>Capítulo 13. Estadística descriptiva</b> .....	108
13.1. Medidas de centralización.....	108
- Historias de la ciencia. La ciencia del Estado.....	110
13.2. Medidas de dispersión.....	111
13.3. Medidas de forma.....	112
13.4. Medidas de posición.....	113
13.5. Medidas de correlación.....	113
- Ampliación. Otros coeficientes de correlación.....	116
- Esquema del capítulo.....	117
<b>Capítulo 14.- Estadística inferencial: distribución muestral de estadísticos</b> .....	118
14.1. Probabilidad.....	120
- Ampliación. Las matemáticas de la probabilidad.....	121
14.2. Distribución muestral de estadísticos.....	122
- Ampliación. Probabilidad en variables continuas.....	128
- Ampliación. El teorema central del límite.....	129
- Historias de la ciencia. La máquina de Galton.....	130
- Historias de la ciencia. El estudiante cervecero.....	131
<b>Capítulo 15.- Estadística inferencial: estimación de parámetros y contraste de hipótesis</b> .....	134
15.1. Estimación de parámetros.....	135
- Ampliación. Tamaño muestral.....	139
15.2. Contraste de hipótesis.....	139
- Ampliación. El valor <i>P</i> .....	146
15.3. Errores en el contraste.....	147
- Historias de la ciencia. Falsos positivos y falsas intuiciones.....	148
- Ampliación. Contraste con otros parámetros.....	149
- Ampliación. Análisis de varianza.....	152
- Esquema del capítulo.....	154

<b>Capítulo 16.- Medidas numéricas</b> .....	155
16.1. Incertidumbre en medidas directas.....	157
- Ampliación. Matices sobre incertidumbres.....	159
16.2. Exactitud y precisión .....	160
- Ampliación. Exactitud y precisión en ciencias sociales .....	161
16.3. Incertidumbre en medidas indirectas.....	162
16.4. Cifras significativas .....	164
16.5. Presentación de medidas .....	166
16.6. Representación gráfica.....	167
- Esquema del capítulo.....	172
 <b>PARTE IV</b>	
<b>Capítulo 17.- La investigación científica en el aula</b> .....	175
17.1. Objetivos y competencias clave .....	175
17.2. Metodologías de aprendizaje de la investigación científica .....	178
17.3. La investigación como proyecto del centro.....	181
17.4. Motivación .....	182
- Historias de la ciencia. Isaac Asimov.....	183
<b>Capítulo 18.- El Certamen Jóvenes Investigadores</b> .....	184
18.1. Requisitos .....	185
18.2. Documentación.....	185
18.3. Criterios de valoración.....	186
18.4. Premios.....	187
18.5. Orientaciones sobre los temas de investigación .....	188
- Ampliación. Otros certámenes para jóvenes investigadores .....	189
<b>Capítulo 19.- La memoria de la investigación</b> .....	191
19.1. Portada e índice.....	192
19.2. Resumen.....	192
19.3. Introducción.....	193
19.4. Antecedentes.....	194
19.5. Hipótesis y objetivos .....	195
19.6. Materiales y métodos .....	196
19.7. Resultados .....	196
19.8. Conclusiones .....	197
19.9. Agradecimientos.....	197
19.10. Bibliografía y webgrafía .....	198
19.11. Anexo de figuras .....	198
- Ampliación. Tipografía y edición de la memoria.....	199
<b>Capítulo 20.- Fuentes documentales</b> .....	200
20.1. Fuentes de datos brutos.....	201
20.2. Fuentes de datos elaborados .....	202
20.3. Cómo citar las fuentes.....	203
20.4. La investigación secundaria .....	206
<b>Capítulo 21.- Presentación pública de la investigación</b> .....	208
21.1. El póster científico .....	209
21.2. La exposición oral .....	212
 <b>Epílogo</b> .....	218
 <b>Para saber más</b> .....	221
 <b>APÉNDICES</b>	
<b>Apéndice I. Actividades para jugar con el método científico</b> .....	227
Actividad 1: El juego de la lógica .....	227
Actividad 2: El juego de la falsación de hipótesis.....	228
Actividad 3: El juego del método científico.....	230
Actividad 4: El juego del sudoku para diseños factoriales .....	232
<b>Apéndice II. Actividades para reflexionar sobre la ciencia</b> .....	234
Actividad 1: Reflexión sobre las fronteras entre las áreas de la ciencia .....	234
Actividad 2: Reflexión sobre la naturaleza de la causalidad .....	234
Actividad 3: Reflexión sobre los límites de la ciencia.....	235

<b>Apéndice III. Actividades para elaborar una investigación científica.....</b>	<b>236</b>
Actividad 1: Planteamiento general de la investigación .....	236
Actividad 2: Análisis de los principales elementos de una investigación general .....	237
Actividad 3: Análisis de los principales elementos de una investigación cuantitativa.....	237
Actividad 4: Elaboración y revisión de la memoria escrita y de los materiales para la presentación pública.....	237
Actividad 5: Simulación de la evaluación de la memoria .....	239

# Introducción

El Certamen Jóvenes Investigadores es convocado anualmente, desde hace más de treinta años, por la Secretaría General de Universidades. Tiene como finalidad promover la cultura científica, fomentar el pensamiento crítico y despertar vocaciones investigadoras en los jóvenes a través de la concesión de importantes premios a investigaciones realizadas por estudiantes de Enseñanza Secundaria Obligatoria, Bachillerato y Formación Profesional de entre 15 y 20 años. Los estudiantes son coordinados por un profesor tutor, que ofrece guía y apoyo durante el desarrollo de la investigación.

Este libro está dirigido a esos profesores tutores e indirectamente, y a través de ellos, a los estudiantes preuniversitarios que se aventuran en el diseño y realización de una investigación científica. Pretende ofrecer una visión general del método científico e introducir las técnicas de investigación más importantes, además de presentar el Certamen Jóvenes Investigadores. El contenido, de un nivel intermedio entre Bachillerato y primeros cursos universitarios, pretende refrescar algunos conocimientos que los profesores pueden tener parcialmente olvidados, y quizás profundizar en otros menos conocidos de antemano. Los temas aquí tratados pueden resultar también útiles e interesantes para los jóvenes autores de las investigaciones, pero es conveniente que sus profesores adapten el nivel y apliquen las técnicas didácticas y pedagógicas más adecuadas a su alumnado, que conocen mejor que nadie.

La participación en el Certamen Jóvenes Investigadores no es solamente un fin en sí mismo, sino que actúa también como excusa para introducir la investigación científica en el aula de secundaria, de manera que la enseñanza de la ciencia no se limite a la transmisión de conocimientos científicos, sino que también invite a poner en práctica los procedimientos que conducen a la creación de esos conocimientos. De hecho, son cada vez más los centros y administraciones educativas que incluyen en el currículo de las etapas previas a la universidad la realización de una investigación científica completa adaptada al nivel de los estudiantes.

Cuando se habla de investigación científica en este libro se incluyen tanto las ciencias naturales y formales (matemáticas, física, química, biología, geología, etc.) como las sociales y humanas (antropo-

logía, sociología, historia, lingüística, etc.), en igualdad de condiciones. Esto añade cierta complejidad al contenido, pero resulta enriquecedor explorar cómo se articula la ciencia en diferentes ámbitos del conocimiento, algunos de los cuales estarán alejados de la especialización académica de cada lector.

El libro se organiza en cuatro partes, cuyo contenido transcurre desde lo más abstracto a lo más práctico. La parte I trata sobre los fundamentos de la ciencia y del método científico moderno, así como de los procedimientos más universales de la ciencia (observación, experimentación, clasificación, definición, deducción lógica, inducción, contraste empírico y falsación de hipótesis, explicación, etc.). En la parte II se describen elementos básicos del diseño de investigaciones de carácter general, y de tipo cualitativo en particular, que comprenden las diferentes metodologías, el papel de las variables y su control, las fuentes de invalidez, etc. La parte III contiene métodos específicos para las investigaciones de tipo cuantitativo, ya sean de ciencias sociales o de ciencias naturales, incluyendo técnicas estadísticas y de manejo y presentación de medidas numéricas. Por último, la parte IV analiza el papel que puede desempeñar la enseñanza de la investigación científica en la educación secundaria, tanto en el aula como a través de certámenes para jóvenes investigadores; se detallan las bases del Certamen Jóvenes Investigadores de la Secretaría General de Universidades, y se ofrecen claves para elaborar la memoria escrita, el póster y la presentación oral de una investigación. Por último, los apéndices recogen actividades para asimilar o reflexionar sobre algunos de los conceptos tratados en las cuatro partes anteriores: juegos sobre el método científico, temas de discusión sobre la ciencia, y preguntas clave para elaborar y revisar una investigación para el Certamen Jóvenes Investigadores.

Este texto ha pasado por varias fases de desarrollo, partiendo de una idea embrionaria que dio lugar a un curso a distancia en línea y que ahora ha sufrido una metamorfosis a la edición como libro. Tanto su origen como su evolución han sido posibles gracias al continuo respaldo y estímulo del equipo de la Subdirección General de Formación del Profesorado Universitario y Programación de la Secretaría General de Universidades y del equipo del Área de Formación en Línea y Competencia Digital Educativa del Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado.

# PARTE I



# Capítulo 1.

## A ciencia cierta

¿Qué es la ciencia? El término proviene del latín *scientia*, que significa «conocimiento», y deriva de *sci-re*, que significa «saber», «conocer». Este verbo procede a su vez, probablemente, de la raíz indoeuropea *sketi-*, que significa «cortar», «rajar». Así, etimológicamente la ciencia requiere «separar una cosa de otra», «discernir», un origen que se refleja en algunos de sus procedimientos más habituales, como son el análisis, la división o la clasificación.

La ciencia, como actividad, consiste en generar y organizar conocimiento sobre un cierto ámbito de la realidad. Se trata de conocimiento científico, que es distinto al ordinario, aunque complementario a él. También se entiende como ciencia el resultado de esa actividad, es decir, el conjunto sistemático y organizado de conocimientos científicos de un cierto ámbito, que permite la explicación de hechos que ya han tenido lugar o la predicción de los que aún no han ocurrido. Cada ámbito de conocimiento está asociado a una ciencia distinta, pero todas tienen objetivos en común. De esos objetivos y de los distintos tipos de ciencias y su clasificación nos ocupamos en este primer capítulo.

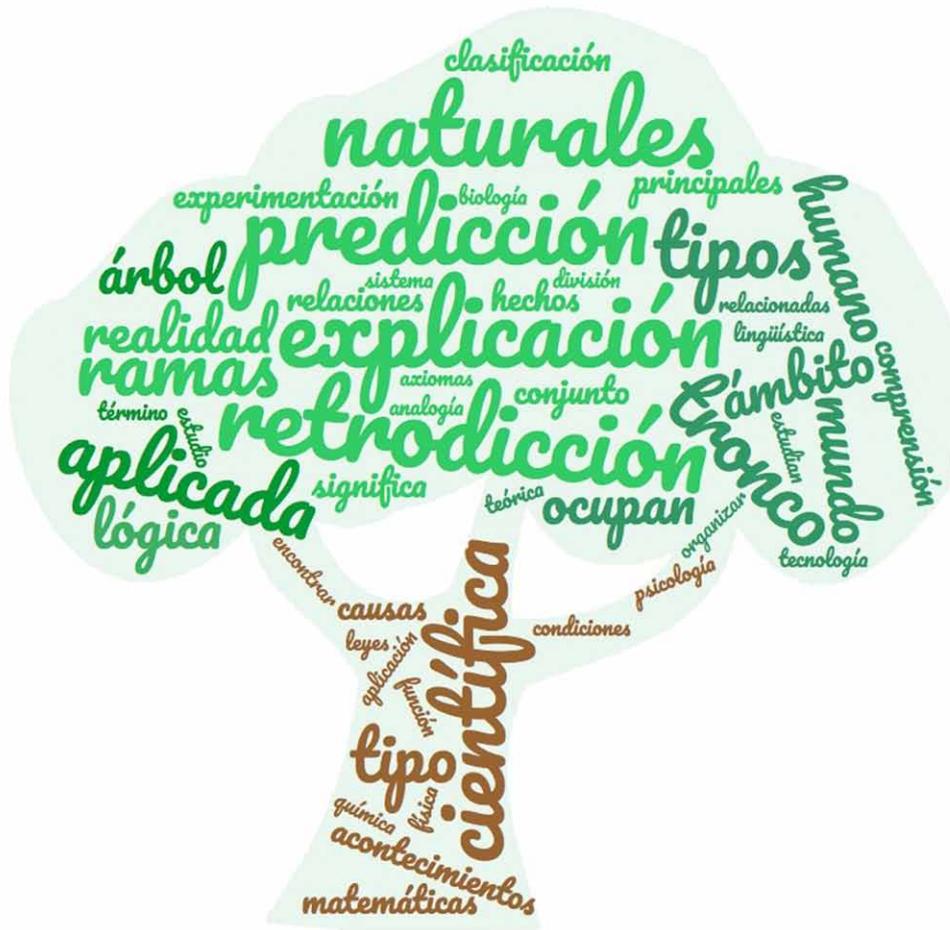
### Conceptos clave:

– Ciencia, explicación, predicción, retrodicción, comprensión, ciencia aplicada, tecnología, ciencia formal, ciencia natural, ciencia social, árbol de las ciencias.

### 1.1. El oráculo científico: explicar, predecir, y otros objetivos de la ciencia

La ciencia, para generar conocimiento, se fija los siguientes objetivos principales:

– Explicar la realidad, es decir, asociar los fenómenos a las causas o condiciones necesarias para que se produzcan, a las leyes científicas que cumplen, a las funciones que desempeñan dentro de un sistema, etc. Cada una de estas posibilidades se refiere a un tipo distinto de explicación, que trataremos en detalle en capítulos posteriores.



Fuente: elaboración propia con nubedepalabras.es.

- Predecir acontecimientos futuros. Si el conocimiento científico permite relacionar unos hechos con otros de manera sistemática (en forma, por ejemplo, de causas y efectos o de leyes que siempre se satisfacen), puede suponerse entonces que, dados unos hechos en un instante de tiempo, tendrán lugar otros hechos en un instante posterior. La predicción puede ir asociada a una certeza o bien a un cierto grado de probabilidad, en función de que la explicación en la que se base sea de tipo determinista o de tipo probabilístico.
- Retrodecir acontecimientos pasados. Muchas ciencias intentan reconstruir lo que ocurrió en el pasado en un lugar o tiempo no accesible al ser humano. Para lograr este objetivo la ciencia emplea la extrapolación (para sucesos que ocurrieron antes de determinados acontecimientos conocidos) y la interpolación (para sucesos que ocurrieron entre otros anteriores y posteriores, que sí son conocidos).
- Comprender el mundo que nos rodea. Este es un objetivo especialmente relacionado con las ciencias sociales, ya que en ellas es a menudo imposible encontrar explicaciones propiamente dichas, aunque a veces se recurre a las de tipo intencional. Se trata de comprender las acciones de los demás seres humanos (o de uno mismo) buscando el sentido que tienen para quien las lleva a cabo, es decir, implica ponerse en el lugar de otro.
- Aplicar los conocimientos obtenidos. Aunque no es un objetivo primordial de la ciencia, constituye una de sus implicaciones más importantes. Se articula a través de la ciencia aplicada y la tecnología, y en el caso de las ciencias sociales puede tener un componente de control o manipulación de los individuos. En cualquiera de los casos, las aplicaciones científicas pueden ser tanto beneficiosas como perjudiciales para el ser humano, pero esa valoración trasciende el ámbito de la propia ciencia.



Licencia: dominio público. Fuente: Pixabay.

## HISTORIAS DE LA CIENCIA. Eddington y las dos mesas

A propósito de la distinción entre conocimiento científico y conocimiento ordinario es interesante el siguiente texto del físico británico Arthur Eddington en *La naturaleza del mundo físico*, donde reflexiona sobre dos maneras distintas de conocer los objetos que le rodean:

Me he puesto a escribir estas conferencias y he acercado mis sillas a mis dos mesas. ¡Dos mesas! Sí; hay duplicados de todos los objetos a mi alrededor: dos mesas, dos sillas, dos plumas. [...]

Con una de las dos mesas estoy familiarizado desde mis primeros años. Es un objeto corriente del ámbito que llamo «el mundo». ¿Cómo la describiría? Tiene extensión, es relativamente permanente; tiene color. Sobre todo es sustancial. Por sustancial no quiero decir simplemente que no se hunde cuando me apoyo en ella; quiero decir que está constituida de «sustancia», y con esa palabra quiero evocar una concepción de su naturaleza intrínseca. Es una cosa: no como el espacio, que es mera negación; ni como el tiempo que es... ¡sabe Dios qué! Pero esto no aclarará lo que quiero decir, porque es la característica distintiva de una «cosa» tener esta sustancialidad, y no creo que la sustancialidad pueda describirse mejor que diciendo que es la clase de naturaleza ejemplificada por una mesa corriente. [...]

La mesa nº 2 es mi mesa científica. Es una conocida más reciente y no me siento tan familiar con ella. No pertenece al mundo que mencioné antes, ese mundo que aparece espontáneamente a mi alrededor cuando abro los ojos -aunque no voy a considerar aquí cuánto de él es objetivo y cuánto es subjetivo-. Es parte de un mundo que se ha impuesto a mi atención en modos más tortuosos.

Mi mesa científica es casi toda vacío. Dispersas en ese vacío hay aquí y allá numerosas cargas eléctricas apresurándose a gran velocidad, pero todo su volumen es, en conjunto, menos de una billonésima del volumen de la mesa. A pesar de su extraña construcción, resulta ser una mesa enteramente eficiente. Sostiene mi papel tan satisfactoriamente como mi mesa nº 1; porque cuando dejo el papel sobre ella, las pequeñas partículas le golpean desde abajo, así que el papel se mantiene a un nivel casi estable, como una pelota de bádmin-ton. Si me apoyo sobre mi mesa no la atravesaré; o, para ser estrictamente preciso, la probabilidad de que mi codo científico pase a través de mi mesa científica es tan exageradamente pequeña, que puede despreciarse en la práctica.

Repasando sus propiedades una a una, parece que para propósitos ordinarios no hay ninguna razón para elegir una u otra de las dos mesas; pero si sobrevienen circunstancias anormales, mi mesa científica demuestra sus ventajas. Si la casa se incendia, mi mesa científica se disolverá de manera muy natural en humo científico, mientras que mi mesa familiar atravesará una metamorfosis en su naturaleza sustancial que sólo puedo contemplar como milagrosa.

No hay nada sustancial en mi segunda mesa. Es casi todo espacio vacío, espacio atravesado, es cierto, por campos de fuerza, pero éstos se asignan a la categoría de «influencias», no de «cosas». Incluso a la minúscula parte que no está vacía no la debemos transferir la vieja noción de sustancia. Al diseccionar la materia en cargas eléctricas, hemos viajado lejos de la imagen que dio origen al concepto de sustancia, y el significado de tal concepto, si alguna vez tuvo alguno, se ha perdido por el camino. Toda la tendencia de la visión científica moderna es romper las categorías separadas de «cosas», «influencias», «formas», etc., y sustituirlas por un trasfondo común a toda experiencia. Ya estudiemos un objeto material, un campo magnético, una figura geométrica o un periodo de tiempo, nuestra información científica se resume en medidas, y ni el aparato científico ni el modo de usarlo sugiere que haya nada esencialmente diferente en estos problemas.

Para saber más: Eddington, A. S. (1928), *The Nature of the Physical World*. Cambridge: Cambridge University Press.

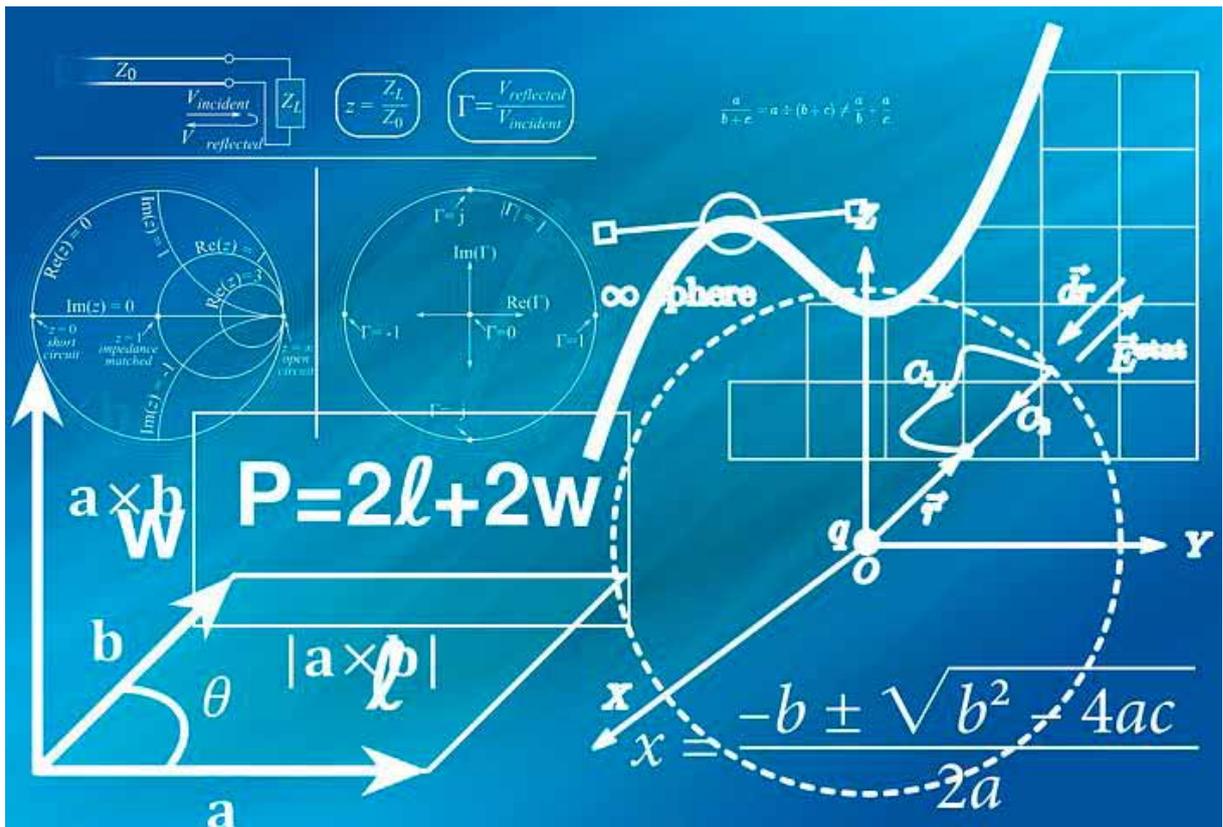


Licencia: dominio público. Fuente: Pixabay.

## 1.2. *Arbor scientiae*: la clasificación de las ciencias

Las diferentes ciencias que se ocupan de distintas áreas de conocimiento pueden clasificarse en tres grupos principales:

- Ciencias formales: no se refieren a hechos observables o sobre los que se pueda experimentar, sino que se ocupan de las consecuencias lógicas que se pueden extraer de un conjunto de axiomas o postulados. No nos dicen nada, en principio, sobre el mundo, sino solamente sobre la estructura lógica que se construye a partir de un conjunto de axiomas que alguien ha inventado y seleccionado. Entre las ciencias formales se encuentran las matemáticas o la lógica, y otras relacionadas con ellas, pero más específicas, como la ciencia computacional teórica o la lingüística teórica.



Licencia: dominio público. Fuente: Pixabay.

- Ciencias naturales: se ocupan del estudio de la naturaleza en muy diversos ámbitos y niveles de complejidad. Se trata de ciencias como la física, la química, la biología, la geología, etc., y son las que están más íntimamente ligadas a la experimentación, es decir, a la posibilidad de reproducir procesos bajo condiciones controladas para obtener mayor cantidad de información y de mayor valor para la elaboración de explicaciones.



Licencia: dominio público. Fuente: Pixabay.

- Ciencias sociales: se ocupan del estudio de la cognición del ser humano, de las relaciones sociales que establece y de su evolución en el tiempo. En otras palabras, estudian el comportamiento humano tanto individual como colectivo. Son ciencias de este tipo la psicología, la sociología, la economía o la historia.



Licencia: dominio público. Fuente: Pixabay.

Dentro de cada grupo no todas las ciencias son igual de fundamentales. Por ejemplo, dentro de las ciencias naturales la biología puede considerarse basada en la bioquímica, que a su vez es una rama de la química, cuyos principios pueden reducirse a los de la física. La jerarquización de las ciencias no siempre es obvia, y a menudo está sujeta a interpretaciones. En cualquier caso, que una ciencia sea más fundamental que otras no implica que deban abandonarse todas las que se sustentan en ella. En el ejemplo anterior, eso significaría que toda la complejidad inherente a los seres vivos podría analizarse directamente con las leyes de la física, sin necesidad de recurrir a las otras ciencias, en particular a la biología. Aunque se puede pensar que, en principio, es posible hacerlo (postura que se denomina reduccionismo), desde el punto de vista práctico resultaría inviable.

Ya en el siglo XIII el filósofo mallorquín Ramon Llull, o Raimundo Lulio, en su obra *Arbor scientiae* (*El árbol de la ciencia*) organizó y jerarquizó diversos grupos de ciencias representándolas como las raíces, el tronco o las ramas de un árbol, aunque evidentemente el conjunto de ciencias que manejaba difería mucho del actual.

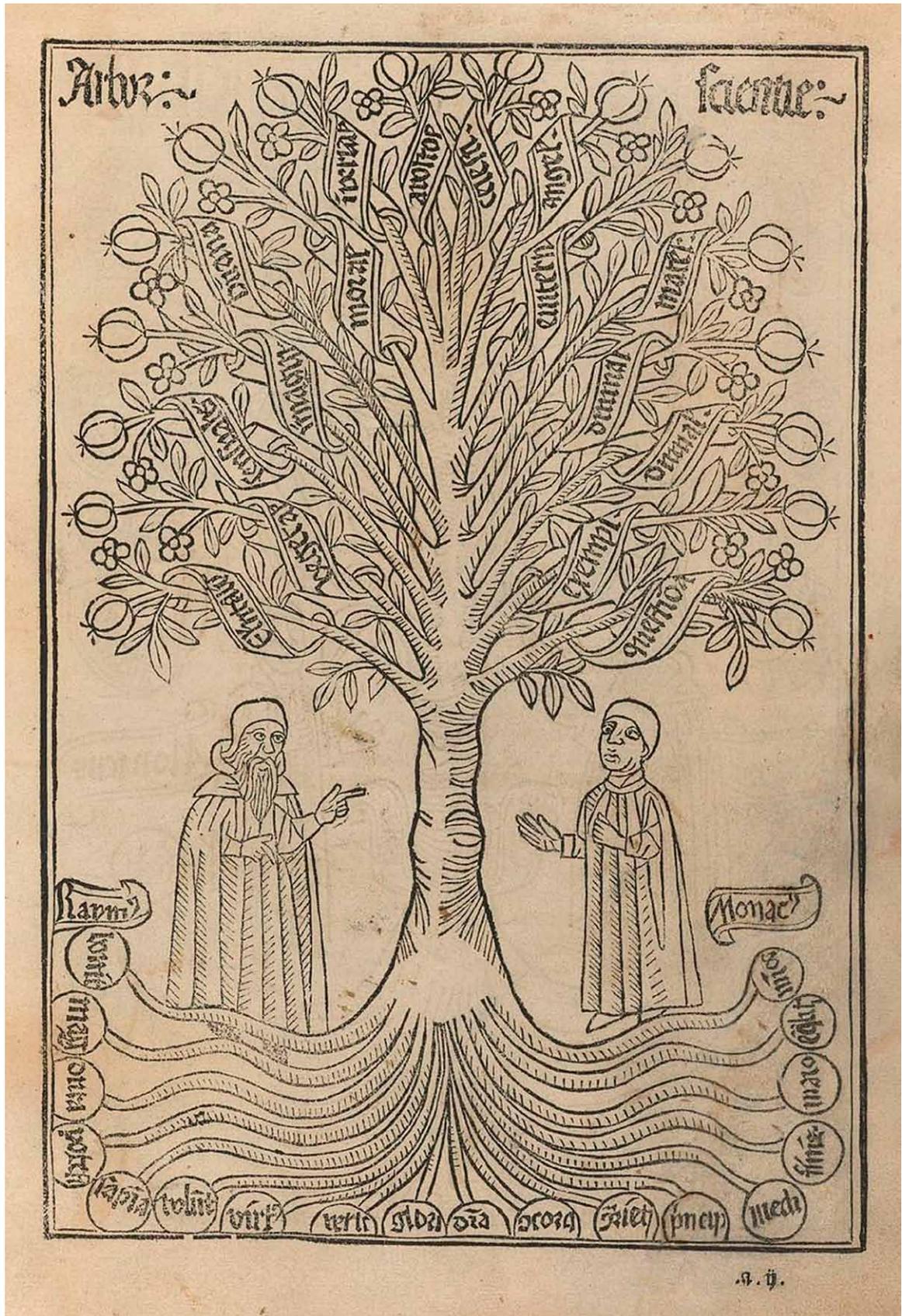


Ilustración de una edición de 1505 de la obra *El árbol de la ciencia*, de Ramon Llull. Licencia: dominio público. Fuente: Wikimedia Commons.

# Capítulo 2.

## Cuestión de protocolo: los procedimientos generales de la ciencia

El conocimiento científico se alcanza mediante una serie de procedimientos generales bien definidos que forman parte de la metodología científica. Entre ellos se encuentran la observación y la experimentación, la división y clasificación, la definición, la inducción o la deducción. Varios de ellos han sido objeto de profunda reflexión ya desde los tiempos de Aristóteles (siglo IV a.C.), destacando la descripción sistemática de la metodología científica que llevó a cabo el filósofo francés René Descartes (1596-1650) en su famoso *Discurso del método*, de 1637; no en vano, el título completo de la obra es *Discurso del método para conducir bien la propia razón y buscar la verdad en las ciencias*.

En este capítulo se describen algunos de los procedimientos más importantes que forman parte del método científico moderno, que se tratará en su conjunto en el siguiente capítulo.

### Conceptos clave:

- División completa, partes excluyentes, criterio de división, análisis, propiedades emergentes, división física, división esencial, relación y clases de equivalencia, clasificación, taxonomía.
- Definición nominal y real, definición esencial (género próximo, diferencia específica), definición descriptiva, definición causal, definición genética, definición operacional, lenguaje natural, imprecisión, polisemia, lenguaje formal, símbolos.
- Inducción, regularidades, ley empírica, hechos brutos, deducción, argumentación, razonamiento, inferencia, premisas, conclusión, enunciado particular (singular), enunciado general (universal), lógica proposicional, lógica de clases, lenguaje formal.
- Explicación, *explanandum*, *explanans*, explicación causal, explicación nomológica, explicación funcional, explicación intencional, proposición o premisa general, proposición o premisa circunstancial, abducción.



En algunas ciencias la experimentación puede ir asociada a dilemas éticos, como la que se lleva a cabo en algunos casos con animales o con humanos. Por ejemplo, en un ensayo clínico para probar la eficacia de un nuevo fármaco, hay que valorar muy cuidadosamente que los beneficios potenciales superen los posibles efectos secundarios, algunos quizás desconocidos. Se suele administrar el fármaco a una serie de individuos, pero a otros solo se les administra un placebo para comprobar que los resultados realmente difieren entre ambos grupos. Esto puede conducir a otro dilema ético: ¿qué ocurre si el fármaco logra curar a los enfermos que lo han tomado, pero el resto de enfermos, en el periodo en que dura el ensayo, empeoran gravemente o mueren porque solo se les administró un placebo? ¿cómo se toma la decisión de quién se arriesga a sufrir efectos secundarios potencialmente graves y quién no, o quién se cura y quién no?

Si bien en ciertas áreas de la biología y de la medicina, o en las ciencias sociales, es difícil recurrir a la experimentación, la tecnología de la instrumentación puede seguir jugando un papel importante no tanto para controlar las condiciones del entorno, sino para la recogida de datos. A ello contribuyen en ciencias sociales, por ejemplo, las tecnologías de la comunicación y de la información, y en particular las asociadas al tratamiento de datos masivos (*big data*).

### AMPLIACIÓN. La ciencia según Descartes y Diderot

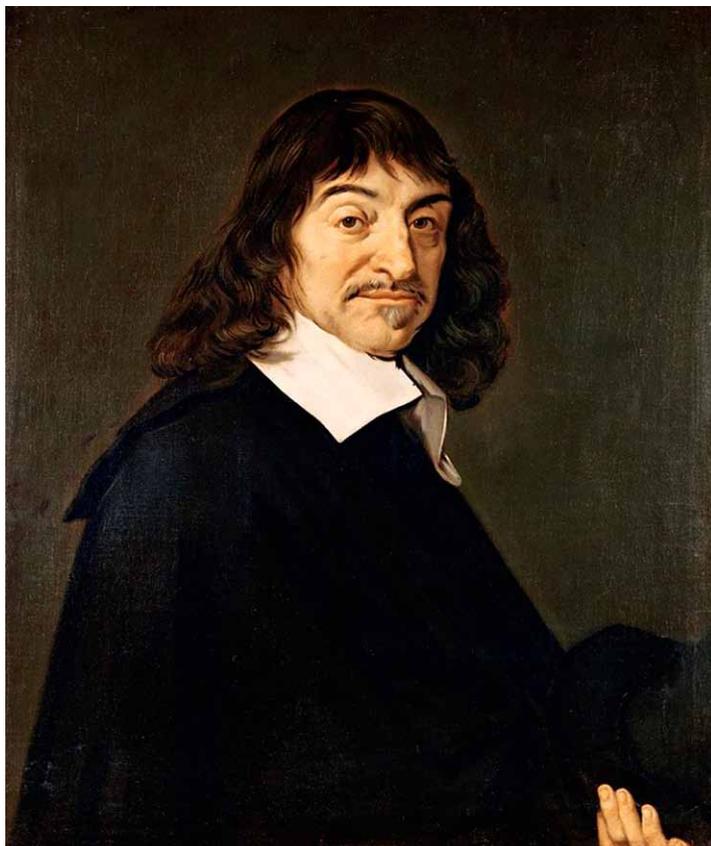
Sin ánimo de volver a romperse la cabeza con el «pienso, luego existo», merece la pena reflexionar sobre el siguiente fragmento del *Discurso del método* de Descartes; se trata de una serie de consejos básicos para la actividad científica, que están muy relacionados con los procedimientos que se tratan en este capítulo:

[...] no admitir como verdadera cosa alguna a menos que se sepa con evidencia que lo es; es decir, evitar cuidadosamente la precipitación y la prevención, y no comprender en mis juicios nada más que lo que se presentase tan clara y distintamente a mi espíritu, que no hubiese ninguna ocasión de ponerlo en duda.

[...] dividir cada una de las dificultades que examinare en tantas partes como sea posible y en cuantas requiriese su mejor solución.

[...] conducir con orden mis pensamientos, empezando por los objetos más simples y más fáciles de conocer, para ir ascendiendo poco a poco, gradualmente, hasta el conocimiento de los más compuestos, e incluso suponiendo un orden entre los que no se preceden naturalmente.

[...] hacer en todo unos recuentos tan integrales y unas revisiones tan generales, que llegase a estar seguro de no omitir nada.



René Descartes. Licencia: dominio público. Fuente: Wikimedia Commons.

Por su parte, el filósofo francés Denis Diderot (1713-1784), en su *Interpretación de la naturaleza*, enumeró de la siguiente manera los procedimientos con los que debe desarrollarse la ciencia:

Poseemos tres medios principales: la observación de la naturaleza, la reflexión y la experiencia. La observación recoge los hechos, la reflexión los combina, la experiencia verifica el resultado de la combinación. Es indispensable que la observación de la naturaleza sea asidua, que la reflexión sea profunda y la experiencia sea exacta.



Denis Diderot. Licencia: dominio público. Fuente: Wikimedia Commons.

## HISTORIAS DE LA CIENCIA. Eddington y el eclipse que probó la relatividad

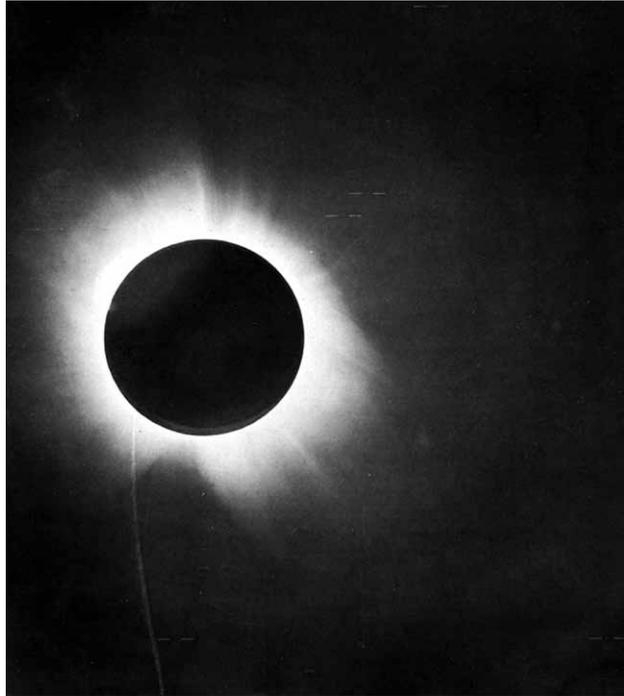


Arthur Eddington. Licencia: dominio público. Fuente: Wikimedia Commons.

Arthur Eddington (1882-1944) fue un físico británico que realizó avances especialmente importantes en astrofísica. En particular, propuso que la energía que producían las estrellas tenía su origen en el núcleo de los átomos; entre los diversos procesos nucleares que manejó, apostó por la fusión nuclear: la unión de dos núcleos ligeros para formar uno más pesado, liberando energía en el proceso. Este mecanismo es el aceptado actualmente como origen principal de la energía en el interior de las estrellas y solo se puede entender en el marco de la teoría especial de la relatividad del físico alemán Albert Einstein (1879-1955) a través de la equivalencia entre masa y energía. La fórmula asociada a esta ley,  $E = mc^2$ , es quizá la más famosa de la ciencia. Precisamente Eddington destacó también por ser un gran divulgador para el público general de las recientemente propuestas teorías de la relatividad, lo que constituye una vertiente igualmente importante de la actividad científica: la difusión y explicación de los avances que se producen. Esta es una tarea especialmente complicada en el caso de la teoría de la relatividad, que es muy abstracta y poco intuitiva.

El propio Einstein también contribuyó a su divulgación con obras destinadas al público general (por ejemplo, *Sobre la relatividad especial y general*). Eddington publicó varias obras sobre los fundamentos de la ciencia, como *Nuevos caminos de la ciencia*, *Filosofía de la ciencia física* o *La naturaleza del mundo físico* (que incluye el relato sobre las dos mesas reproducido parcialmente en el capítulo anterior).

Eddington encabezó en mayo de 1919 una expedición a la isla de Príncipe en el Golfo de Guinea para observar un eclipse de Sol, para comprobar una de las predicciones de la teoría de la relatividad general que Einstein había publicado cuatro años antes. En el marco de esta teoría los cuerpos con grandes masas deforman el espacio a su alrededor; en consecuencia, la luz de las estrellas situadas detrás de uno de esos cuerpos, como el Sol, sigue una trayectoria curva para llegar a la Tierra, en lugar de recta, lo que se traduce en que su posición aparente en el cielo se ve un poco desplazada respecto a la real.



Fotografía de Eddington del eclipse solar del 1919. Licencia: dominio público. Fuente: Wikimedia Commons.

El problema con el Sol es que su gran brillo impide que veamos las estrellas del firmamento a su alrededor, por lo que no se puede saber si cambian su posición aparente cuando el Sol se sitúa próximo a ellas en el cielo. Para comprobar esta predicción teórica, habría que modificar las condiciones de la observación a través de un experimento que encendiese y apagase el Sol y lo desplazara en la bóveda celeste según la necesidad. Evidentemente, tal manipulación no es posible, así que el experimento queda descartado, pero siempre queda el recurso a la paciencia. Si se espera a que ocurra un eclipse total de Sol, y se viaja al lugar de la Tierra donde sea visible, será posible observar la posición aparente de las estrellas a su alrededor, porque la intensa luz del disco solar estará apantallada por la Luna interpuesta.

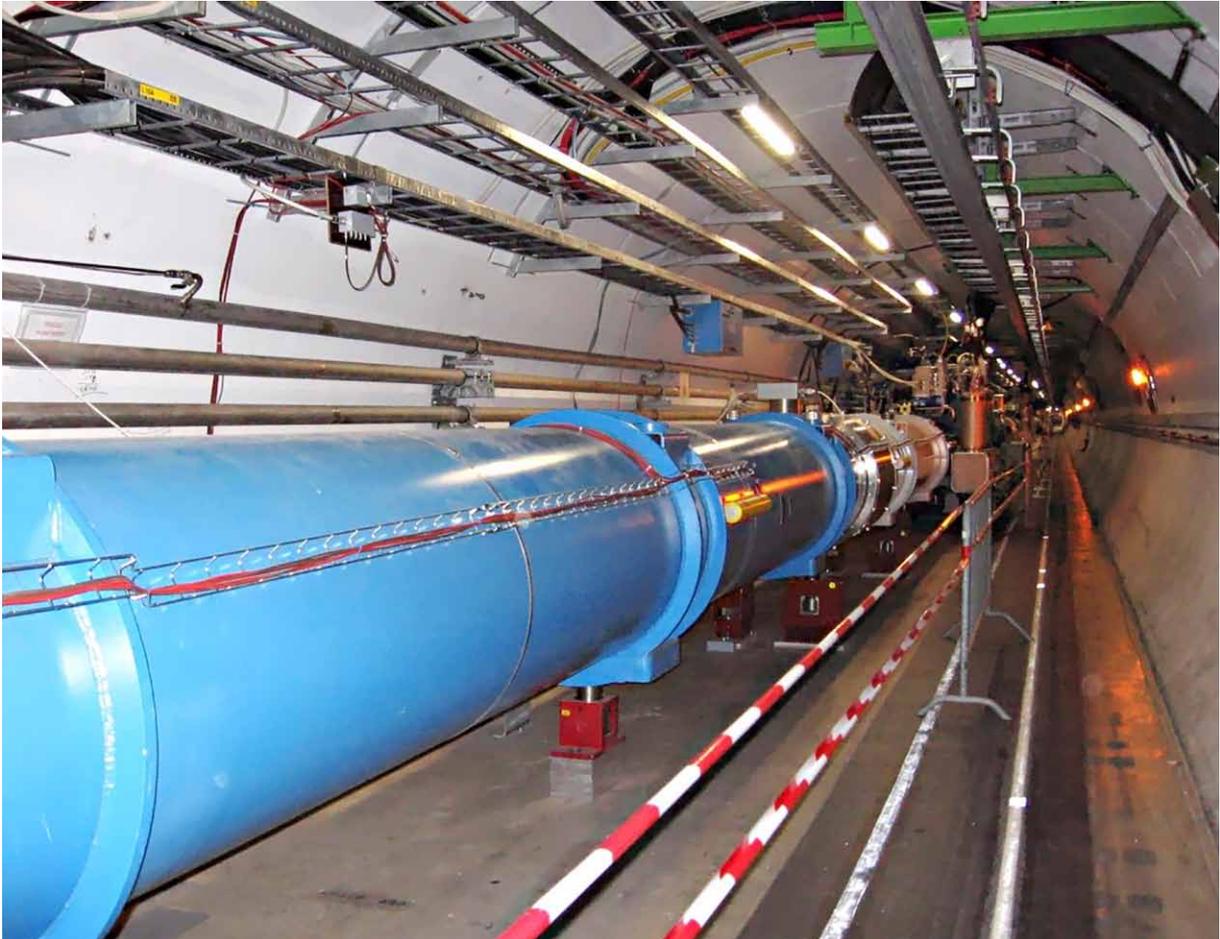
Así pues, Eddington y su equipo compensaron la imposibilidad de manipulación experimental mediante una buena dosis de ingenio y de paciencia, que resultó en la observación necesaria para comprobar la teoría. La conclusión, por cierto, fue que efectivamente la posición de las estrellas cambia al curvar el Sol la trayectoria de la luz que emiten. La teoría de la relatividad general de Einstein cumplía así uno de los más importantes objetivos de la ciencia, la predicción de fenómenos. Gracias al contraste empírico de hipótesis como el llevado a cabo con esta observación, la teoría de la relatividad fue aceptada por la comunidad científica.

## HISTORIAS DE LA CIENCIA. Experimentación al límite en el LHC

Un ejemplo de instalación moderna para la experimentación es el acelerador de partículas LHC (por las siglas en inglés de Gran Colisionador de Hadrones) situado en el CERN, la Organización Europea para la Investigación Nuclear en Ginebra (Suiza). El objetivo de esta instalación internacional es acelerar partículas subatómicas a grandes energías y hacerlas colisionar entre sí, para desvelar los secretos más profundos de la física a través de los fragmentos formados tras los choques.

En el LHC se controlan las condiciones físicas del entorno mediante tecnología puntera para lograr las temperaturas más bajas (inferiores incluso a la del espacio interestelar) y el mayor grado de vacío del universo. La energía que alcanzan las partículas, solo superada por algunos rayos cósmicos, se

consigue acelerándolas a través de un túnel subterráneo circular de 27 km de longitud, consumiendo al año tanta electricidad como medio millón de hogares españoles. En cuanto a la recogida de datos, se emplean detectores extraordinariamente sensibles y precisos, que generan una ingente cantidad de información (unos 50 mil millones de *megabytes* al año), para cuyo manejo y almacenamiento se requiere un *hardware* y *software* especialmente diseñados.



Túnel del LHC en el CERN. Licencia: Creative Commons Attribution-Share Alike 3.0 Unported. Autor: Julian Herzog.

## 2.2. Divide y vencerás. División y clasificación en la ciencia

La división consiste en separar un todo en sus partes, y es un procedimiento esencial en ciencia porque permite analizar un conjunto complejo de manera más simple a través de sus componentes; el propio término «análisis» significa etimológicamente «disolución de un conjunto en sus partes». Esto no implica que el estudio de las partes por separado sea suficiente para entender el conjunto, porque normalmente esas partes interactúan de manera compleja dando lugar a propiedades emergentes. Así, un conocimiento de todos los órganos en los que podemos dividir el cuerpo humano, por muy exhaustivo que sea, no es un conocimiento completo de nuestro organismo en su conjunto.

Una división adecuada para la tarea científica tiene que reunir las siguientes características:

- Tiene que ser completa, es decir, la unión de todas las partes tiene que ser igual al todo. Por ejemplo, no sería adecuado dividir el átomo en electrones y protones, porque faltarían los neutrones para realizar la división completa.

- Tiene que estar formada por partes mutuamente excluyentes, es decir, unas partes no pueden estar incluidas total o parcialmente en otras. Por ejemplo, no sería apropiado dividir el átomo en corteza electrónica, núcleo, protones y neutrones, porque estas dos últimas partículas ya forman parte del núcleo; sí sería adecuado dividirlo en corteza y núcleo, que son distintos entre sí, o en electrones, protones y neutrones, que también son componentes distintos entre sí, y además cualquiera de esas dos divisiones sería completa.
- Todas las partes tienen que surgir de un mismo criterio de división, esto es, no se puede cambiar la regla con la que se forman las partes antes de que la división esté completa. Así, no tendría sentido dividir el átomo en partículas con carga negativa, partículas con carga positiva, partículas ligeras y partículas pesadas, porque se ha cambiado el criterio de división, pasando de ser la carga eléctrica de las partes a ser su masa (además de que las partes resultantes no son mutuamente excluyentes).

La clasificación es otro importante procedimiento científico relacionado con la división, que consiste en ordenar jerárquica y sistemáticamente un conjunto de objetos. El todo del que forman parte no es tanto un objeto sino una entidad más abstracta o teórica, como por ejemplo «las partículas subatómicas»; otro ejemplo es «las plantas terrestres», que se pueden clasificar en briofitas (no vasculares) y traqueofitas (vasculares), estas últimas a su vez en pteridofitas (sin semillas) y espermatofitas (con semillas), y estas últimas en angiospermas (con flor) y gimnospermas (sin flor). Como se ve en este ejemplo, las clasificaciones sistemáticas y jerarquizadas (donde cada parte se divide en otras, y estas a su vez en otras, y así sucesivamente) son muy importantes en ciencias como la biología. La taxonomía es la disciplina que, dentro de esas ciencias, se ocupa de los principios y métodos de la clasificación. Una de las clasificaciones taxonómicas más famosas es la que llevó a cabo el botánico y zoólogo sueco Carl Linnaeus (1707-1778) para plantas y animales, que sirvió como base para las clasificaciones actuales.

Tabla de clasificación del reino animal del *Systema Naturae* de Linneo. Licencia: dominio público. Fuente: Wikimedia Commons.

A pesar de la importancia de la clasificación, ninguna ciencia puede limitarse simplemente a «coleccionar mariposas», es decir, a clasificar objetos o especímenes conforme van siendo descubiertos. Si bien se trata de un primer paso necesario para poner orden en las observaciones, la ciencia tiene que extraer conclusiones a partir de esa clasificación, y ofrecer una explicación de por qué se produce de esa manera; es decir, tiene que formular un marco teórico acerca de su campo de estudio.

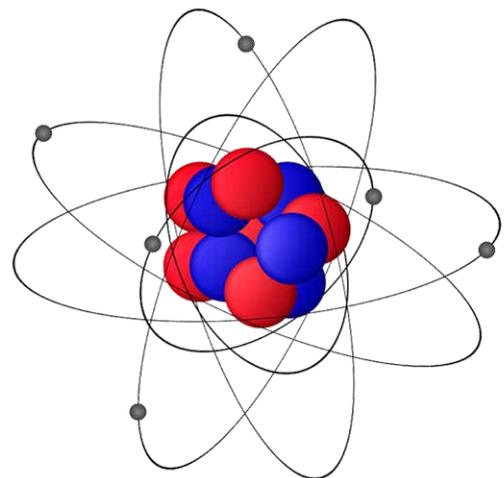


Licencia: dominio público. Fuente: Pixabay.

### ▶ AMPLIACIÓN. División física y división esencial

Cabe distinguir entre la división física de los objetos y la división esencial. La primera consiste en identificar todas las partes que forman físicamente un objeto. Por ejemplo, en el caso del átomo de oxígeno más abundante, su división física en partículas constituyentes daría como resultado 8 electrones, 8 protones y 8 neutrones. En la división esencial, en cambio, solo interesan los tipos distintos de partes, no el número de cada una de ellas. Así, el átomo de oxígeno, o de cualquier otro elemento, se puede dividir esencialmente en electrones, protones y neutrones.

En ciencia son importantes ambas divisiones, la física en el plano más empírico y la esencial en el plano más teórico, pero ambas están relacionadas entre sí, como veremos a continuación. Entre las partes de una



Licencia: dominio público. Fuente: Pixabay.

división física se pueden establecer relaciones de equivalencia, que son las que fijan el criterio de división. Las relaciones de equivalencia se caracterizan por ser reflexivas (un elemento está relacionado consigo mismo), simétricas (si un elemento está relacionado con otro, este segundo está entonces relacionado con el primero) y transitivas (si un elemento está relacionado con otro, y este lo está con un tercero, el primero está entonces relacionado con el tercero). Por ejemplo, entre las 24 partes en que se divide por completo el átomo de oxígeno (en partículas subatómicas) se puede establecer la relación de equivalencia «tener la misma carga eléctrica que» (se comprueba inmediatamente que esta relación cumple las tres propiedades indicadas antes); al aplicarla, las partes quedan clasificadas en tres grupos: con carga negativa, con carga positiva y con carga neutra. Con este procedimiento, la división física del átomo se ha transformado en una división esencial según el criterio de la carga eléctrica, resultando tres partes, en forma de tres clases de equivalencia, correspondientes a electrón, protón y neutrón. La división esencial, formada por clases de elementos que comparten propiedades, es además el fundamento del procedimiento de clasificación.

### ► HISTORIAS DE LA CIENCIA. Los átomos, invisibles e ¿indivisibles?

Hemos estado usando el átomo como ejemplo para ilustrar el procedimiento de división en ciencia, pero ¿hasta dónde puede dividirse físicamente la materia que nos rodea? Aunque en la antigua Grecia no existía capacidad experimental para encontrar ese límite, el filósofo Demócrito (460-370 a.C.), basándose en ideas de sus maestros, propuso que la materia estaba formada por unas entidades eternas, homogéneas, incompresibles, invisibles e indivisibles. Esta última característica es la que da origen al nombre de esas unidades mínimas, los átomos, que proviene del griego *atomos*, que significa «que no puede cortarse o dividirse». Las primeras evidencias empíricas de los átomos no aparecieron hasta principios del siglo XIX, cuando el químico inglés John Dalton (1766-1844) dedujo su existencia a partir del hecho de que las sustancias químicas se combinaban para formar compuestos según proporciones de masa sencillas que variaban de unos compuestos a otros.

Hoy día sabemos que los átomos en realidad sí son divisibles, y poseen una estructura interna formada por partículas, que reciben el poco sorprendente calificativo de subatómicas. Las evidencias de la existencia de estos componentes comenzaron a aparecer a finales del siglo XIX. Los descubrimientos de físicos como el francés Henri Becquerel (1852-1908) y el neozelandés Ernest Rutherford (1871-1937) sobre las radiaciones provenientes del interior de los átomos y su interacción con la materia permitieron deducir que el átomo podía descomponerse en porciones más pequeñas y que contenía un núcleo con carga positiva en el que se concentraba la mayor parte de su masa; las unidades más simples del núcleo que pueden encontrarse libres son el protón y el neutrón, este último descubierto en 1932 por el físico inglés James Chadwick (1891-1974). Alrededor del núcleo y muy alejado de él se encuentra la corteza, formada por electrones que orbitan a su alrededor; estas partículas fueron descubiertas también a finales del siglo XIX por el físico inglés Joseph Thomson (1856-1940) al estudiar los rayos catódicos.

El electrón, el protón y el neutrón son las partes resultantes de una división esencial del átomo. Pero los diferentes elementos químicos se distinguen entre sí por el número de electrones presentes en la corteza de sus átomos, que a su vez depende del número de protones en el núcleo, según sabemos hoy día. Por tanto, las propiedades químicas de la materia solo se pueden entender a través de la división física de los átomos, es decir, considerando el número de partículas de cada tipo que contienen.

Cuando se clasifican las partes de un conjunto, este no solo adquiere un aspecto más simple, sino que además surgen evidencias sobre la naturaleza y propiedades de esas partes. El químico ruso Dimitri Mendeleev (1834-1907) clasificó los elementos según sus propiedades químicas observadas (en esa época se desconocía que la estructura y composición de los átomos es el origen de esas propiedades); al ordenarlos en una tabla según esa clasificación, que es el origen de la tabla periódica moderna, se dio cuenta de que quedaban huecos, como los de un puzzle cuando

faltan algunas piezas. Mendeleev fue incluso capaz de predecir las propiedades de esos elementos desconocidos a partir de las propiedades de los que estaban a su alrededor en la tabla. Les asignó nombres como eka-boro, eka-aluminio, eka-silicio o eka-manganeso, que contienen el nombre de un elemento contiguo en la tabla junto con el prefijo sánscrito «eka», que significa «uno». Hoy día esos elementos, ya descubiertos, se denominan escandio, galio, germanio y tecnecio, respectivamente. En el caso del tecnecio, sus átomos no existen en la naturaleza en cantidades apreciables (porque son radiactivos) y debe ser creado en laboratorio para ser estudiado (de ahí su nombre, que significa artificial en griego). ¡Se descubrió más de 60 años después de haber sido predicho por Mendeleev!

Reihen	Gruppe I. — R'O	Gruppe II. — RO	Gruppe III. — R'O <sup>3</sup>	Gruppe IV. RH <sup>4</sup> RO <sup>4</sup>	Gruppe V. RH <sup>5</sup> R'O <sup>5</sup>	Gruppe VI. RH <sup>6</sup> RO <sup>6</sup>	Gruppe VII. RH R'O <sup>7</sup>	Gruppe VIII. — RO <sup>8</sup>
1	H=1							
2	Li=7	Be=9,4	B=11	C=12	N=14	O=16	F=19	
3	Na=23	Mg=24	Al=27,3	Si=28	P=31	S=32	Cl=35,5	
4	K=39	Ca=40	—=44	Ti=48	V=51	Cr=52	Mn=55	Fe=56, Co=59, Ni=59, Cu=63.
5	(Cu=63)	Zn=65	—=68	—=72	As=75	So=78	Br=80	
6	Rb=86	Sr=87	?Yt=88	Zr=90	Nb=94	Mo=96	—=100	Ru=104, Rh=104, Pd=106, Ag=108.
7	(Ag=108)	Cd=112	In=113	Sn=118	Sb=122	Te=125	J=127	
8	Cs=133	Ba=137	?Di=138	?Ce=140	—	—	—	— — — —
9	(—)	—	—	—	—	—	—	
10	—	—	?Er=178	?La=180	Ta=182	W=184	—	Os=195, Ir=197, Pt=198, Au=199.
11	(Au=199)	Hg=200	Tl=204	Pb=207	Bi=208	—	—	
12	—	—	—	Th=231	—	U=240	—	— — — —

Tabla periódica de Mendeleev de 1871. Licencia: dominio público. Fuente: Wikimedia Commons.

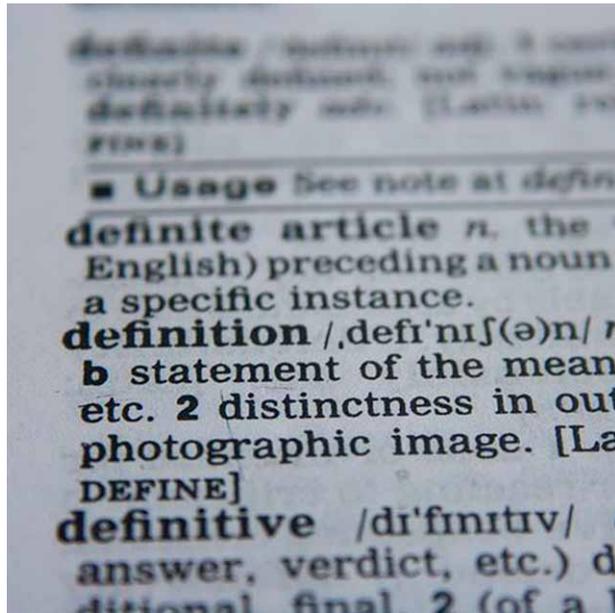
Una historia similar ha ocurrido en la física del siglo xx con las partículas, tanto las que forman parte de los átomos como con otras muchas. Se propusieron ingeniosos patrones de clasificación del numeroso conjunto de partículas descubiertas, a partir de los cuales los físicos establecieron teorías que explicaban por qué se ordenaban de esa manera. Como ocurrió con los elementos químicos, las tablas de clasificación de las partículas presentaban algunos huecos, que condujeron a la predicción de partículas desconocidas hasta el momento, con unas propiedades muy concretas. Sin conocer esas propiedades habría sido muy difícil detectarlas experimentalmente, porque habría sido como buscar agujas en un pajar. Poco a poco todas las partículas predichas fueron descubiertas en aceleradores de partículas, como el LHC del CERN.

### 2.3. Definir los límites: definición de conceptos y terminología científica

Otro de los procedimientos más importantes en ciencia es el de la definición precisa de conceptos. Etimológicamente, definir significa «poner límites», ya que proviene del latín *finis*, que es límite o frontera. Precisamente esto que acabamos de hacer es una definición, en concreto de tipo nominal, porque se aclara el significado de una palabra recurriendo, en este caso, a su etimología (también podrían haberse usado sinónimos de la palabra en cuestión).

En ciencia no interesan tanto esas definiciones nominales, sino las definiciones «reales», que transmiten las características y naturaleza de las cosas, que pueden ser a su vez de varios tipos.

La habitual y deseable en ciencia es la definición esencial, que contiene dos elementos: el género próximo y la diferencia específica. El primero se refiere a una característica que comparte con otros conceptos de una misma clase, mientras que la segunda se refiere a una característica que lo distingue del resto de conceptos incluidos en esa clase. Como se puede deducir, este tipo de definición está muy relacionada con la división en clases de equivalencia que vimos en el apartado anterior. Por ejemplo, la definición esencial de cuadrado es «polígono de cuatro lados», donde «polígono» es el género próximo y transmite que se trata de una figura geométrica plana formada por segmentos rectos consecutivos que encierran una región del plano (como lo son otras muchas figuras: triángulos, pentágonos, etc.); y «de cuatro lados» es la diferencia específica, que lo distingue del resto de polígonos.



Licencia: dominio público. Fuente: Pixabay.

En ciencia, las definiciones deben especificar el concepto al que se refieren, pero a menudo lo hacen de manera tan precisa que no existe en el lenguaje común un término específico para referirse a ello, y es necesario inventarlo. La terminología científica es uno de los principales obstáculos para transmitir la ciencia (otro muy importante es el lenguaje formal que suele emplearse, el de las matemáticas). Si bien el uso de vocabulario científico es imprescindible en la ciencia propiamente dicha, es decir, en la investigación, hay que hacer lo posible para evitarlo en la divulgación, e introducirlo muy gradualmente en la docencia.

► **AMPLIACIÓN. Otras definiciones y el lenguaje natural**

Cuando no se cuenta con una definición esencial se puede recurrir a otros tipos, como la definición descriptiva, que contiene una serie de características del objeto distintas del género próximo o la diferencia específica, la definición causal, que alude a la causa que produce el objeto, la definición genética, que expresa cómo se crea o construye el objeto, o la definición operacional, que determina cómo medirlo experimentalmente. Esta última es muy importante en ciencia, porque asocia de manera unívoca el concepto que se está manejando en una investigación dada con el procedimiento exacto para medirlo, de manera que cualquier científico pueda hacerlo de la misma manera y se eviten confusiones. Por ejemplo, una definición operacional del concepto «masa» en física es «lo que marca una balanza cuando se sitúa en ella un cuerpo»; con esta definición todo el mundo sabe cómo determinar la masa, y se evitan complicaciones asociadas a su definición esencial o descriptiva; esta última se emplea en las leyes y teorías, pero es mucho más abstracta.

Evidentemente, una definición debe aclarar el significado del término que se quiere definir; esto implica, entre otras cosas, que el propio término no puede formar parte de la definición. Además, lo que se quiere definir y la definición asociada deben ser equivalentes e intercambiables; esto quiere decir que la definición no puede ser ni más general ni más restringida que el concepto al que se aplica. Por ejemplo, una definición adecuada de «cuchillo» no podría ser «instrumento para cortar formado por una hoja de metal», porque hay muchos objetos distintos con esas propiedades, y tampoco podría ser «instrumento para cortar formado por una hoja de metal de 20 cm y un mango de color negro», porque no todo lo que se entiende por cuchillo reúne esas características. La definición de la RAE, por

cierto, es «instrumento para cortar formado por una hoja de metal de un corte solo y con mango», que parece más adecuada.

La ciencia hace uso de una terminología específica y de unas definiciones exactas para evitar la imprecisión y la polisemia que a menudo aparecen en el lenguaje natural, que es el que usamos en el día a día. La imprecisión impide diferenciar cosas parecidas entre sí o el grado de las cualidades; por ejemplo, decimos habitualmente cosas como «he visto pasar un pájaro» o «esa montaña es muy alta», mientras que en el ámbito científico se distinguiría una especie concreta de pájaro (en zoología) o se daría la altitud de la montaña en metros (en geografía). La polisemia, por su parte, se refiere a la multiplicidad de significados distintos que puede tener una misma palabra o frase. Por ejemplo, se puede comparar la variedad de significados de la palabra «fuerza» según la RAE con el muy específico que tiene en física (que la RAE recoge en la acepción 14). En realidad, en ciencias como la física, las definiciones son tan precisas que a veces conviene expresarlas directamente con el lenguaje formal de las matemáticas.

## 2.4. El color de los cisnes: la inducción

La inducción, o método inductivo, consiste en recopilar un conjunto de datos mediante la observación o la experimentación y extraer de ellos relaciones que se repiten sistemáticamente, es decir, regularidades o patrones. Estas regularidades se recogen en una ley empírica universal, que se supone válida no solo para ese conjunto de datos, sino para los recogidos en cualquier otra observación del mismo fenómeno en el pasado o en el futuro.

La inducción permite pasar de lo particular, que son los hechos observados, a lo general o universal, que es la ley empírica que contiene las regularidades extraídas y que se supone válida para todos los hechos a los que se refiere, no solo los observados previamente.

La inducción presenta un problema, y es que no existe un razonamiento lógico válido que la sustente. Aunque el enunciado «*algunos* hechos del tipo X cumplen la relación Y» sea verdadero (y lo es porque los hechos a los que se refiere han sido observados cumpliendo esa relación), no existe ninguna garantía de que el enunciado «*todos* los hechos del tipo X cumplen la relación Y» también sea verdadero. Desde luego, difícilmente podrán ser observados todos los hechos de ese tipo que se puedan producir, que a menudo son infinitos.

Además del problema de la justificación lógica, también es discutible la posibilidad de registrar hechos «brutos» para extraer regularidades a partir de ellos. La mera selección de los hechos que se deben observar y registrar, la manera de hacerlo, la comprensión del funcionamiento de los instrumentos usados o de nuestros sentidos, etc., ya implican una cierta teoría sobre el mundo, previa a la generada por la inducción; esto pone en duda la posibilidad de obtener datos experimentales independientes, ajenos a cualquier constructo teórico.

### HISTORIAS DE LA CIENCIA. Cisnes negros

Un ejemplo clásico en el que el procedimiento inductivo falla es el del color de los cisnes. Históricamente, cada vez que se veía un cisne, este era invariablemente de color blanco, asociación de hechos que se había observado en innumerables ocasiones. Evidentemente, no se tenía constancia de todos los cisnes del mundo, o al menos era imposible saberlo, pero eran tantos los ejemplos (particulares) de esa relación, que el procedimiento inductivo permitía establecer la ley empírica universal «todos los cisnes son blancos». Aunque no hubiese una justificación lógica detrás, la ley parecía ser plenamente válida; tanto es así que existe una expresión latina, *rara avis*, que significa algo extraordinario o inusual, cuyo origen es un verso del autor romano Juvenal, «*rara avis in terris nigroque simillima cygno*»: un ave rara en la tierra, como un cisne negro.



Licencia: dominio público. Fuente: Pixabay.

Llegados a este punto parece fácil predecir cómo termina esta historia. En 1697, una expedición europea descubrió en Australia la existencia de cisnes negros (*Cygnus atratus*), endémicos de aquel continente.

El filósofo británico Bertrand Russell (1872-1970) aportó este otro ejemplo de aplicación fallida de la inducción en *Los problemas de la filosofía*, esta vez llevada a cabo por animales y con un final más trágico:

Este género de asociación no se limita al hombre; es también muy fuerte en los animales. [...] Los animales domésticos esperan su alimento cuando ven la persona que habitualmente se lo da. Sabemos que todas estas expectativas, más bien burdas, de uniformidad, están sujetas a error. El hombre que daba de comer todos los días al pollo, a la postre le tuerce el cuello, demostrando con ello que le hubiesen sido útiles al pollo opiniones más afinadas sobre la uniformidad de la naturaleza.

Una variante de este ejemplo se popularizó como la historia del pavo inductivista: un pavo, al ver que le daban comida un día tras otro a las nueve de la mañana, extrajo la regla general de que recibiría su alimento en el futuro todos los días a esa misma hora, hasta que llegó el Día de Acción de Gracias y, esperando su alimento, lo que recibió fue un corte en el cuello.

## 2.5. Lógica en estado puro: la deducción

Una deducción lógica, también llamada argumentación, razonamiento o inferencia, consiste en la extracción de conclusiones a partir de ciertos enunciados, denominados premisas. Si las premisas son verdaderas, entonces las conclusiones deducidas son necesariamente verdaderas, y ello se debe a la estructura lógica de la argumentación, no al significado concreto que tengan las premisas.

Las argumentaciones pueden establecerse de manera sencilla, o al menos de forma inequívoca, recurriendo al lenguaje de la lógica, que es formal y simbólico. Formal quiere decir que se ocupa solamente de las relaciones entre sus símbolos, y no del significado que tienen o de su uso. Simbólico

implica que esos símbolos carecen de significado y características de uso concretos, pudiendo representar cualquier concepto. Un lenguaje de este tipo es por tanto adecuado para analizar deducciones, que se basan en la estructura y relaciones entre los elementos y son independientes de sus significados.

Un ejemplo de razonamiento lógico es el siguiente:

Premisa 1: Si las manzanas son frutas, entonces son comestibles.

Premisa 2: Las manzanas son frutas.

Conclusión: Las manzanas son comestibles.

En la premisa 1 se establece como verdadera una relación condicional y la premisa 2 dice que la condición de esa relación se da en la realidad, es decir, también es verdadera. Como conclusión, se deduce que la consecuencia de la relación condicional también es verdadera.



Licencia: dominio público. Fuente: Pixabay.

Mediante un lenguaje formalizado y parcialmente simbólico, se puede expresar este razonamiento de la siguiente manera:

Premisa 1: Si M son F, entonces M son C.

Premisa 2: M son F.

Conclusión: M son C.

Aquí, M sustituye a «manzanas», F a «frutas» y C a «comestibles». Se puede ir un paso más allá y expresar el razonamiento de forma plenamente simbólica de la siguiente manera:

Premisa 1:  $p \rightarrow q$

Premisa 2:  $p$

Conclusión:  $q$

Aquí,  $p$  reemplaza el enunciado completo «las manzanas son frutas» (o «M son F») y  $q$  reemplaza el enunciado «las manzanas son comestibles» (o «M son C»). El símbolo de la flecha ( $\rightarrow$ ) relaciona una condición con una consecuencia ( $p \rightarrow q$  significa «si se da la condición  $p$ , entonces ocurre la consecuencia  $q$ »).

Los símbolos empleados pueden tomar significados muy diferentes al del ejemplo inicial, y aun así el razonamiento descrito sigue siendo correcto desde el punto de vista formal, aunque no tiene por qué serlo atendiendo a su significado; pueden darse de hecho enunciados claramente falsos o absurdos. Así ocurriría, por ejemplo, si en el razonamiento anterior el símbolo M tomara el significado de «piedras»: la premisa 2 sería falsa, y por tanto también lo sería la conclusión deducida. La lógica solo se ocupa de la validez formal de los razonamientos, no de la veracidad de las premisas, que ha de determinarse con la observación o la experimentación.

**AMPLIACIÓN. Proposiciones y clases, reglas y leyes**

La manera en que se han expresado los razonamientos anteriores forma parte de la lógica proposicional, porque sus elementos fundamentales son proposiciones, es decir, pensamientos en los que se afirma algo y que se expresan mediante enunciados descriptivos, como por ejemplo «las manzanas son comestibles». Existe otra manera de formular el mismo razonamiento, que es la siguiente:

- Premisa 1: Todas las frutas son comestibles.
- Premisa 2: Las manzanas son frutas.
- Conclusión: Las manzanas son comestibles.

De manera simbólica, se puede escribir como:

- Premisa 1: Todo F es C (o también  $F \subset C$ ).
- Premisa 2: Todo M es F (o también  $M \subset F$ ).
- Conclusión: Todo M es C (o también  $M \subset C$ ).

El símbolo « $\subset$ » expresa una relación de inclusión de una clase en otra, que se emplea también en teoría de conjuntos. Esta forma de expresar los razonamientos es propia de la lógica de clases, donde las clases son conjuntos de elementos que comparten alguna propiedad, como vimos al hablar de los procedimientos de división y clasificación en ciencia.

Por su parte, los razonamientos lógicos pueden darse o bien en forma de reglas de inferencia lógica (conjunto de instrucciones para transformar unas proposiciones en otras) o bien en forma de leyes lógicas (relaciones entre proposiciones que siempre resultan verdaderas), pero ambas formas son equivalentes.

Por ejemplo, un razonamiento lógico típico es el modus ponendo ponens (que es el de los ejemplos anteriores), cuya aplicación nos dice que si tenemos una relación condicional verdadera (premisa 1:  $p \rightarrow q$ ), y su condición es verdadera (premisa 2: p), entonces la consecuencia es verdadera (conclusión: q). Así, la regla lógica indica que de las premisas puede extraerse esa conclusión, y que incluso pueden ser remplazadas por ella, porque siempre que sean ciertas las primeras, lo será la última. En forma de ley lógica, ese mismo razonamiento se expresaría como un enunciado que siempre es verdadero (se dice «tautológico»): siempre es verdad que si p implica q y ocurre p, entonces ocurre q; simbólicamente se escribe como:  $[(p \rightarrow q) \wedge p] \rightarrow q$ , donde la cuña ( $\wedge$ ) indica la conjunción copulativa «y».

El número de reglas de inferencia lógica, o su equivalente en forma de leyes lógicas, es infinito, pero se puede dar una lista de la más simples, importantes y habituales, como la que aparece en la tabla siguiente, precedida de la definición de los principales conectores lógicos y de los tres principios fundamentales:

DEFINICIÓN DE CONECTORES LÓGICOS			
Nombre	Símbolo	Significado	Condición de verdad
Negación	$\neg$	no p	Si p es falsa
Conjunción	$\wedge$	p y q	Si p y q son ambas verdaderas
Disyunción	$\vee$	p o q	Siempre, excepto si p y q son ambas falsas
Condicional	$\rightarrow$	p si q	Siempre, excepto si p es verdadera pero q no lo es
Bicondicional	$\leftrightarrow$	p si y solo si q	Si p y q son ambas verdaderas o ambas falsas

PRINCIPIOS FUNDAMENTALES	
Identidad	$p \rightarrow p$
No contradicción	$\neg (p \wedge \neg p)$
Exclusión del término medio	$p \vee \neg p$

	REGLA DE INFERENCIA LÓGICA			LEY LÓGICA	
	Premisa 1	Premisa 2	Conclusión		
<i>Modus ponendo ponens</i>	$p \rightarrow q$	$p$	$q$	$[(p \rightarrow q) \wedge p] \rightarrow q$	
<i>Modus tollendo tollens</i>	$p \rightarrow q$	$\neg q$	$\neg p$	$[(p \rightarrow q) \wedge \neg q] \rightarrow \neg p$	
<i>Modus tollendo ponens</i>	$p \vee q$	$\neg p$	$q$	$[(p \vee q) \wedge \neg p] \rightarrow q$	
	$p \vee q$	$\neg q$	$p$	$[(p \vee q) \wedge \neg q] \rightarrow p$	
Silogismo hipotético	$p \rightarrow q$	$q \rightarrow r$	$p \rightarrow r$	$[(p \rightarrow q) \wedge (q \rightarrow r)] \rightarrow (p \rightarrow r)$	
Silogismo disyuntivo	$p \vee q$	$p \rightarrow r$	$q \rightarrow s$	$r \vee s$	$(p \vee q) \wedge (p \rightarrow r) \wedge (q \rightarrow s) \rightarrow (r \vee s)$
Doble negación	$\neg \neg p$		$p$	$\neg \neg p \rightarrow p$	
	$p$		$\neg \neg p$	$p \rightarrow \neg \neg p$	
Simplificación	$p \wedge q$		$p$	$p \wedge q \rightarrow p$	
	$p \wedge q$		$q$	$p \wedge q \rightarrow q$	
Adjunción	$p$	$q$	$p \wedge q$	$p \wedge q \rightarrow q \wedge p$	
Adición	$p$		$p \vee q$	$p \rightarrow (p \vee q)$	
	$q$		$p \vee q$	$q \rightarrow (p \vee q)$	
Simplificación disyuntiva	$p \vee p$		$p$	$(p \vee p) \rightarrow p$	
Conmutación	$p \wedge q$		$q \wedge p$	$(p \wedge q) \rightarrow (q \wedge p)$	
	$p \vee q$		$q \vee p$	$(p \vee q) \rightarrow (q \vee p)$	
Proposiciones bicondicionales	$p \leftrightarrow q$		$p \rightarrow q$	$(p \leftrightarrow q) \rightarrow (p \rightarrow q)$	
	$p \leftrightarrow q$		$q \rightarrow p$	$(p \leftrightarrow q) \rightarrow (q \rightarrow p)$	
	$p \rightarrow q$	$q \rightarrow p$	$p \leftrightarrow q$	$[(p \rightarrow q) \wedge (q \rightarrow p)] \rightarrow (p \leftrightarrow q)$	
De Morgan	$\neg (p \wedge q)$		$(\neg p \vee \neg q)$	$\neg (p \wedge q) \rightarrow (\neg p \vee \neg q)$	
	$\neg (p \vee q)$		$(\neg p \wedge \neg q)$	$\neg (p \vee q) \rightarrow (\neg p \wedge \neg q)$	

Desde el punto de vista práctico para la actividad científica, no es imprescindible tener un conocimiento inmediato ni exhaustivo de esta lista, porque nuestra mente emplea las reglas de manera intuitiva e inconsciente, generalmente. Por esa razón decimos, en el lenguaje común, que esos razonamientos nos resultan «lógicos», en el sentido de evidentes, aunque en algunos casos no lo son tanto.

Lo que sí puede resultar más interesante es reconocer las falacias lógicas, es decir, los razonamientos que creemos correctos pero que no lo son. Encontraremos algunos de ellos en los próximos capítulos, y será importante identificarlos porque son los grandes enemigos de la investigación científica, ya que proporcionan conclusiones erróneas disfrazadas de corrección lógica.

## 2.6. Nos merecemos una explicación: la explicación en ciencia

Las explicaciones se refieren a las razones, causas o motivos por los que sucede un hecho, y son respuestas a porqués. La pregunta se formula como «¿por qué X?», donde X representa lo que hay que explicar, que se denomina *explanandum*. La respuesta se formula como «X porque Y», donde la Y es la razón o motivo y se denomina *explanans*.

Existen diversos tipos de explicación según el ámbito al que hace referencia el *explanans*. Así, se distingue entre la explicación causal, donde se atribuye una causa propiamente dicha; la explicación nomológica, que se apoya en la existencia de una ley científica; y la explicación funcional,

basada en la función desempeñada dentro de un sistema. Sobre cada una de ellas se profundizará en los próximos capítulos, además de mencionar otros tipos como la explicación teleológica o la explicación intencional, fundamentadas en las decisiones conscientes de los individuos y que son propias de las ciencias sociales.

### AMPLIACIÓN. La lógica de la explicación

Desde el punto de vista de la lógica, explicar consiste en aportar un *explanans* que implique necesariamente la ocurrencia del *explanandum*. Por ejemplo, ante la pregunta de por qué la acera está mojada, un razonamiento lógico que sustentaría la explicación podría ser el siguiente:

*Explanandum* (lo que se quiere explicar): «la acera está mojada» (o, en forma de pregunta, «¿por qué la acera está mojada?»).

*Explanans* (la explicación): «si llueve, entonces se moja la acera» (proposición general) y «ha llovido» (proposición circunstancial).

Es decir, la acera está mojada porque se moja siempre que llueve, y además se da la circunstancia de que recientemente ha llovido. Como se observa en el ejemplo, el *explanans* se ha separado en una proposición general y en otra circunstancial (o específica); en un mismo *explanans* pueden darse varias de cada una de ellas. Esta es la estructura típica de la explicación en el marco de una teoría científica, aunque habitualmente no se formule de manera tan explícita. La proposición o premisa general es una ley o principio que relaciona unos hechos con otros (en ocasiones se puede decir que relaciona unas causas con unas consecuencias). Una premisa circunstancial, por su parte, recoge uno de esos hechos. De ambas se deduciría lógicamente una consecuencia, que es el *explanandum*, lo que se quería explicar. Un *explanans* dado no tiene por qué ser el único posible, ya que pueden existir otros de los que se deduzca lógicamente el mismo *explanandum* (por ejemplo, «cuando el jardinero riega, entonces se moja la acera», junto con «el jardinero ha regado»).

La simple asociación entre un hecho observado y una ley o regla (premisa general), en ausencia de premisa circunstancial, no representa una certeza con rigor lógico, aunque sí puede resultar muy probable. Así, si vemos la acera mojada podemos pensar que probablemente haya llovido, haciendo uso de la bien conocida regla de que cuando llueve, la acera se moja. Aunque no tenemos certeza de que haya llovido realmente (habría que comprobarlo empíricamente), lo podemos suponer con cierta probabilidad, pero sin olvidar que podemos estar equivocados: ¿y si ha venido el jardinero y ha regado, o se ha roto una tubería? Este procedimiento lógico se denomina abducción, y junto con la deducción y la inducción, constituyen las tres formas de razonamiento básicas de la lógica, introducidas ya por Aristóteles en su obra *Primeros Analíticos* (siglo iv a.C.) y aplicadas a la lógica moderna y a la filosofía de la ciencia por el norteamericano Charles Peirce (1839-1914). En cierto modo, la abducción es una deducción en sentido inverso, y desde ese punto de vista es más parecida a una inducción; pero, a diferencia de esta última, en la abducción no se busca establecer una regla general, sino asociar un hecho observado con una regla general previamente conocida. Según Peirce, la abducción, que él denomina conjetura, es el primer paso del razonamiento científico, y para Umberto Eco (filósofo además de escritor), representa el modo de razonar del detective. Precisamente Sherlock Holmes tenía como máxima que «cuando se ha descartado lo imposible, lo que queda, por improbable que resulte, debe ser la verdad», es decir, cuando no existen alternativas al resultado de una abducción, su probabilidad aumenta hasta convertirse en una certeza.



Licencia: dominio público. Fuente: Pixabay.

# Capítulo 3.

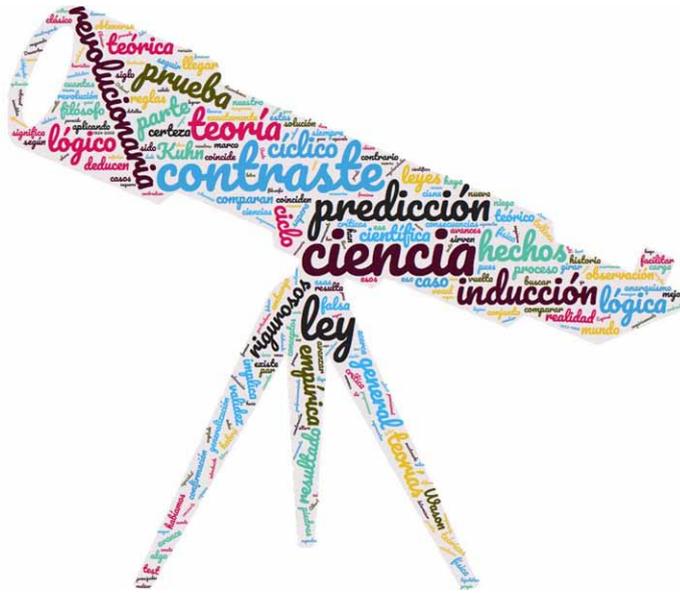
## No hay ciencia sin método: una visión actual del método científico

Los procedimientos más importantes de la metodología científica que introdujimos en el capítulo anterior forman parte de lo que se entiende globalmente como método científico, que consta de una serie de pasos bien definidos, ordenados y estructurados para avanzar en el conocimiento científico. El origen del método se remonta a la filosofía griega clásica y ha recibido aportaciones continuas a lo largo de la historia, como las de Descartes o Diderot que vimos en el capítulo anterior, pero el mayor impulso al método moderno procede de científicos y filósofos de la ciencia del siglo xx. Describiremos aquí el método general de las ciencias naturales y sociales, pero es importante señalar que para cada ciencia en concreto es posible adaptar el método a las particularidades de su campo de estudio.

El método científico no es una cuestión cerrada, ni mucho menos. Existen hoy día numerosas críticas y controversias acerca de su validez, alcance o adecuación para avanzar en los diversos campos del conocimiento. Mencionaremos algunas de ellas y los autores que las sostienen, sin pretender ser exhaustivos y con el principal objetivo de fomentar la reflexión.

### Conceptos clave:

- Método científico, inducción, deducción, empirismo lógico, heurística, hipótesis, teoría, ley empírica, ley teórica.
- Anarquismo metodológico, carga teórica, revolución científica, inconmensurabilidad, paradigma científico, ciencia normal, ciencia revolucionaria, serendipia.
- Contraste de hipótesis, verificación, confirmación (provisional), falsación, *modus tollendo tollens*, falacia de afirmación del consecuente.
- Ciencia formal, matemáticas, lógica, axioma, postulado, teorema, sistema axiomático, independencia, consistencia, completitud, ciencia normativa.



Fuente: elaboración propia con nubedepalabras.es.

### 3.1. Espiral de conocimiento: el ciclo del método científico

Los procedimientos inductivo y deductivo que vimos en el capítulo anterior se combinaron en un método científico propuesto por el físico italiano Galileo Galilei (1564-1642), resultando uno de los pilares de la Revolución Científica de los siglos *xvi* y *xvii*. El método evolucionó y se asentó durante la primera mitad del siglo *xx* a partir de corrientes filosóficas como el empirismo lógico, que compagina la observación del mundo con el razonamiento lógico.

En este método se parte de una hipótesis, de la que se deducen consecuencias que se pueden comparar con observaciones o resultados experimentales. Es esencial que las consecuencias deducidas de las hipótesis puedan ser contrastadas empíricamente, es decir, han de obtenerse predicciones sobre el mundo, susceptibles de ser observadas. Por ejemplo, si la hipótesis de partida, de carácter general, es que los objetos caerán al suelo si nada se interpone en su camino, se puede deducir de ella el resultado de un experimento particular que consiste en soltar una piedra desde cierta altura; ese experimento se puede entonces llevar a cabo y comparar el resultado con la predicción deducida de la hipótesis (evidentemente, las hipótesis científicas suelen ser más precisas, y pueden tener carácter cuantitativo y ser expresadas con lenguaje matemático). Si las predicciones coinciden con las observaciones, la hipótesis pasa la prueba del contraste empírico y puede llegar a convertirse en una ley. Un conjunto de hipótesis y leyes relativas a un determinado campo del conocimiento y compatibles entre sí forman una teoría.

El término teoría se entiende a veces en el lenguaje común como algo con carácter meramente hipotético, pero no debe confundirse con el papel que representa en la ciencia. Así, se dice en ocasiones con carácter despectivo que la teoría de la relatividad en física o la teoría de la evolución en biología son «solo teorías», como si fueran meras ocurrencias sin coherencia lógica ni base experimental. Esas apreciaciones no son correctas, ya que una teoría científica contiene un conjunto de leyes que ya han sido sometidas a exigentes contrastes empíricos, aunque bien es cierto que eso no implica que sean definitivas (algo que, como veremos después, nunca se puede afirmar).

Las hipótesis que se someterán a contraste empírico pueden obtenerse mediante inducción a partir de una generalización de hechos observados previamente. El resultado de esta generalización es lo que habíamos denominado una ley empírica, sobre la que no existe certeza de validez (recordemos la historia de los cisnes negros). Esta ley se toma como hipótesis de partida y de ella se extraen predicciones mediante un razonamiento deductivo, que luego se comparan con el experimento. Si supera la prueba, la hipótesis pasa a considerarse ley teórica.

En realidad, no es necesario que las hipótesis que se emplean en este método hayan surgido de un proceso inductivo a partir de hechos observados. Pueden haber sido inventadas directamente, porque su validez vendrá dada por el contraste con el experimento, no por su origen, si bien es cierto que el procedimiento inductivo es una fuente de inspiración habitual para la construcción de hipótesis. También resulta muy valioso para el avance de la ciencia encontrar hechos que carecen de explicación en las teorías o leyes vigentes. Es decir, no se trata tanto de registrar un gran número de hechos compatibles con las leyes conocidas, sino en buscar los que no encajan con ellas, los hechos problemáticos. Para explicar esos hechos se elabora una hipótesis nueva (quizá como modificación de la anterior), cuyas predicciones se someterán a contraste empírico.

El método científico que acabamos de describir recibe el nombre de hipotético-deductivo, porque parte de una hipótesis, y de ella se deducen predicciones sobre casos particulares que se comparan con el experimento. Este método se puede entender como un proceso cíclico: las observaciones sugieren por inducción una hipótesis general, de la cual se obtienen por deducción unas predicciones que se comparan con las observaciones, que pueden apoyar la hipótesis inicial o sugerir, por inducción, una nueva hipótesis, y así sucesivamente. Cada vez que se da una vuelta en este ciclo, la hipótesis que se maneja se ajusta mejor a la realidad; si se añade un eje temporal al esquema, se convierte en una espiral de avance del conocimiento científico.

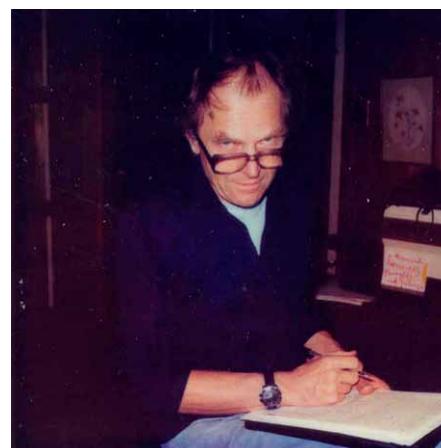
Una parte de este proceso cíclico tiene carácter empírico, al referirse al contraste con la observación o el experimento, mientras que otra parte tiene carácter teórico, porque involucra generalizaciones o deducciones. También se puede asociar una parte del ciclo al método inductivo y la otra mitad al método deductivo; en la primera se pasa de lo particular a lo general siguiendo procedimientos heurísticos (no rigurosos, como la extracción de regularidades empíricas), mientras que en la segunda se pasa de lo general a lo particular siguiendo procedimientos de deducción lógica, estos sí totalmente rigurosos. Como señalamos antes, la parte inductiva del ciclo no tiene por qué darse necesariamente, ya que lo importante de las hipótesis científicas no es su origen sino el éxito con el que explican o predicen; no obstante, los métodos heurísticos, y en particular la inducción, suelen ser su germen.



Fuente: elaboración propia.

### 3.2. ¿Un modelo a seguir?: críticas al método científico

El filósofo austríaco Paul Feyerabend (1924-1994) defendió que los principales avances de la ciencia ocurrían al ignorar o incluso al desobedecer las reglas del método científico, postura que lleva el nombre de anarquismo metodológico. Para Feyerabend, seguir las reglas del método científico implica mantenerse en los marcos teóricos ya establecidos, impidiendo la aparición de nuevas ideas. La razón es que incluso los propios hechos observados, de donde podrían surgir hipótesis nuevas por inducción, se interpretan ya desde el principio a partir de las teorías científicas vigentes, es decir, van provistos de una carga teórica. Por ejemplo, al identificar por observación un cisne blanco, ¿cómo estamos definiendo exactamente un cisne? (puede haber distintas variedades dentro de la misma



Paul Feyerabend. Licencia: libre uso. Fuente: Wikimedia Commons. Autor: Grazia Borrini-Feyerabend.

especie, o ejemplares poco habituales que no compartan ciertas características con los demás); ¿qué entendemos exactamente por blanco? (puede haber pequeñas variaciones en el tono). La respuesta a estas y otras preguntas similares viene determinada, o al menos influida, por los constructos teóricos que se manejan en cada momento.

Otra crítica hacia el método científico en la forma en que lo hemos descrito fue elaborada por el físico y filósofo de la ciencia norteamericano Thomas Kuhn (1922-1996). Kuhn, al igual que Feyerabend, se interesó por el mecanismo que hace abandonar un marco teórico y adoptar uno nuevo muy diferente al anterior. Puede llegar a ser tan distinto, que los nuevos conceptos, términos, procedimientos, etc. pueden no ser ni siquiera comparables o equivalentes a los de la teoría precedente, lo que se conoce como inconmensurabilidad. Los grandes avances de la ciencia, según Kuhn, tienen lugar a través de revoluciones científicas que desembocan en un cambio de paradigma científico, y no tanto mediante una sucesión de pequeños pasos de perfeccionamiento de la teoría vigente en ese momento. Estos últimos son los que se alcanzan aplicando el método científico habitual, que para Kuhn constituye los periodos de «ciencia normal», en contraste con los periodos de «ciencia revolucionaria». Un ejemplo de revolución científica es la que precipitaron en la física las teorías de la relatividad de Albert Einstein, que cambiaron por completo los conceptos de espacio y de tiempo respecto al paradigma anterior, el de la mecánica de Newton.



Thomas Kuhn. Licencia: CC BY-SA 4.0. Fuente: Wikimedia Commons. Autor: Davitrip.

### HISTORIAS DE LA CIENCIA. La serendipia de Fleming y Becquerel

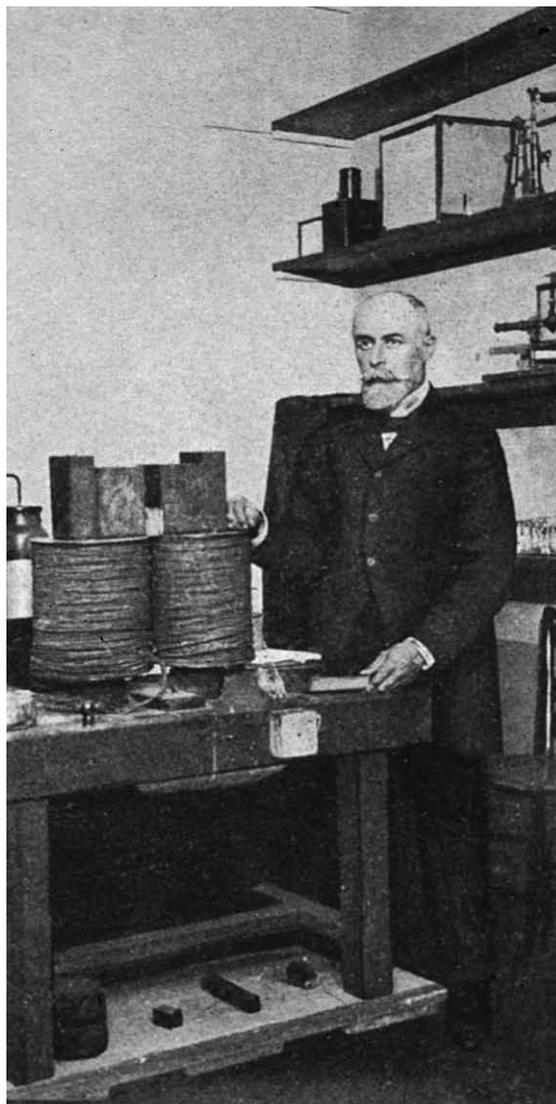
El término serendipia se emplea para referirse a descubrimientos casuales o accidentales que se dan en la vida cotidiana, y también en ciencia. La palabra tiene su origen en Serendip, nombre persa de la actual Sri Lanka, y pasó a emplearse en las lenguas europeas a través de una leyenda sobre tres príncipes procedentes de aquel lugar a los que algunas casualidades y su ingenio les sacaron de varios aprietos.

Un caso de serendipia científica muy conocido es el descubrimiento de la penicilina por el médico y microbiólogo escocés Alexander Fleming (1881-1955). Antes de irse de vacaciones en agosto de 1927, Fleming dejó unos cultivos de estafilococos en su laboratorio. A la vuelta, observó que uno de los cultivos se había contaminado por un hongo, y que alrededor de él los estafilococos habían desaparecido. Tras varios meses de investigaciones, concluyó que el hongo que había aparecido, del género *Penicillium*, segregaba una sustancia, que llamó penicilina, que actúa como antibiótico eliminando bacterias causantes de numerosas enfermedades. Su uso médico, que se extiende hasta nuestros días, no comenzó hasta 1942.

Otro caso muy interesante de serendipia en la ciencia, quizás menos conocido, es el del físico francés Henri Becquerel (1852-1908). En 1896 se encontraba en París estudiando el fenómeno de la fosforescencia, según el cual algunas sustancias, como las sales de uranio, absorben la luz y posteriormente la reemiten en forma de otro tipo de rayos capaces de atravesar materiales. Esto lo averiguó haciendo experimentos con placas fotográficas envueltas en papeles opacos, sobre las que colocaba las sales de uranio; al exponerlo a la luz solar, las sales absorbían la luz y emitían esa otra radiación que atravesaba el papel opaco y dejaba una sombra sobre la placa fotográfica. Mientras realizaba estos experimentos, tuvo la «mala» suerte de que se sucedieron muchos días nublados en París, así que decidió guardar la placa fotográfica en un cajón junto con las sales de uranio para seguir más adelante con sus investigaciones. Como

las sales no habían estado expuestas a la luz, no deberían haber dejado ninguna huella en la placa, pero lo que descubrió fue que habían dejado una sombra muy intensa. Becquerel concluyó que esa sustancia emitía unas radiaciones, independientes de la luz solar (es decir, sin relación con la fosforescencia), que también atravesaban materiales y que no disminuían de intensidad a corto plazo: había descubierto la radiactividad. Los años siguientes siguió estudiando este fenómeno, junto con otros científicos tan insignes como Marie y Pierre Curie, descubriendo detalles sobre esas radiaciones (que son de varios tipos) y sobre su origen, que es el núcleo de los átomos. Hoy la física nuclear, además de ser un campo esencial para entender la estructura de la materia, tiene numerosas aplicaciones, como la obtención de energía en las centrales nucleares o el diagnóstico y el tratamiento de enfermedades.

Con estos y otros muchos casos de serendipia, cabe preguntarse si el método científico es realmente imprescindible para avanzar en ciencia. Y también cabe preguntarse si es necesario ser un buen científico, o si por el contrario cualquier persona con suerte puede realizar esos descubrimientos de la misma manera que puede ganar la lotería. Un poco de reflexión acerca de estas historias de la ciencia nos obligaría a concluir que estos grandes descubrimientos casuales solo les ocurren a grandes mentes científicas, muy entrenadas para distinguir la relevancia de lo que observan. Además, el hallazgo casual es solo un punto de partida, porque entender en profundidad ese hecho y sus implicaciones requiere de muchos años de experimentos, formulación y contrastación de hipótesis, es decir, de aplicación del método científico.



Henri Becquerel. Licencia: dominio público.  
Fuente: Wikimedia Commons.

### 3.3. Al traste con el contraste: de la verificación a la falsación

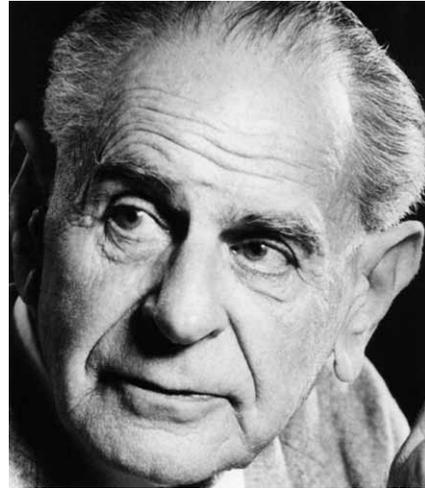
El contraste empírico de hipótesis exige una reflexión cuidadosa. Un posible procedimiento es la verificación, es decir, la comprobación de que las predicciones que se deducen de las hipótesis coinciden con las observaciones. Llevar a cabo una verificación exhaustiva y completa es imposible, porque habría que realizar infinitas observaciones, en todos los lugares posibles y en cualquier momento de tiempo. A lo máximo que se puede aspirar es a lograr una confirmación provisional, más robusta cuantas más verificaciones se hayan realizado, pero nunca completa. Este es un problema parecido al que ocurría con la inducción, que puede proporcionar expectativas de veracidad sobre la ley extraída, pero nunca certeza.

Debido a esa limitación, el filósofo austríaco Karl Popper (1902-1994) consideraba que el grado de confirmación de las hipótesis, a través de su continua verificación, permanecía en realidad siempre nulo, porque siempre quedaría una cantidad inabarcable de situaciones por comprobar. En su lugar, Popper propuso que el contraste de hipótesis debía llevarse a cabo a través de la falsación, es decir, buscando los hechos que contradicen las predicciones de las hipótesis, y no tanto los que las confirman. Basta con que una observación no coincida con la predicción para que la hipótesis deba ser rechazada.

Si, por el contrario, la hipótesis supera las pruebas de falsación a las que se ve sometida, queda aceptada provisionalmente. En resumen, la verificación de una hipótesis nunca llega a ser completa por muchas predicciones deducidas de ella que se hayan cumplido, pero una única predicción no cumplida es suficiente para falsar la hipótesis, es decir, para rechazarla (aunque también existe la posibilidad, de la que hablaremos más adelante, de mantenerla mediante la incorporación de hipótesis accesorias, llamadas *ad hoc*).

Para favorecer la falsación de hipótesis o de las teorías que las engloban, estas deben ser construidas para facilitar predicciones audaces, inverosímiles, inauditas o fuera de lo habitual. Si esas predicciones no se dan en la realidad, se rechazará la hipótesis. Si, por el contrario, esas extrañas predicciones se cumplen, la hipótesis habrá pasado una dura prueba de falsación.

El objetivo de la falsación es, pues, facilitar lo máximo posible el rechazo de nuestras hipótesis, por mucho trabajo que nos haya costado elaborarlos. En cierto modo, este procedimiento consiste en echar piedras sobre nuestro propio tejado, cuantas más piedras y más pesadas, mejor, para asegurarnos de que el edificio teórico que estamos construyendo es sólido. Para Popper, el espíritu científico no consiste en la búsqueda de la certeza, sino en la permanente crítica y revisión de lo que se da por sabido.



Karl Popper. Licencia: sin restricciones.  
Fuente: Wikimedia Commons.

### AMPLIACIÓN. La lógica de la verificación y la falsación

Detrás de la verificación y de la falsación existen dos razonamientos lógicos distintos. En el caso de la verificación, el razonamiento tiene la siguiente estructura:

Premisa 1: Si la hipótesis H es verdadera, entonces su predicción empírica P es verdadera.

Premisa 2: La predicción P es verdadera (coincide con los resultados experimentales).

Conclusión \*: La hipótesis H es verdadera.

Este razonamiento resulta ser lógicamente incorrecto (lo que se indica con un asterisco al enunciar la conclusión), como se puede entender más fácilmente con el siguiente ejemplo simplificado:

Premisa 1: Si se acaba la gasolina del coche, entonces el coche deja de funcionar.

Premisa 2: El coche ha dejado de funcionar.

Conclusión \*: Se ha acabado la gasolina del coche.

Esta conclusión es errónea, o falaz, porque el coche puede dejar de funcionar por razones distintas a que se haya acabado la gasolina, como, por ejemplo, debido a una avería. Dicho de otro modo, que un coche no funcione no implica necesariamente que no tenga gasolina. Este razonamiento incorrecto se denomina falacia de la afirmación del consecuente, y consiste en pensar que, si es cierta una relación condición-consecuencia entre dos hechos (premisa 1), y ocurre la consecuencia (premisa 2), entonces ocurre necesariamente la condición incluida en la premisa 1 (lo cual se expresa como conclusión, pero que es errónea). Es importante señalar que este ejemplo es solamente una versión simplificada de la estructura lógica de la verificación, pues en él no se está tratando de verificar una hipótesis, sino de concluir si se había dado una condición (gasolina agotada) al haber observado su consecuencia (coche que no funciona).



Licencia: dominio público. Fuente: Pixabay

En el caso de la falsación, por su parte, el razonamiento lógico puede expresarse como:

Premisa 1: Si la hipótesis H es verdadera, entonces su predicción empírica P es verdadera.

Premisa 2: La predicción P es falsa (al menos un resultado experimental no coincide con ella).

Conclusión: La hipótesis H es falsa.

Este razonamiento lógico sí es correcto y riguroso, y es conocido desde la antigüedad. Se denomina *modus tollendo tollens*, que significa «modo que niega negando», es decir, razonamiento que establece que si un hecho implica otro, y no se da (se niega) el segundo, entonces no se puede haber dado (se niega) el primero. En este caso, se niega la hipótesis que implica una cierta predicción mediante la negación de esa predicción. La ley lógica correspondiente se expresa simbólicamente como:  $[(p \rightarrow q) \wedge \neg q] \rightarrow \neg p$ , donde el símbolo  $\neg$  indica negación (no es verdadero el enunciado que le sigue). No hay que confundir esta ley con el *modus ponendo ponens* que se introdujo en uno de los ejemplos del capítulo anterior.

El ejemplo de la gasolina quedaría en el caso de la falsación como sigue:

Premisa 1: Si se acaba la gasolina del coche, entonces el coche deja de funcionar.

Premisa 2: El coche NO ha dejado de funcionar.

Conclusión: NO se ha acabado la gasolina del coche.

Es decir, si el coche deja de funcionar cuando se acaba la gasolina, pero ahora está funcionando, significa necesariamente que la gasolina no se ha acabado. Se pueden buscar excepciones a esta deducción, como por ejemplo que tengamos un coche híbrido que pueda funcionar también con la electricidad acumulada en una batería. Pero esos casos afectarían a la veracidad de la premisa 1, que podría no ser cierta, pero no invalidarían el razonamiento lógico en sí, según el cual, si las premisas 1 y 2 son verdaderas, entonces necesariamente la conclusión es verdadera.

### 3.4. Formalidad ante todo: el método de las ciencias formales

El método de las ciencias formales no contiene una fase de contraste de hipótesis mediante observación o experimentación, porque estas ciencias no se ocupan de describir o explicar la realidad. Su cometido es encontrar las consecuencias lógicas de un conjunto de proposiciones, que reciben el nombre

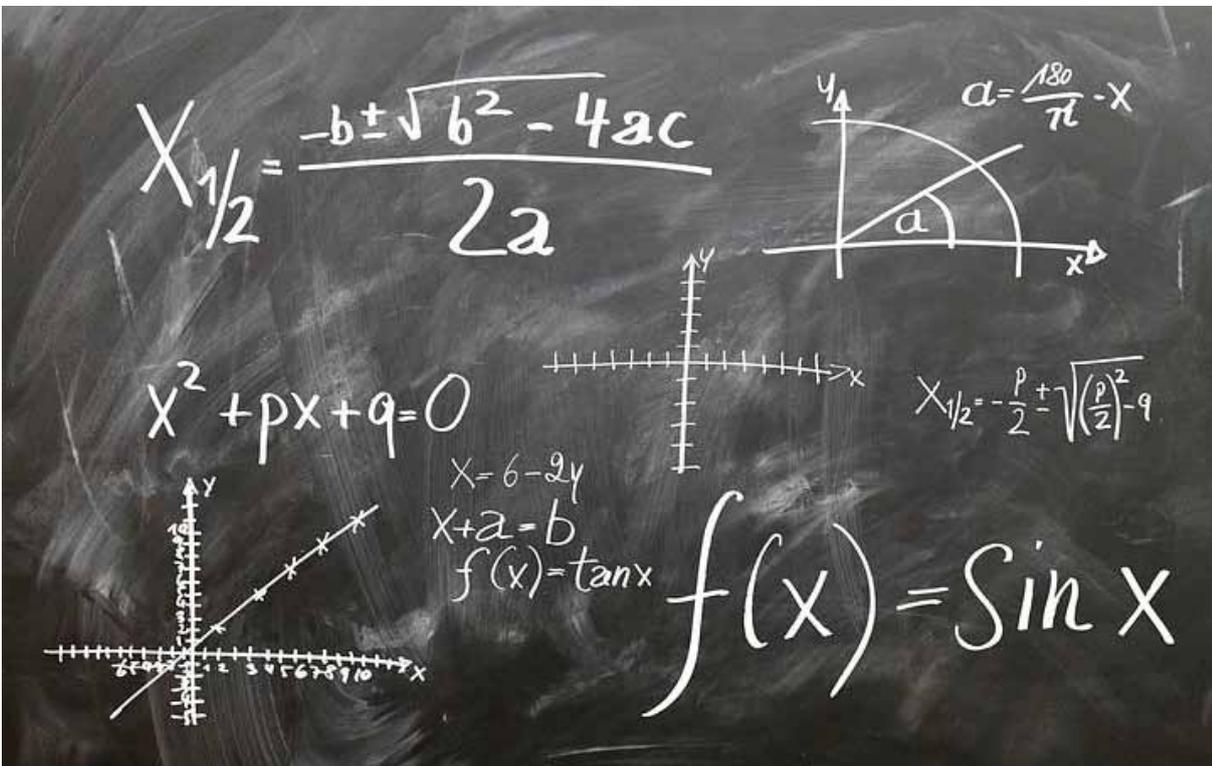
de axiomas o postulados. Antiguamente se distinguía entre estos dos términos, reservando el primero para proposiciones cuya verdad era evidente, y el segundo para proposiciones cuya verdad no lo era; actualmente no se hace distinción, y ambos se emplean para proposiciones básicas de un sistema sobre las que no se discute ni su veracidad ni si esta es evidente o no.

A partir de los axiomas y mediante los razonamientos o deducciones lógicas que vimos en el capítulo anterior, las ciencias formales generan nuevas proposiciones ciertas, que se denominan teoremas. La diferencia con los axiomas o postulados es que estos se dan por ciertos sin demostración, mientras que los teoremas han de ser demostrados a partir de los axiomas. A su vez, los axiomas y los teoremas ya demostrados sirven para generar nuevos teoremas, y así sucesivamente, hasta construir una estructura lógica formalmente válida que se sustenta en un pequeño conjunto o sistema de axiomas.

Un sistema axiomático adecuado debe cumplir los siguientes requisitos:

- Independencia: ninguno de los axiomas debe poder ser deducido o demostrado a partir de los otros.
- Consistencia: partiendo de ese conjunto de axiomas no se debe poder deducir un teorema y su contrario.
- Completitud: cualquier enunciado que se pueda formular dentro del sistema debe poder ser demostrado como cierto (siendo entonces un teorema) o como falso a partir únicamente de los axiomas del sistema.

Los ejemplos más importantes de ciencias formales son las matemáticas y la lógica. De esta última se dice además que es una ciencia normativa, porque no describe cómo razona el ser humano, sino cómo debe hacerlo para que las conclusiones obtenidas sean formalmente válidas. Ambas ciencias se utilizan en mayor o menor medida como herramienta por las ciencias naturales y sociales; se puede decir que, de la misma manera que la componente empírica de esas ciencias recurre a menudo a instrumentos para la observación, la componente teórica recurre a instrumentos formales como las matemáticas o la lógica.



Licencia: dominio público. Fuente: Pixabay

## HISTORIAS DE LA CIENCIA. Los postulados de Euclides

El matemático griego Euclides (siglo IV a.C.) estableció un sistema de cinco postulados sobre los cuales se podían demostrar todos los teoremas de la geometría. Ese sistema cumplía las propiedades de independencia, consistencia y completitud, es decir, ni faltaba ni sobraba ningún postulado para poder demostrar cualquier teorema geométrico, razón por la cual permaneció inalterable durante más de dos mil años.

Dado que la veracidad de los postulados no es un atributo significativo, es posible cambiarlos siempre que se mantengan las buenas propiedades del sistema axiomático. Los teoremas que se construyan a partir del nuevo sistema serán diferentes a los anteriores, pero igualmente válidos desde el punto de vista formal.

El último de los postulados de Euclides establece que «por un punto exterior a una recta solo pasa una única recta paralela a ella» (en realidad Euclides lo formuló de una manera más enrevesada, pero totalmente equivalente a esta). A partir del siglo xviii se sucedieron numerosos intentos de modificar este postulado, llegándose a dos nuevas versiones del mismo que, junto con los otros cuatro, formaban dos nuevos sistemas. Por un lado, un quinto postulado de la forma «por un punto exterior a una recta pasan infinitas rectas paralelas a ella» define una geometría denominada aguda o hiperbólica, en la que, entre otras propiedades, la suma de los ángulos de los triángulos es menor de 180 grados (en lugar de exactamente 180, como ocurre en la geometría euclídea o plana). Por otro lado, un quinto postulado de la forma «por un punto exterior a una recta no pasa ninguna recta paralela a ella» define una geometría denominada obtusa o elíptica, en la que la suma de los ángulos de los triángulos es mayor de 180 grados, entre otras propiedades.

Estas nuevas geometrías son formalmente tan válidas como la euclídea, que define la geometría plana con la que estamos más familiarizados, pero parecerían no tener utilidad en el mundo real. Pero eso no es así; por ejemplo, la geometría sobre la superficie terrestre, que es una esfera (aproximadamente), es de tipo elíptica, aunque la diferencia con la geometría plana no se nota si no es a muy largas distancias. Según la teoría de la relatividad de Einstein, la geometría de nuestro universo a gran escala podría ser elíptica o hiperbólica.



Euclides. Licencia: dominio público.  
Fuente: Wikimedia Commons.

# Capítulo 4.

## Rebelde con causa: la explicación causal

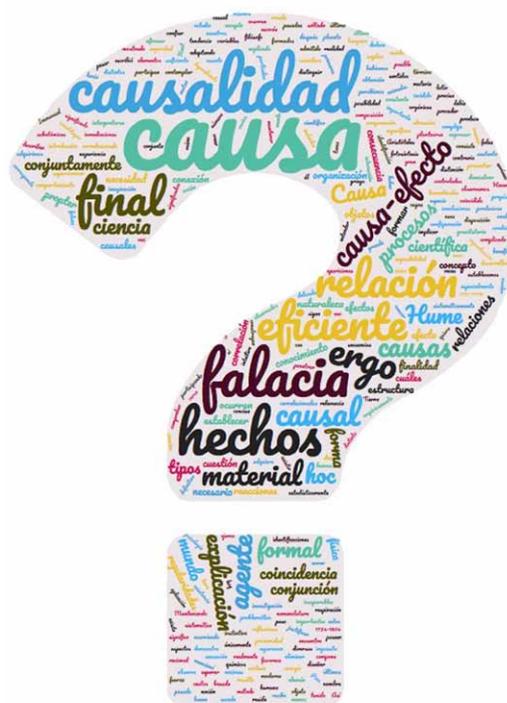
La explicación causal consiste en justificar la existencia de un hecho a través de la causa que lo produce. Esta explicación establece que, si ha ocurrido la causa, entonces habrá tenido lugar el hecho en cuestión, que es su consecuencia o efecto.

Por ejemplo, una explicación causal en el ámbito de la física es que una piedra soltada cerca del suelo cae hacia él porque se encuentra en las inmediaciones de un objeto de mucha masa, el planeta Tierra, que ejerce una fuerza de atracción gravitatoria sobre la piedra.

El término causa reúne varios significados diferentes, como el de causa material, causa formal, etc., que describiremos en este capítulo. Además, advertiremos sobre algunas relaciones que no son de causa-efecto, pero que a veces se confunden con ellas, y que representan una seria amenaza para la comprensión de la ciencia.

### Conceptos clave:

- Causa, relación causa-efecto, causa material, causa formal, causa final, causa agente o eficiente, falacia *cum hoc ergo propter hoc*, causalidad, correlación.

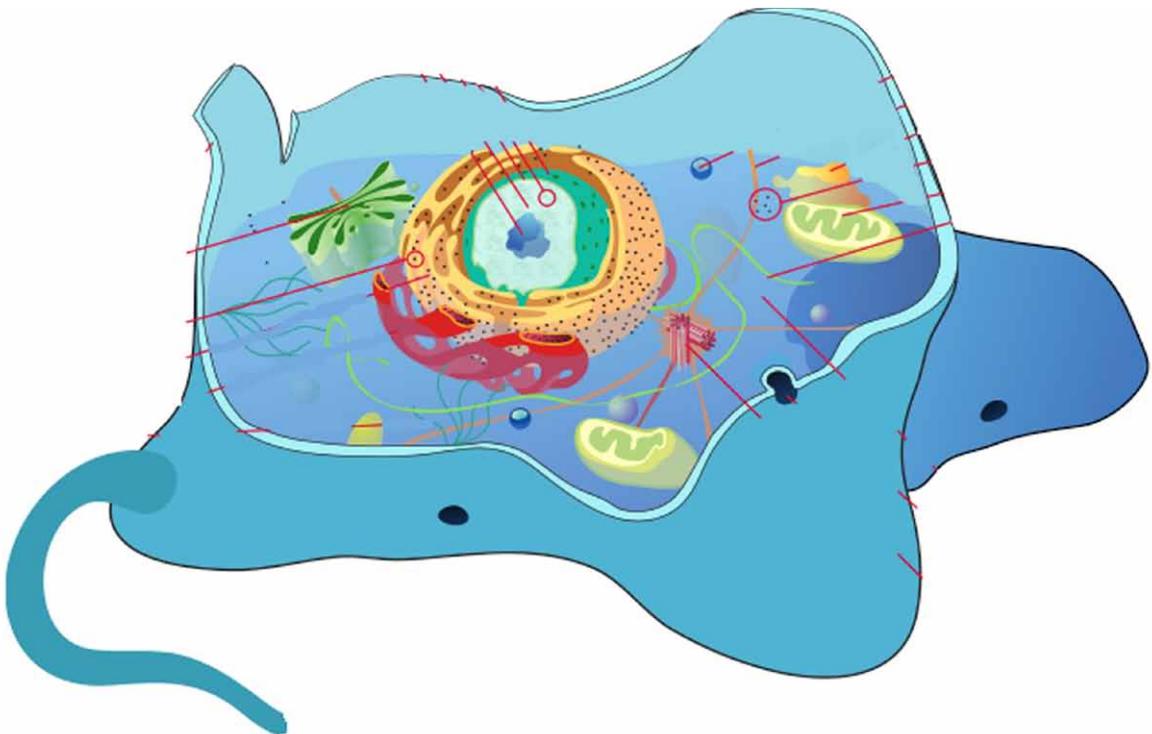


Fuente: elaboración propia con nubedepalabras.es.

#### 4.1. Inventario de causas: tipos de relaciones causa-efecto

En las diversas ciencias se emplean distintos tipos de causas, algunas de cuyas primeras identificaciones se deben al filósofo griego Aristóteles (384 a.C - 322 a.C.). Manteniendo en buena parte su nomenclatura, pero adaptando el significado allí donde resulta necesario, esos mismos tipos de causas tienen relevancia en diferentes aspectos de la ciencia moderna:

- Causa material: Hace referencia a la composición material de los cuerpos. Por ejemplo, la causa material de los procesos que tienen lugar en una célula viva son los átomos que la constituyen, o las partículas subatómicas de las que están formados esos átomos, o las moléculas que se construyen a partir de ellos, dependiendo de la escala física en la que se quiera analizar la cuestión.
- Causa formal: Hace referencia a la disposición o configuración que toma la materia que compone los cuerpos, es decir, a su estructura u organización. Por ejemplo, la causa formal de los procesos que tienen lugar en una célula viva se establece a partir de la precisa organización de sus átomos para formar moléculas orgánicas de estructura compleja, que participan en largas secuencias de reacciones químicas controladas en cada paso por otras moléculas, las enzimas.
- Causa final: Hace referencia a la finalidad, utilidad o beneficio. Por ejemplo, la causa final de los procesos que tienen lugar en una célula viva puede hacer referencia a la obtención de energía (respiración), a la síntesis de biomoléculas (fotosíntesis), a la reproducción, etc.; en definitiva, la finalidad última es mantener la célula viva.
- Causa agente o causa eficiente: Hace referencia a la acción de unos cuerpos sobre otros. Por ejemplo, la causa eficiente de los procesos que tienen lugar en una célula viva se puede establecer como la interacción electromagnética de unos átomos con otros para formar moléculas, o de unas moléculas con otras participando en reacciones químicas.



Licencia: dominio público. Fuente: Wikimedia Commons.

Estos cuatro tipos de causas no son mutuamente excluyentes, y todas pueden emplearse en la explicación científica causal, pero algunas son más típicas de unas ciencias que de otras. Por ejemplo, en biología y ciencias afines, así como en ciencias sociales, son habituales las referencias a causas finales, que surgen en el marco de la explicación funcional y se tratarán en un capítulo posterior. En cualquier caso, desde hace unos siglos la ciencia tiende a considerar más adecuada y a usar con más frecuencia la causa eficiente de los hechos, es decir, trata de identificar cómo unos hechos o cuerpos actúan sobre otros, y estos últimos sobre otros distintos, y así sucesivamente.

## 4.2. Oro parece, causa no es: falsas relaciones causales

Probar estrictamente la relación causal entre dos hechos no es posible, porque lo único que realmente se observa es la coincidencia o proximidad de esos hechos en el espacio y en el tiempo. Para el filósofo escocés David Hume (1711-1776), la relación de causalidad era una sensación creada por nuestra mente, que nos hace esperar que ciertos hechos sigan ocurriendo conjuntamente si nuestra experiencia nos dice que así ha sucedido siempre en el pasado. Para otros filósofos, como el alemán Immanuel Kant (1724-1804), la relación de causalidad era una necesidad impuesta por nosotros para poder entender el mundo de manera racional. Desde este último punto de vista, se puede decir que la causalidad es necesaria para la comprensión científica del mundo, que es el asunto que nos ocupa aquí.

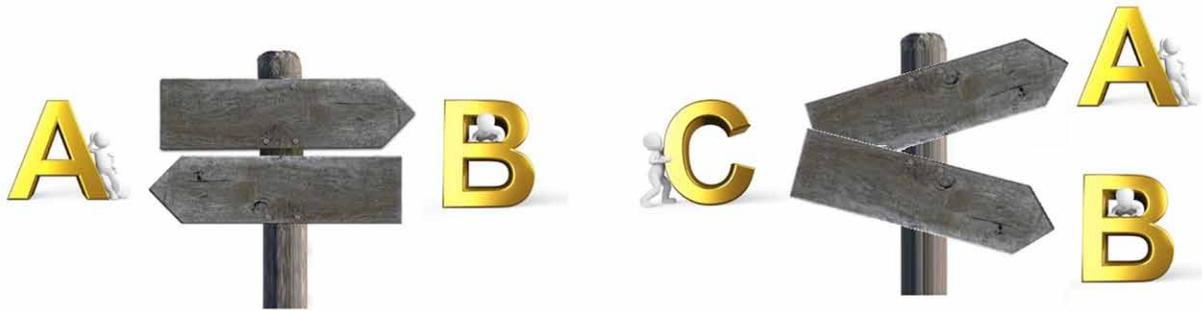
Una vez admitida la necesidad, o al menos la utilidad científica de la identificación de las causas, es importante ser muy cautos para no establecer relaciones causales allí donde en realidad no las hay. Este error recibe el nombre de falacia *cum hoc ergo propter hoc*, que en latín significa «con esto, por tanto a causa de esto». Es decir, al producirse un hecho unido a otro se deduce que el segundo es causa del primero, pero de forma errónea o falaz. Cuando dos hechos, A y B, ocurren sistemáticamente de manera conjunta (es decir, no se trata de una coincidencia casual, sino que ocurre un gran número de veces), se dice que están correlacionados. Por eso, otra forma de expresar esta falacia es decir que la correlación no implica causalidad.

En efecto, la coincidencia sistemática entre hechos no puede interpretarse siempre como una relación causal entre ellos. En primer lugar, de existir relación causal entre los hechos A y B, habría que distinguir si A es la causa de B o si, por el contrario, B es la causa de A, o incluso si se da entre ellos una relación más complicada basada en la causalidad mutua. En segundo lugar, habría que contemplar la posibilidad de que A y B ocurran conjuntamente porque ambos sean consecuencia de otro hecho distinto C, que es la causa de ambos. De este modo, si ocurre el hecho C, ocurren conjuntamente A y B, pero estos dos últimos no son causa uno del otro.

Estas reflexiones son especialmente importantes a la hora de diseñar y desarrollar una investigación científica. Antes de dar por buenas las conclusiones sobre causas y efectos es necesario plantearse, en primer lugar, cuáles son las causas y cuáles son los efectos, y también si puede haber circunstancias no incluidas en el estudio que sean la causa común de los hechos observados. En el contexto del diseño de una investigación científica, que trataremos en un capítulo posterior, hablaríamos de identificar correctamente las variables independientes, dependientes y las posibles variables perturbadoras.



David Hume. Licencia: dominio público.  
Fuente: Wikimedia Commons.



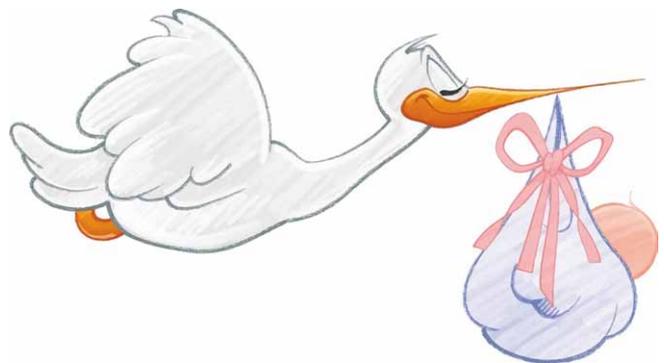
Licencia: dominio público. Fuente: Pixabay y elaboración propia.

## HISTORIAS DE LA CIENCIA. Cigüeñas con bebés, niños miopes y Premios Nobel golosos

Un reciente estudio estadístico observó una cierta coincidencia entre el número de parejas de cigüeñas y el número de nacimientos anuales de bebés (humanos) en 17 países europeos. En estadística se calcula un valor, llamado coeficiente de correlación, para expresar el grado de coincidencia entre dos variables, que en este caso son, para cada uno de los países estudiados, el número de parejas de cigüeñas y el número de bebés nacidos. El valor del coeficiente de correlación obtenido en este estudio no era demasiado alto, pero tampoco muy bajo; es decir, parecía cumplirse que cuantas más parejas de cigüeñas, mayor número de bebés. Estas y otras inferencias estadísticas, también mal interpretadas, parecían apoyar la absurda hipótesis de que la causa de los nacimientos de bebés era su transporte y entrega por parte de parejas de cigüeñas, como cuenta la leyenda (aunque no se decía nada de si procedían de París).

Un ejemplo de estudio más serio (publicado en la prestigiosa revista *Nature*), en el que una fuerte correlación se interpreta como causalidad, es el de la aparición de miopía en los niños en relación con la luz ambiental con la que duermen. Los datos recogidos por los autores mostraban que los niños que dormían con más luces encendidas eran los que más miopía padecían, y a partir de ello se podía formular la hipótesis de que altos niveles de luz ambiental durante el sueño provocaba el desarrollo de miopía (para probar esta relación causal habría que encontrar el mecanismo biofísico que relacionase ambas circunstancias, algo seguramente muy complejo). Pero otros autores propusieron enseguida una explicación causal mucho más plausible, introduciendo una nueva variable en el estudio: existía una fuerte correlación entre padres miopes e hijos también miopes, sustentada en muchos más estudios y con una probable causa genética; los padres miopes tendían a dejar luces encendidas en la casa durante la noche, precisamente por sus problemas de visión. Es decir, la causa común de la miopía de los niños y de las luces nocturnas era la miopía de los padres.

Otro ejemplo muy llamativo, publicado también en una prestigiosa revista, *The New England Journal of Medicine*, relacionaba el consumo de chocolate con la mejora de la capacidad cognitiva. Las dos variables empleadas en el estudio eran el consumo de chocolate per cápita y el número de premios Nobel recibidos por cada millón de habitantes en diferentes países, y la correlación entre ambas resultaba muy alta. La conclusión era, por tanto, que la causa de una mayor capacidad cognitiva en los habitantes de ciertos países era su mayor consumo de chocolate. Como se ha puesto de manifiesto en análisis posteriores, parece mentira que el autor



Licencia: dominio público. Fuente: Pixabay.

de este estudio no valorase la posibilidad de que ambas variables sean consecuencia de una causa común (y no una la causa de la otra), como por ejemplo el diferente nivel de vida de cada país. En realidad, seguramente serán muchas circunstancias, relacionadas de formas muy complejas, las que influyan en la capacidad cognitiva de la población. El efecto del chocolate en la cognición se asocia a unas sustancias llamadas flavonoides, pero ¿por qué no se estudiaron los patrones de consumo de otros alimentos que también contienen flavonoides, como el té o el vino tinto? Y, en cualquier caso, ¿realmente la mejor manera de medir la capacidad cognitiva de una población al completo es a través del número de premios Nobel conseguidos?



Licencia: dominio público. Fuente: Pixabay.



ley en cuestión, son relevantes para el hecho que se quiere explicar. Estas se denominan condiciones iniciales, y junto con la ley, constituyen el *explanans* (la explicación). Según la estructura lógica de la explicación que se introdujo en un capítulo anterior, del *explanans* se deduce el *explanandum* (el hecho que se quiere explicar), y por eso este tipo de explicación se denomina también nomológico-deductiva, según la terminología del filósofo alemán Carl Hempel (1905-1997).

### Conceptos clave:

– Explicación nomológica, ley, regularidad, condiciones iniciales, ley determinista, ley probabilística (estadística), ley de tendencia, cláusula *ceteris paribus*, hipótesis *ad hoc*.

## 5.1. El código legal: tipos de leyes científicas

Se distinguen los siguientes tipos de leyes científicas:

### – Leyes deterministas

Establecen relaciones entre hechos que se producen siempre de la misma manera. Por ejemplo, al soltar una piedra sobre el suelo, esta cae.

### – Leyes probabilísticas

Establecen relaciones entre hechos que ocurren con una cierta probabilidad, y se proporciona el valor de esta última. Según este tipo de leyes, un mismo hecho puede estar relacionado con otros de diversas maneras, y cada relación ocurre con una cierta probabilidad. Por ejemplo, se puede imaginar una ley que afirme que el resultado de una cierta medida arrojará el valor A con una probabilidad de 0,999 y el valor B con una probabilidad de 0,001 (las probabilidades siempre se encuentran entre 0 y 1, y la suma de las probabilidades de todos los resultados posibles debe ser 1).

Si los hechos implicados en este tipo de leyes ocurren un gran número de veces, las probabilidades que proporciona la ley pueden interpretarse como frecuencias, es decir, como el número de veces que ocurre cada resultado. Siguiendo con el ejemplo anterior, si la medida se realiza 1000 veces, cabe esperar obtener el valor A aproximadamente 999 veces y el valor B aproximadamente 1 vez. Estas frecuencias resultan de multiplicar la probabilidad de obtener cada valor por el número de repeticiones de la medida. Debido a esta interpretación, estas leyes se denominan también estadísticas, y se cumplirán con mayor precisión cuanto mayor sea el número de repeticiones de las relaciones descritas por ellas.

### – Leyes de tendencia

Establecen relaciones habituales entre hechos, pero que no ocurren siempre de la misma manera ni siempre con la misma probabilidad. Es decir, esas relaciones no pueden ser reflejadas por leyes deterministas ni por leyes probabilísticas, pero presentan una frecuencia suficiente como para resultar científicamente relevantes.

La existencia de estas leyes puede deberse a que los hechos que relacionan no estén definidos o cuantificados con la suficiente precisión, por lo que la relación entre ellos tampoco puede estarlo. También puede ocurrir que ciertas relaciones sean demasiado complejas e involucren demasiados factores como para establecer relaciones bien definidas, ya sean deterministas o probabilísticas, y que la única opción al alcance del investigador sea establecer tendencias.

### ▶ AMPLIACIÓN. Cada ciencia tiene su explicación

Cada uno de los diversos tipos de ley científica es más habitual en unas ciencias que en otras. En ciencias sociales, dada la naturaleza de las interacciones humanas, es casi imposible establecer leyes que

se cumplan siempre, o siquiera probabilidades que se mantengan regulares y estables, por lo que a lo máximo que se puede aspirar es a formular leyes de tendencia.

Las leyes deterministas son típicas de las ciencias naturales más fundamentales, como la física o la química. Las leyes probabilísticas, por su parte, aparecen con frecuencia en ciencias que se ocupan de sistemas más complejos, como puede ser la biología. En ellas, dada la complejidad del objeto de estudio, es muy difícil conocer con precisión todas las circunstancias, lo que puede justificar en algunos casos la ausencia de certezas.

Pero las áreas de aplicación de las leyes deterministas y probabilísticas no siempre están bien diferenciadas. En la propia física también son habituales las leyes probabilísticas, por dos motivos diferentes. Por un lado, en ciertas áreas de la física se estudian sistemas con un gran número de variables, cuyo conocimiento no puede ser completo, por lo que las leyes que se obtienen son de carácter estadístico; sin embargo, a menudo, los valores promedio de esas variables siguen leyes que se aproximan mucho a las deterministas. Así ocurre, por ejemplo, en la física estadística y en la termodinámica. Por otro lado, la física cuántica es de carácter intrínsecamente probabilístico, lo que implica que sus leyes solo asignan probabilidades a los diferentes resultados posibles. Esto no se debe a que falte información sobre las condiciones iniciales o a que nuestro conocimiento sobre las leyes que rigen el mundo sea solo aproximado o imperfecto, sino a la propia naturaleza del mundo físico en las pequeñas escalas (moléculas, átomos, partículas subatómicas). En el mundo microscópico, la naturaleza no admite certezas, solamente probabilidades y por tanto cierto grado de incertidumbres.

Este carácter probabilístico de la física cuántica, que se iba confirmando conforme avanzaba el siglo xx, resultaba muy sorprendente y difícil de creer. Hasta entonces, las leyes de la física parecían describir la naturaleza como el mecanismo de un reloj, como ocurría con las teorías de Newton, sustentando la denominada concepción mecanicista de la naturaleza. Uno de los más extrañados por el comportamiento que mostraba el mundo microscópico, hasta el punto de no llegar a creérselo, fue Einstein. En el contexto de sus críticas a la física cuántica formuló su conocida frase «Dios no juega a los dados», dando a entender que no era aceptable que la naturaleza (o quien la crease) dejase en manos del azar el desarrollo de los acontecimientos. Conforme avanza la física, más se confirma que Einstein estaba equivocado, y que en los niveles más profundos de la ciencia natural las leyes probabilísticas han venido para quedarse, dejando sin hueco a las deterministas.



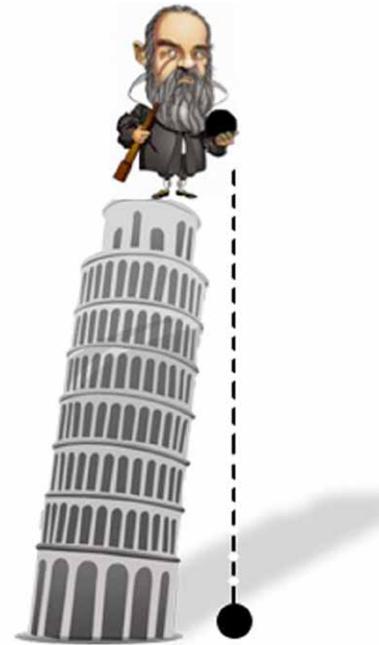
Licencia: dominio público. Fuente: Pixabay.

## 5.2. La letra pequeña de la ley: la cláusula *ceteris paribus*

Las leyes científicas suelen ser concisas, estableciendo de la manera más clara y directa posible la relación existente entre dos o más hechos. Pero es muy difícil encontrar ámbitos del conocimiento donde los hechos ocurran de forma aislada: tanto el mundo físico analizado por las ciencias naturales como las actividades humanas que estudian las ciencias sociales encierran una enorme complejidad. Por esa razón las leyes científicas van acompañadas implícitamente de la denominada cláusula *ceteris paribus*, del latín que significa que el resto (de cosas) permanecen igual. Ese resto se refiere a todos los hechos o circunstancias que pueden ocurrir a la vez que los hechos que sí están expresados en la ley, y sobre los que podrían influir.

En el ejemplo que venimos empleando, la ley acerca de la caída de los objetos al suelo viene acompañada implícitamente por una cláusula *ceteris paribus* que recogería condiciones como, por ejemplo, la ausencia de viento, que podría afectar a la trayectoria del objeto, sobre todo si es muy ligero. También supondría que no hay una persona que coge el objeto antes de que alcance el suelo, o un meteorito que choca con la Tierra y la destruye, y una infinidad de otras situaciones. La ocurrencia de cualquiera de estos hechos podría dar resultados distintos a los que predice la ley, porque el objeto ya no caería al suelo, pero eso no la invalida. Simplemente, la ley no sería aplicable porque se viola la cláusula *ceteris paribus*, al haberse modificado las circunstancias no especificadas por la ley.

Algunas de las circunstancias cambiantes que se han dado como ejemplo resultan bastante inverosímiles, como la del meteorito, pero existen otras muchas que sí son habituales y con las que es necesario ser más cauto. En el contexto de la física, por ejemplo, no resulta tan descabellado considerar los efectos de la atmósfera al analizar la caída de los cuerpos; no hay más que pensar en la diferencia entre lanzarse desde un avión con y sin paracaídas.

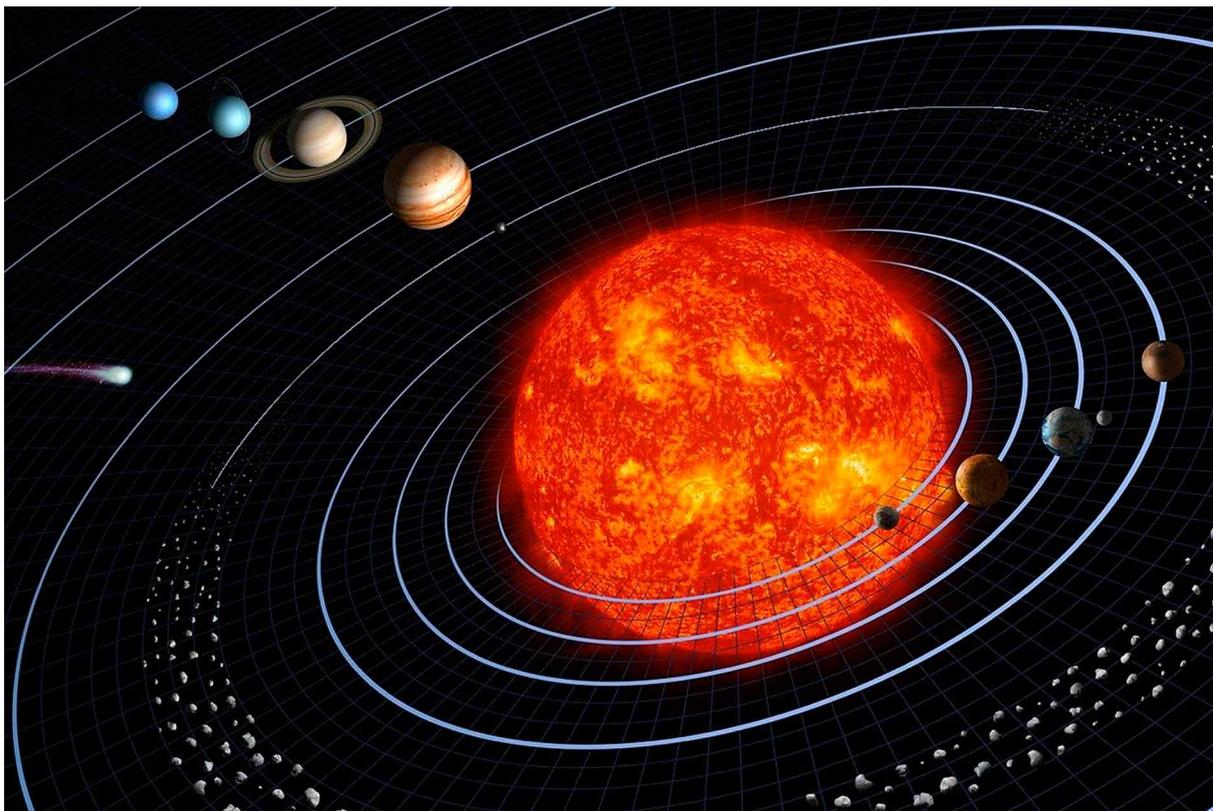


Galileo Galilei (1564-1642) realizando experimentos de caída de objetos desde la Torre de Pisa (según la leyenda). Fuente: elaboración propia.

## 5.3. Tener razón a toda costa: las hipótesis *ad hoc*

Las hipótesis *ad hoc*, del latín que significa «para esto», se añaden a una ley para conseguir que no quede invalidada por ciertos hechos que la contradicen. Al igual que ocurría con la cláusula *ceteris paribus*, las hipótesis *ad hoc* pueden resultar necesarias, porque las leyes son a veces tan concisas que no contemplan todas las circunstancias que pueden afectar al desarrollo de los acontecimientos. Evidentemente, se introducen solamente cuando la ley en cuestión resulta válida y útil en un contexto muy amplio, y quiere evitarse que sea rechazada por no cumplirse en alguna circunstancia muy concreta.

Por ejemplo, cuando se descubrió el planeta Urano se observó que su órbita alrededor del Sol no cumplía la ley del movimiento de los cuerpos celestes que orbitan alrededor del Sol establecida por Newton en el marco de su teoría de gravitación universal. El resto de planetas conocidos sí la obedecían, pero este único caso fallido sería suficiente para rechazar la ley. En lugar de ello, se propuso la hipótesis *ad hoc* de que existía un planeta desconocido hasta el momento cuya atracción gravitatoria perturbaba la órbita de Urano. Esta hipótesis era perfectamente contrastable empíricamente, lo que de hecho se llevó a cabo y condujo al descubrimiento de ese nuevo planeta, Neptuno. La historia se repitió con este último, ya que para explicar su órbita anómala se tuvo que proponer de nuevo una hipótesis



Licencia: dominio público. Fuente: Pixabay.

*ad hoc* sobre la existencia de un cuerpo celeste desconocido, que esta vez se trataba de Plutón. Las hipótesis *ad hoc* de este ejemplo contribuyeron a ampliar y fortalecer la teoría de gravitación universal de Newton.

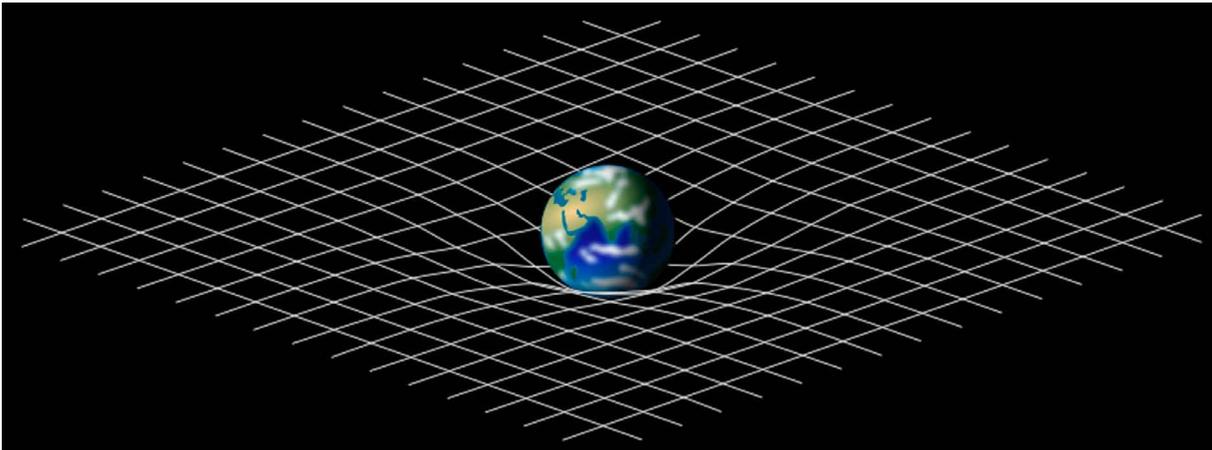
Lo cierto es que rechazar leyes que no concuerdan con las observaciones es uno de los fundamentos del método científico, que dejaría de ser útil si tras cada fracaso se recurriese a una hipótesis *ad hoc* (o «excusa *ad hoc*») para mantener la vigencia de la ley. De hecho, el abuso de hipótesis *ad hoc* es típico de las pseudociencias. Por ello, estas hipótesis deben introducirse en casos excepcionales, en número reducido, y han de ser susceptibles de comprobación experimental de manera independiente a la propia ley a la que acompañan, de manera que se puedan descartar una por una si no son correctas. Las leyes o teorías más simples, que vienen acompañadas de un menor número de hipótesis *ad hoc*, son más fáciles de falsar, en el sentido usado por Popper en el marco del contraste de hipótesis que vimos en el capítulo 3, y por tanto son las «más científicas».

La distinción entre la cláusula *ceteris paribus* y las hipótesis *ad hoc* es sutil en cuanto al papel que juegan en el contraste de hipótesis científicas. Cuando una teoría falla a la hora de explicar o predecir una observación y la razón es el incumplimiento de la cláusula *ceteris paribus*, la ley no se rechaza, sino que simplemente se considera que no podía aplicarse en esas circunstancias. Cuando el fallo no se debe al incumplimiento de la cláusula *ceteris paribus*, sino que es genuino porque la ley sí era aplicable, entonces la hipótesis *ad hoc* sirve para mantener su vigencia.

#### ► HISTORIAS DE LA CIENCIA. La errónea hipótesis *ad hoc* de Einstein que acabó siendo necesaria

Un ejemplo de hipótesis *ad hoc* fue la que introdujo el célebre físico alemán Albert Einstein (1879-1955) en relación a su teoría de la relatividad general. Se trata de una historia muy curiosa porque esa hipótesis trajo de cabeza al propio Einstein y sigue siendo una cuestión abierta hoy en día.

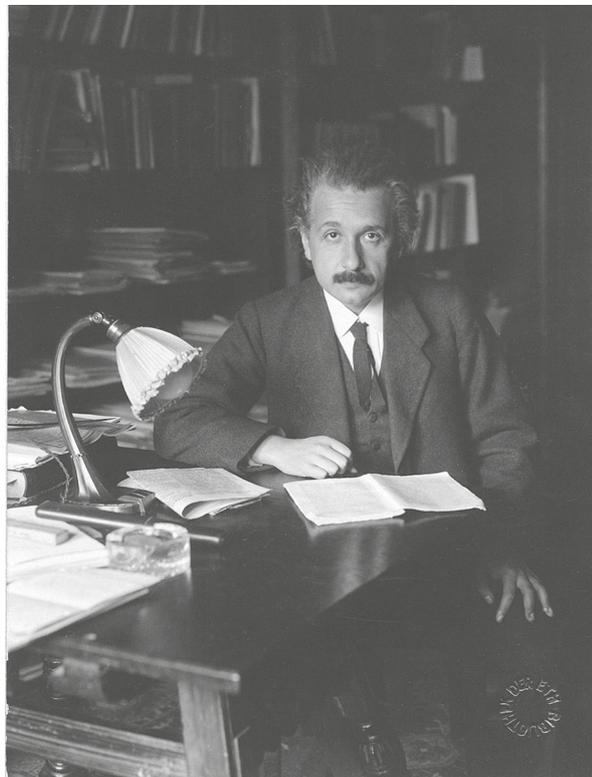
De manera simplificada, la ley física propuesta por Einstein afirma que los cuerpos con masa deforman el espacio vacío a su alrededor, de manera que los movimientos de otros cuerpos próximos tienen que adaptarse a esa nueva geometría. Este efecto, que solo resulta relevante en cuerpos con grandes masas (planetas, estrellas, etc.), permite reinterpretar la fuerza de atracción gravitatoria a partir de la geometría del espacio. Con esta idea tan revolucionaria se podían explicar fenómenos gravitatorios cotidianos, como la caída de una piedra al suelo, así como el movimiento de objetos astronómicos, incluyendo algunos detalles que la ley de gravitación universal de Newton, del siglo xvii, no podía describir.



Representación de la curvatura del espacio por la presencia de un cuerpo de gran masa, como la Tierra.  
Licencia: CC BY-SA 3.0. Fuente: Wikimedia Commons. Autor: Mysid.

La ley de Einstein también ofrecía una predicción para la evolución del universo en su totalidad: este tenía que disminuir o aumentar de tamaño conforme transcurría el tiempo. Einstein no se sentía cómodo con esta consecuencia, ya que consideraba obvio que el universo era estático, es decir, que su tamaño no cambiaba con el tiempo. Para conseguir que la ley fuese compatible con un universo estático, y por tanto que siguiera siendo válida, introdujo la siguiente hipótesis *ad hoc*: la existencia de «algo» en el universo que lo mantiene con un tamaño fijo, y que hace necesario sumar un cierto valor muy concreto a uno de los lados de la ecuación matemática que describe su ley. Este término añadido *ad hoc* se denominó constante cosmológica.

Observaciones astrofísicas posteriores demostraron que el tamaño del universo sí que cambia con el tiempo, aumentando (expandiéndose). Así pues, la creencia de Einstein acerca del tamaño fijo del universo era solo eso, una creencia, que no tenía base empírica y que se demostró equivocada. Cabe recordar que las hipótesis *ad hoc* pueden ser necesarias para especificar la



Albert Einstein. Licencia: dominio público.  
Fuente: Wikimedia Commons.

aplicación de una ley científica, y que, como ocurre con la propia ley a la que acompañan, pueden ser corroboradas o refutadas por los hechos empíricos. Einstein calificó la introducción de la hipótesis *ad hoc* de la constante cosmológica como el mayor error de su vida, nada menos.

Este es un ejemplo de hipótesis *ad hoc* que resultó refutada, pero la vida da muchas vueltas, y la ciencia, si cabe, aún más. Y además, grandes genios de la ciencia como Einstein aciertan en parte hasta cuando se equivocan. Observaciones astrofísicas más recientes han demostrado que el universo no solo se expande, sino que lo hace aceleradamente, es decir, cada vez más deprisa. Este hecho hace necesario introducir de nuevo la hipótesis *ad hoc* de Einstein, en una versión ligeramente distinta (esta vez no para explicar un universo estático sino uno con expansión acelerada). La nueva constante cosmológica se asocia a la presencia en el universo de una sustancia de naturaleza desconocida que causa repulsión (al contrario que la gravedad), denominada energía oscura. Pero esa ya es otra historia de la ciencia.

# Capítulo 6.

## Comienza la función: la explicación funcional

La explicación funcional consiste en justificar un hecho a partir de la función que tiene en un sistema. Un ejemplo de explicación funcional es que las aves tienen alas «porque» estas les permiten volar; la capacidad de vuelo es la función que las alas tienen en el organismo de las aves, y lo que explica su existencia.

Se trata de una explicación muy diferente a las descritas en los capítulos anteriores, en particular a la causal. De hecho, en cierto sentido son opuestas: en la causal la explicación del hecho reside en su causa, que siempre es anterior al hecho mismo, mientras que en la funcional la explicación reside en una consecuencia o efecto (la función), que tiene lugar *a posteriori*. La explicación funcional es muy común en biología y en ciencias afines, y también en ciencias sociales. Especialmente en estas últimas se puede confundir con explicaciones teleológicas o intencionales, que describiremos también en este capítulo.

### Conceptos clave:

- Explicación funcional, función, efecto, sistema, retroalimentación (*feedback*), retroalimentación causal.
- Explicación intencional, explicación teleológica, lamarckismo, heredabilidad de rasgos adquiridos, darwinismo, selección natural, adaptación al entorno, síntesis evolutiva.
- Estructuralismo (funcional-estructuralismo), falacia estructuralista (funcionalista), materialismo cultural.



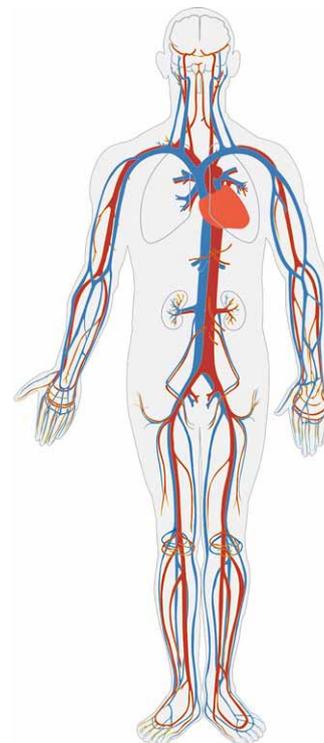
Fuente: elaboración propia con nubedepalabras.es.

## 6.1. El fin no justifica los medios: la función como efecto

La función es el papel que juega un proceso o elemento dentro de un sistema estructurado, y que contribuye a que permanezca en un cierto estado. Por ejemplo, podemos asignar una o varias funciones al corazón, que contribuyen a que el sistema del que forma parte, nuestro organismo, se mantenga vivo.

La explicación funcional se articula a través de los efectos o contribuciones que tiene el proceso o elemento sobre el sistema, y que resultan beneficiosos para él, en el sentido de que permiten mantener su estado. Así, la presencia del corazón en nuestro organismo se puede explicar funcionalmente porque permite la circulación sanguínea, siendo esa su función. Los efectos asociados a esa función, como la distribución de oxígeno y nutrientes o la regulación térmica, son beneficiosos (de hecho, imprescindibles) para mantener el organismo vivo.

Las funciones son consecuencias o efectos, y por tanto tienen lugar *a posteriori*. Así, el bombeo de sangre, que es la función del corazón, es evidentemente posterior a la existencia del propio corazón. Lo contrario ocurre con las causas, que suceden antes del hecho en cuestión. Esta diferencia hace que la explicación funcional se ponga en entredicho: ¿cómo puede la explicación de un hecho suceder después que el propio hecho? ¿no deberían ser todas las explicaciones de tipo causal, y no funcional? Para resolver este problema ha de introducirse un elemento adicional en la fundamentación de la explicación funcional, la retroalimentación causal, de la que hablaremos después.



Licencia: dominio público.  
Fuente: Pixabay.

### ▶ AMPLIACIÓN. La lógica de la explicación funcional

Desde el punto de vista lógico, una explicación funcional ingenua seguiría el siguiente razonamiento:

- Premisa 1: Si tiene lugar un hecho X (o está presente el elemento X) en un sistema, entonces ocurre el efecto Y en el sistema, que contribuye a que permanezca en cierto estado (en otras palabras, el efecto Y es la función de X).
- Premisa 2: El sistema muestra al efecto Y y se mantiene en un cierto estado en el tiempo.
- Conclusión \*: Tiene lugar el hecho o elemento X en el sistema.

Así es como la ocurrencia o existencia de X quedaría explicada a través de su efecto o función Y en un sistema. Para el ejemplo del corazón, este esquema de razonamiento quedaría del siguiente modo:

- Premisa 1: Si está presente el corazón, entonces se puede bombear sangre en nuestro organismo y este permanece vivo (la función del corazón es bombear sangre).
- Premisa 2: En nuestro organismo se bombea sangre y se mantiene vivo.
- Conclusión \*: Nuestro organismo contiene un corazón.

Así, la presencia del corazón quedaría explicada a través de su función, que es la de bombear sangre, manteniendo nuestro organismo vivo.

Ahora bien, los razonamientos anteriores son incorrectos desde el punto de vista lógico (razón por la cual las conclusiones están señaladas con un asterisco), ya que son ejemplos de la falacia de la

afirmación del consecuente. Ya habíamos encontrado estos razonamientos al hablar de la validación de hipótesis en el método científico, y habíamos explicado allí que el razonamiento es falaz porque la consecuencia podría deberse a condiciones distintas a la expresada en la relación condicional.

Así pues, es necesario precisar la naturaleza de la explicación funcional si se quiere sustentar firmemente en la lógica. Una alternativa consistiría en establecer una ley lógica según la cual, si es cierto que el elemento o proceso X tiene el efecto Y, entonces necesariamente ocurre X. De manera simbólica:  $(X \rightarrow Y) \rightarrow X$ . En otras palabras, esta ley establece que el simple hecho de que X produzca el efecto Y es razón suficiente para que se produzca X. Con esta sutil reformulación, la lógica detrás de la explicación funcional es totalmente correcta, pero ahora el peso de la prueba se traslada a la ciencia: es necesario demostrar que esa ley es cierta. Para ello, se ha de encontrar el mecanismo que hace que la producción del efecto Y por parte de X implique necesariamente la existencia de X: la retroalimentación causal.

## 6.2. La intención es lo que cuenta: explicación teleológica e intencional

El concepto de función de un elemento o proceso dentro de un sistema se puede confundir a veces con el de una finalidad u objetivo marcado para tal sistema. Esto último correspondería a una explicación teleológica, cuyo nombre proviene del griego *telos*, que significa fin u objetivo. La explicación teleológica es en general inadecuada en ciencias naturales, puesto que requiere establecer conscientemente unos objetivos, algo que no tiene sentido en la mayoría de sistemas, ya sean inertes o vivos.

En cambio, la explicación teleológica sí puede resultar apropiada en ciencias sociales, donde las acciones humanas pueden ir dirigidas a lograr ciertos objetivos, y se basan en la unión de dos circunstancias: la intención de lograr esos objetivos y la creencia en que una determinada acción es la adecuada para alcanzarlos. Se puede hablar entonces de explicación intencional: que una persona quiera actuar de cierta manera, empujada por sus deseos, necesidades o creencias, puede considerarse una explicación para el hecho de que la acción tenga lugar. De esta manera, en la explicación intencional están involucrados los conceptos de creencia y quizá también, aunque no necesariamente, el de racionalidad. Un ejemplo de explicación intencional es el de un profesor que realiza un curso sobre el método científico y sus procedimientos porque quiere ayudar a sus alumnos a llevar a cabo una investigación científica en el aula. El profesor cree que ese curso es un medio adecuado para contribuir a lograr sus fines, es decir, hay una intención detrás de la acción y una creencia en que tal acción contribuirá a lograr el objetivo.

En ocasiones se atribuyen intenciones a cualquier tipo de ser vivo o a la naturaleza en su conjunto. Esto ocurre en el mejor de los casos por abuso del lenguaje, y en el peor de los casos para justificar doctrinas anticientíficas, como el diseño inteligente de los seres vivos. Las explicaciones teleológica e intencional tampoco resultan adecuadas en ciencias sociales cuando se refieren a conceptos abstractos o colectivos en lugar de a individuos; por ejemplo, no tiene sentido dar explicaciones basadas en los objetivos e intenciones de la sociedad en su conjunto, o de la economía de mercado.



Licencia: dominio público. Fuente: Pixabay.

## 6.3. Con vuelta atrás: la retroalimentación causal

La explicación puramente funcional de un hecho (un elemento o un proceso) se apoya en los efectos beneficiosos que tiene para un sistema. Para diferenciarla claramente de la explicación teleológica cuando hay seres humanos involucrados, se requiere la ausencia de intencionalidad y que los sujetos no sean conscientes del beneficio sobre el sistema en cuestión.

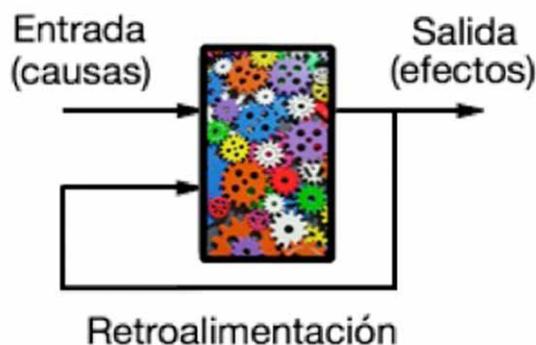
Además, resulta necesario establecer un mecanismo de retroalimentación causal entre el efecto sobre el sistema y el hecho que se pretende explicar funcionalmente. Es decir, de algún modo el efecto tiene que ser responsable de que el hecho aparezca o se mantenga en el tiempo. Este es un eslabón crucial en la explicación funcional, y de hecho la convierte en una explicación causal indirecta. El hecho en cuestión puede tener muchos efectos distintos sobre el sistema, pero solo uno de ellos, que es el que denominamos función, tiene capacidad explicativa sobre la existencia del hecho, y es el que va ligado a la retroalimentación causal. La mejor forma de entender este mecanismo es a través del caso paradigmático de la evolución biológica, como haremos enseguida.

### ▶ AMPLIACIÓN. El mecanismo de retroalimentación en la ciencia

El concepto de retroalimentación en general (también usado a menudo con su término en inglés, *feedback*) es muy importante en ciencia. Por ejemplo, cuando impartimos un curso se producen, con un poco de suerte, una serie de efectos en nuestros alumnos, que pueden cuantificarse a través de los resultados de sus exámenes. El profesor puede usar esa información para mantener o modificar su forma de dar clase: si las notas son muy bajas, puede decidir cambiar algún aspecto de la metodología. De esta manera, la acción del profesor se ve influida por los efectos que ha tenido su propia acción en ocasiones anteriores, esto es, hay una retroalimentación de información. Otro ejemplo es el del uso de la información sobre cifras de ventas o resultados de encuestas de satisfacción de los clientes por parte de las empresas.

La información retroalimentada no tiene por qué generarse siempre a través de una reflexión consciente. La retroalimentación automática es de hecho un mecanismo muy común en sistemas tanto naturales como artificiales. Un ejemplo muy simple es el de una calefacción con termostato, en el que se mantiene encendida la caldera hasta que la temperatura ambiente supera cierto valor, momento en el que el termostato la apaga; cuando, al cabo de cierto tiempo, la temperatura ambiente baja de cierto valor, el termostato enciende de nuevo la caldera. Así, el funcionamiento de la caldera se ve influido por el efecto que tiene su propio funcionamiento (que es el aumento de temperatura), de manera totalmente automática.

En el ámbito de la biología también ocurren innumerables procesos autorregulados por retroalimentación. Por ejemplo, cuando el nivel de glucosa en sangre sobrepasa cierto límite, se activa en el páncreas la producción de insulina, hormona que retira la glucosa del torrente sanguíneo para almacenarla en otros órganos; y cuando el nivel de glucosa ha descendido por debajo de cierto límite, el páncreas deja de producir insulina.



Fuente: elaboración propia.

#### 6.4. El cuello de las jirafas: la explicación funcional en biología

La llamativa longitud del cuello de las jirafas tiene una explicación funcional bastante obvia: es así de largo porque permite a las jirafas alcanzar el alimento de las copas de los árboles. Esta explicación reúne las condiciones expresadas anteriormente: se basa en los efectos beneficiosos que el hecho (el largo cuello) tiene para el sistema (el organismo de la jirafa), y no existe por parte del animal intencionalidad ni capacidad de dar lugar a ese hecho. Lo que falta es encontrar el mecanismo de retroalimentación causal, es decir, cómo y por qué las jirafas desarrollaron ese cuello tan beneficioso.



Licencia: dominio público. Fuente: Pixabay.

El naturalista francés Jean-Baptiste de Lamarck (1744-1829) propuso que los rasgos como el largo cuello de las jirafas surgían debido al uso intensivo y continuado de esa parte del organismo: cuanto más se estira el cuello para alcanzar el alimento de las copas de los árboles, más crece. Propuso también que ese rasgo surgido como adaptación al ambiente de un individuo se transmitía a su descendencia, ya que las jirafas ya nacían con un cuello largo. Aunque el aumento de la longitud del cuello debido a su continuo estiramiento fuese muy lento, casi imperceptible, al cabo de muchas generaciones, y debido al carácter hereditario del rasgo adquirido, podría dar lugar a un cambio notable que resultaría en una adaptación de la especie a su entorno.

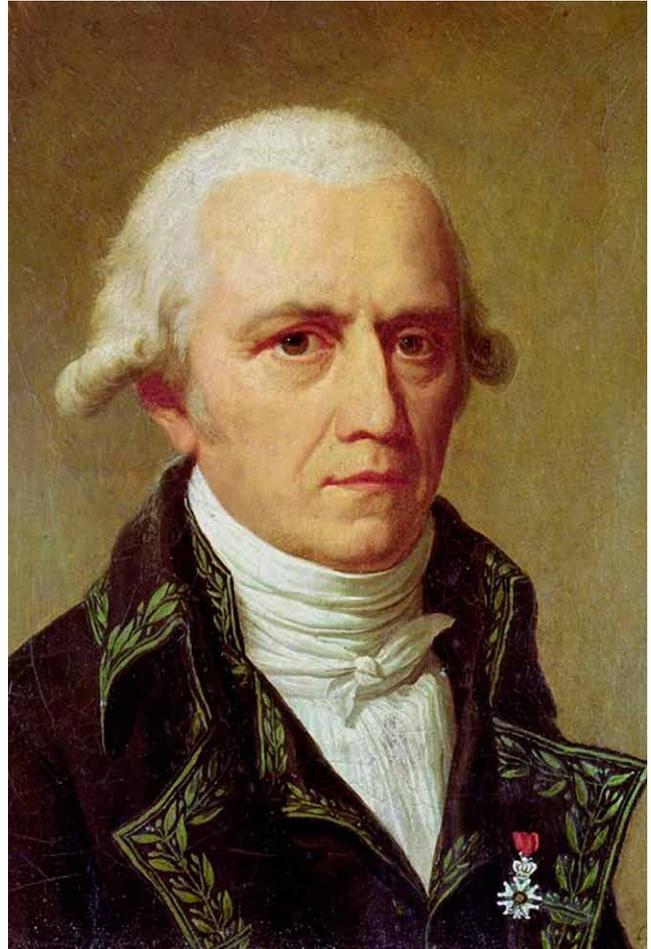
Estas pautas recopiladas por Lamarck (aunque basadas parcialmente en creencias más antiguas) constituyen una hipótesis para la retroalimentación causal requerida por la explicación funcional, ya que proporciona una cadena de causas y efectos razonable para la aparición y permanencia del elemento del sistema, en este caso el cuello largo en las jirafas. Resultan razonables porque la adquisición de rasgos por el uso continuo de ciertos órganos es un efecto habitual; por ejemplo, los individuos que llevan a cabo una intensa actividad física pueden desarrollar considerablemente algunas partes de su musculatura. Por otro lado, también es bien conocido el carácter hereditario de muchos rasgos de los individuos. Pero esta hipótesis razonable tiene que ser contrastada experimentalmente, y aquí es donde se encuentran discrepancias: solo son hereditarios aquellos rasgos presentes desde el nacimiento en el material genético, no los que adquiere el individuo durante su vida (sin entrar aquí en discutir mecanismos epigenéticos más complicados). Por ejemplo, la descendencia de los individuos muy entrenados en el gimnasio no nace con los músculos más desarrollados que el resto.

El argumento fallido de retroalimentación causal de Lamarck fue perfeccionado con la hipótesis de la selección natural del naturalista inglés Charles Darwin (1809-1882). Darwin partió del hecho de que las nuevas generaciones de una especie presentan desde el nacimiento una amplia variedad de rasgos, algunos de los cuales no estaban presentes en la generación anterior. La aparición de esos nuevos rasgos puede suponerse debida al puro azar, no dirigida inicialmente a producir ningún beneficio adaptativo a los

individuos, lo que resulta mucho más verosímil. Entre los individuos de la nueva generación, algunos de ellos presentarán rasgos que los hagan especialmente adaptados al medio en el que habitan, por lo que tendrán más probabilidades de sobrevivir y más descendencia, que es lo que se entiende como selección natural. Dado que esos rasgos beneficiosos venían determinados desde el nacimiento, forman parte del material hereditario y serán transmitidos a su descendencia. De este modo, y con cambios al azar en el material hereditario, los rasgos que presentan una ventaja adaptativa se van instalando y potenciando en la especie generación tras generación.

En el caso del cuello de las jirafas, el argumento específico, aunque simplificado, se puede describir como sigue. Hace millones de años, el antepasado de las jirafas actuales tuvo algunos descendientes con el cuello anormalmente largo, debido a un cambio accidental y aleatorio en el material hereditario. Esos individuos se alimentarían con mayor facilidad de las copas de los árboles, lo que les proporcionaría una ventaja adaptativa sobre el resto, especialmente si la vegetación baja escaseaba. En tales condiciones favorables, esos individuos tendrían más posibilidades de reproducirse, transmitiendo su cuello anormalmente largo a su descendencia. Con el paso de las generaciones, el rasgo acabaría implantándose en la especie.

Conviene resaltar cómo, una vez conocido el mecanismo completo, una explicación funcional se puede reducir en realidad a una causal. La explicación funcional «las jirafas tienen un cuello muy largo porque les sirve para alcanzar comida de los árboles» se puede expresar ahora como: «las jirafas tienen el cuello muy largo porque la aparición al azar de ese rasgo hereditario supuso una ventaja adaptativa en el proceso de selección natural, de modo que, con el transcurso de las generaciones, ese rasgo acabó formando parte de todos los individuos de la especie».



Jean-Baptiste de Lamarck. Licencia: dominio público. Fuente: Pixabay.

### AMPLIACIÓN. La síntesis evolutiva

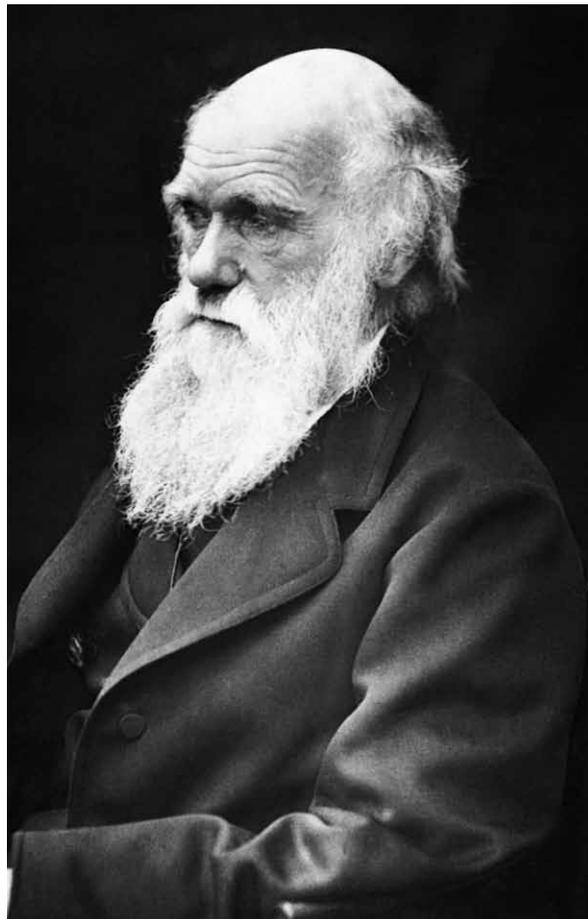
El mecanismo de retroalimentación causal darwinista, basado en la selección natural, es más detallado que el lamarckiano y no presenta el problema de la transmisión hereditaria de un rasgo adquirido durante la vida del individuo. De todos modos, la ciencia debe aclarar los detalles que siguen quedando pendientes. Por ejemplo, ¿es posible que la descendencia presente rasgos anómalos, ausentes en los progenitores, y que surjan de manera aleatoria? ¿es posible que esas anomalías se transmitan a su vez a la descendencia de la descendencia?

Esos detalles se resuelven con el neodarwinismo, también llamado síntesis evolutiva, que es la descripción más moderna del mecanismo. La aparición de nuevos rasgos en la descendencia se debe a que el material genético de los progenitores puede sufrir cambios al azar, denominados mutaciones. La descendencia puede presentar rasgos asociados a esas mutaciones en los genes y

además pueden transmitirlos a su propia descendencia. Aunque quedan muchos detalles por conocer acerca de los procesos bioquímicos que expresan la información codificada en los genes, la replican o la alteran dando lugar a mutaciones, los avances en el campo son continuos. Pero, en términos generales, la síntesis neodarwinista representa ya un mecanismo muy completo, verosímil y con amplia base empírica para dotar a las explicaciones funcionales biológicas de la retroalimentación causal necesaria.

### HISTORIAS DE LA CIENCIA. Darwin y el pico de los pinzones

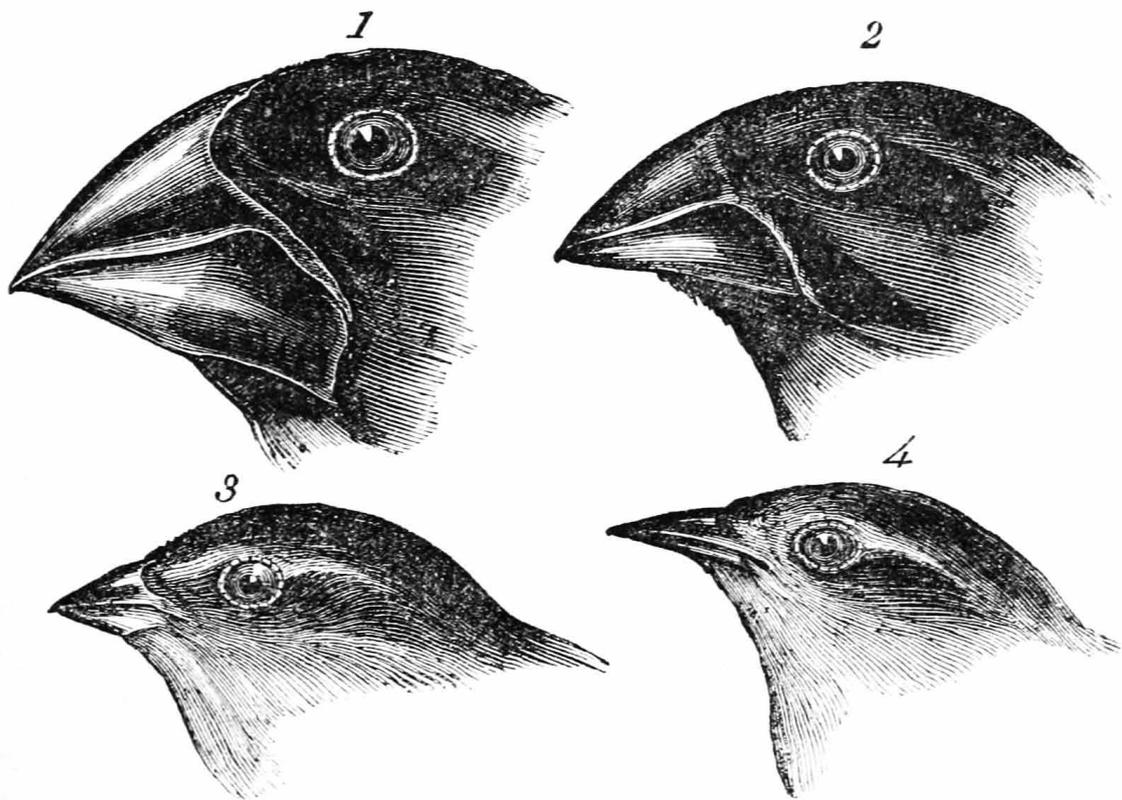
En 1831 Charles Darwin embarcó en el HMS Beagle, nave de la Marina Real Británica que emprendía un viaje alrededor del mundo para cartografiar regiones poco conocidas en aquella época, como las costas de América del Sur y Oceanía. Darwin viajaba en calidad de naturalista, aunque sin percibir un sueldo por ello. Durante los casi cinco años de viaje, recogió numerosos especímenes de animales y vegetales, así como de fósiles, para que fuesen estudiados por expertos al regresar a Inglaterra, aunque él también tomó una gran cantidad de notas durante el viaje. En las islas Galápagos, un archipiélago del Pacífico situado frente a la costa de Ecuador, Darwin recogió ejemplares de aves que se parecían mucho entre sí excepto por algunos rasgos, principalmente el tamaño y la forma del pico. Se trataba de diferentes especies de pinzones (hoy llamados pinzones de Darwin), cada una de las cuales habita en una isla diferente del archipiélago y tiene el pico adaptado a la principal fuente de alimentación en cada una de ellas (semillas, frutos, insectos...). En su recopilación de hallazgos durante el viaje (*Journal of researches into the natural history and geology of the countries visited during the voyage of HMS Beagle round the world, under the Command of Capt. Fitz Roy, R.N.*, de 1845), Darwin escribió a este respecto:



Charles Darwin. Licencia: dominio público.  
Fuente: Wikimedia Commons.

Cuando se considera esa gradación y esa diversidad de configuración en un pequeño grupo de pájaros muy afines unos a otros, realmente podría creerse que en virtud de una pobreza originaria de aves en ese archipiélago, una sola especie se ha modificado para alcanzar objetivos diferentes.

Observaciones como estas fueron el origen de las ideas de Darwin sobre la adaptación al entorno mediante selección natural y sobre la diferenciación entre especies, es decir, fueron el germen del mecanismo de retroalimentación causal que sustentaría a partir de entonces la evolución biológica.



1. *Geospiza magnirostris*.  
3. *Geospiza parvula*.

2. *Geospiza fortis*.  
4. *Certhidea olivacea*.

Pinzones de las Galápagos dibujados por Darwin. Licencia: dominio público. Fuente: Wikimedia Commons.

### 6.5. Fallo estructural: la falacia estructuralista

Las explicaciones funcionales en biología poseen actualmente, como hemos visto, una base muy firme y un carácter plenamente científico, porque el mecanismo de retroalimentación causal está bien establecido. Gracias a ello, se puede hablar de función, beneficio, utilidad, etc. sin que ello implique una intencionalidad o un diseño premeditado por parte de un ente superior (que correspondería a la hipótesis del diseño inteligente).

La explicación funcional es también muy habitual en ciencias sociales, especialmente en corrientes como el estructuralismo o el funcional-estructuralismo, en las que la sociedad se describe como un sistema en el que las instituciones, las normas, las costumbres, etc., tienen una función específica que contribuye al mantenimiento y estabilidad de la estructura social. No debe resultar extraño que uno de los impulsores de esta corriente, el filósofo y biólogo británico Herbert Spencer (1820-1903), compare la sociedad con el organismo humano; ambos funcionan gracias a los órganos que los constituyen, que en el caso de la sociedad serían los elementos mencionados antes (instituciones, normas, etc.). Es interesante señalar que este mismo filósofo fue quien acuñó la conocida expresión de «supervivencia del más apto», en relación al mecanismo de la selección natural, que también aplicó a los individuos o grupos dentro de la sociedad en el marco del darwinismo social.

La explicación funcional en ciencias sociales, como ocurre en las demás ciencias, no es plena si no se identifica el mecanismo de retroalimentación causal involucrado. En ciencias sociales este defecto da lugar a la denominada falacia estructuralista o falacia funcionalista: se establece como causa de las acciones de los individuos los efectos beneficiosos que proporcionan a la sociedad en su conjunto, pero sin identificar el mecanismo que sustenta esa causalidad.

No se contempla que la causa de esas acciones individuales sea la búsqueda intencionada del bien común (explicación intencional), porque en el marco funcionalista los individuos ni siquiera son conscientes de los efectos beneficiosos de sus acciones, y por tanto éstas no son intencionadas ni van dirigidas a ningún objetivo. Si las acciones continúan llevándose a cabo, y no de manera intencionada, es porque existe un mecanismo de retroalimentación causal, pero en ciencias sociales no se ha identificado uno tan general como el de la biología. Para cada explicación funcional en ciencias sociales, si no se quiere caer en la falacia funcionalista, es necesario identificar un mecanismo de retroalimentación causal particular, y demostrar que efectivamente actúa de ese modo, lo que representa una gran dificultad añadida a la investigación científica en ciencias sociales.

### HISTORIAS DE LA CIENCIA. Vacas sagradas

Como ejemplo de explicación funcional en ciencias sociales, se puede recurrir al análisis que hizo el conocido antropólogo norteamericano Marvin Harris (1927-2001) sobre el carácter sagrado de las vacas en la India. Esta creencia prohíbe su sacrificio y el consumo de su carne, incluso aunque las necesidades nutricionales de la población no estén plenamente cubiertas. ¿Por qué los individuos actúan de un modo tan opuesto a sus propios intereses inmediatos en cuanto a una necesidad básica como es la alimentación?

Según Harris, la función o utilidad de cuidar y mantener con vida a las vacas reside en que sirven para arar campos de cultivo, contribuyen a la salubridad del entorno al consumir desperdicios, y proporcionan leche, fertilizante para los campos en forma de estiércol y combustible para cocinar a partir del estiércol seco. Además, contribuyen a la supervivencia, aunque sea en los límites de la pobreza, de cientos de millones de agricultores sin otros recursos, y a mantener el orden social, aunque se trate del injusto sistema de castas. Estas explicaciones constituyen lo que en antropología se denomina perspectiva *etic*, que es la que sostiene el investigador cuando analiza e intenta explicar racionalmente los hechos y acciones que observa, pero en los que no interviene.

Sin embargo, la población india no es en general consciente de estos beneficios, y en cualquier caso es dudoso que los considerase más ventajosos que el inmediato consumo de la carne. Descartada la explicación intencional, sería necesario sustentar firmemente la explicación funcional de partida con un mecanismo de retroalimentación causal, del que Harris no da detalles. Se puede suponer, por ejemplo, un mecanismo análogo al de la selección natural en biología, de modo que las comunidades indias que dejaron de sacrificar vacas obtuvieron tales ventajas socioeconómicas (como las descritas arriba), que prosperaron más que las demás y aumentaron más rápidamente su población, que además heredaría culturalmente esa costumbre. Esta no es la única hipótesis causal posible, y de hecho se le pueden atribuir muchas lagunas, pero sirve como ejemplo. En cualquier caso, cualquier hipótesis al respecto debería ser contrastada empíricamente, algo que es muy difícil de llevar a cabo en ámbitos como este.

Quizá hayamos descartado demasiado pronto la explicación intencional en el caso de las vacas sagradas. La prohibición del consumo de carne de vaca se enmarca en una creencia religiosa, como indica su atributo de sagradas. En efecto, las vacas son consideradas deidades en la religión hinduista, y, siempre según esta creencia, su sacrificio conlleva graves consecuencias para el responsable, tanto en esta vida como en sus futuras reencarnaciones. Bajo esas amenazas, es lógico que los creyentes se abstengan incluso de incomodar al ganado vacuno, que campa a sus anchas por aldeas y ciudades. Esto constituye una explicación intencional del fenómeno, pues el comportamiento de cada individuo se basa en una decisión consciente, sustentada en creencias religiosas. Este punto de vista se denomina perspectiva *emic*, que es la que sostienen los sujetos que realizan las acciones y es por tanto subjetiva (en contraste con la *etic*).

Es dudoso que una explicación intencional de ese tipo sea lo suficientemente científica: ¿una creencia religiosa o superstición es motivo suficiente para realizar una acción, y por tanto debe considerarse

una explicación (intencional) de esa acción? ¿no existirá en realidad otra causa subyacente más mundana? Para antropólogos como Harris, creencias religiosas o costumbres culturales de este tipo «disfrazan» la verdadera explicación, que es mucho más terrenal. Ese carácter se denomina «material», de donde surge el nombre de la corriente antropológica a la que pertenece Harris, el materialismo cultural. Como hemos visto antes, las explicaciones materialistas pueden seguir siendo incompletas si no aportan ni contrastan empíricamente un mecanismo de retroalimentación causal adecuado. Además, se puede interpretar que los comportamientos sociales que se pretenden explicar se acaban «institucionalizando» en forma de creencia religiosa o costumbre prácticamente obligatoria, a veces incluso en forma de preceptos legales, lo que desplaza la explicación funcional en favor de una intencional basada en la religión o la legalidad. Es evidente que la investigación científica en ciencias sociales puede requerir explicaciones muy complejas, que a veces invaden ámbitos ajenos a la propia ciencia.

Para saber más: M. Harris (2007) *Bueno para comer*. Madrid: Alianza Editorial.



Licencia: dominio público. Fuente: Pixabay.

# Capítulo 7.

## Fantasía y realidad: los límites de la ciencia

La postura según la cual la ciencia describe y explica una realidad externa y objetiva, ya sea la del mundo natural (ciencias naturales) o la de las actuaciones y relaciones humanas (ciencias sociales), se denomina realismo científico.

El realismo puede referirse a los conceptos que maneja la ciencia, a las leyes que expresan relaciones entre esos conceptos, o a todo a la vez. Por ejemplo, en la teoría del electromagnetismo en física aparecen conceptos como carga eléctrica negativa y carga eléctrica positiva, y también forma parte de la teoría la ley que dice que cargas de distinto signo se atraen y las del mismo signo se repelen. No todos estos elementos corresponden necesariamente a entidades que existen realmente. También cabe la posibilidad de plantearse, como lo hicimos en capítulos anteriores, si las leyes científicas expresan relaciones causa-efecto reales o solamente coincidencias habituales de unos hechos con otros.

Un realismo científico completo encuentra diversos obstáculos. Por una parte, la realidad puede ser cada vez más compleja conforme profundizamos en ella, lo que la haría incomprendible para el ser humano a partir de cierto punto. Por otro lado, es posible que nuestros sentidos, ya sea por sí solos o a través de instrumentos, solo registren una pequeña parte de la realidad, y que haya otra parte de alcance desconocido a la que nunca tendremos acceso. También hay que considerar la posibilidad de que, como en toda actividad humana, la ciencia se vea afectada por influencias psicológicas, culturales y sociales, que hagan que los resultados obtenidos no reflejen fielmente la realidad.

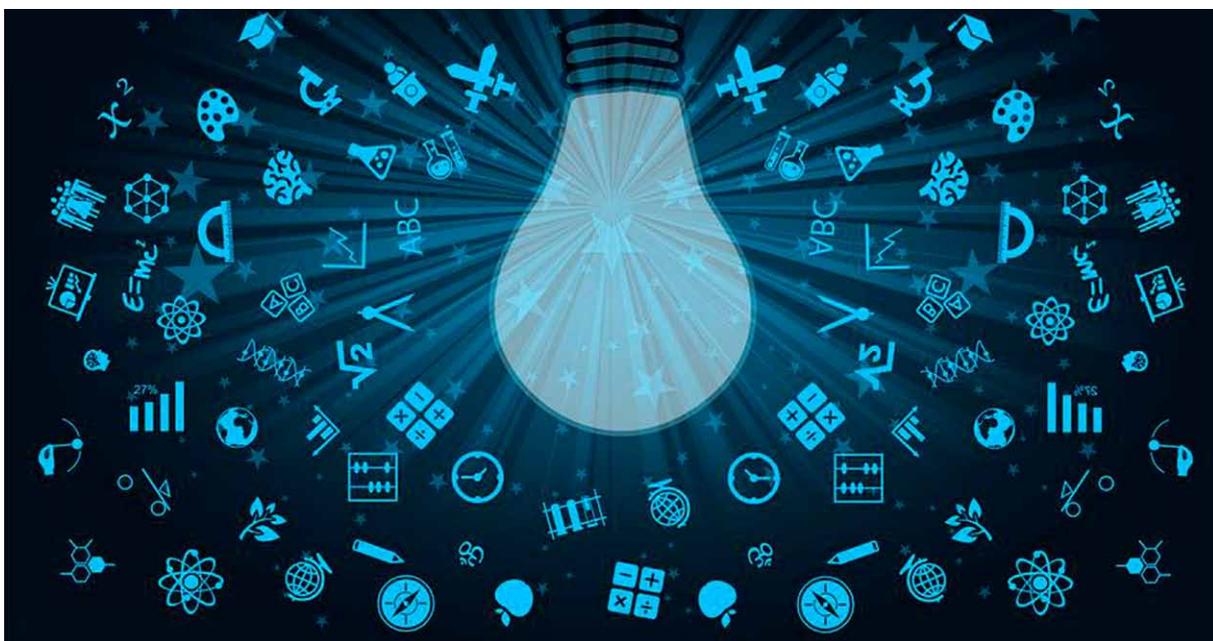
Otro punto de vista distinto al realismo, denominado instrumentalismo, considera la ciencia únicamente como una herramienta predictiva, capaz de generar leyes que nos permitan saber con antelación cuáles serán los resultados de una observación o experimento (ya sea con certeza o asignando probabilidades). A pesar de esa capacidad predictiva, se ignoraría qué realidad hay detrás de ello. El avance de la ciencia iría dirigido entonces principalmente a las necesidades humanas de predicción y control de su entorno, es decir, se vería influido por factores psicológicos, culturales y sociales.

Lo que parece comúnmente aceptado hoy día es alguna forma de naturalismo científico, que establece que la propia ciencia en su conjunto (sus fundamentos, sus métodos, el conocimiento que genera) ha de considerarse un fenómeno natural y social (ya que se lleva a cabo por seres humanos), y por tanto ha de ser analizada con una componente empírica, y no simplemente teórica, abstracta y aislada de todo lo demás.

### ▶ AMPLIACIÓN. Ciencia y creencia

Los fundamentos de la ciencia que han sido introducidos hasta ahora no solo son necesarios para llevar a cabo con éxito una investigación científica que avance en el conocimiento. También son imprescindibles para poder valorar en su justa medida los conocimientos científicos con los que ya contamos, y que forman parte de la sociedad actual tanto en su vertiente tecnológica como en la de creencias compartidas (cultura científica). En otras palabras, conocer los fundamentos de la ciencia sirve para establecer las limitaciones de la ciencia y para distinguir lo que es la ciencia de lo que no lo es.

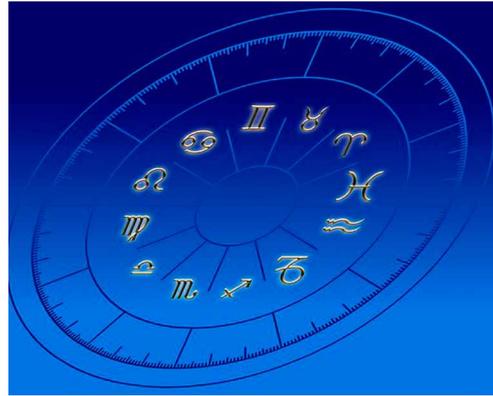
Como ya mencionamos al inicio, el conocimiento científico se distingue del conocimiento ordinario, y lo hace principalmente porque el primero se obtiene aplicando el método científico, basado en la observación o experimentación por un lado (excepto en las ciencias formales) y en la racionalidad de la lógica por otro. La filosofía también emplea la racionalidad, pero no como instrumento para avanzar en el conocimiento de la realidad externa, sino para reflexionar sobre asuntos propiamente humanos como los fines, los valores, el sentido o las ideologías. Estas cuestiones también pueden ser abordadas en el ámbito de la religión, no desde la racionalidad sino desde la fe o la creencia. A pesar de las distinciones que se puedan hacer entre ciencia, filosofía y religión, las fronteras son difusas y algunos ámbitos invaden lentamente, con el paso de los siglos, el territorio de otros.



Licencia: dominio público. Fuente: Pixabay.

Mención aparte merecen las pseudociencias, cuyo contenido, aunque pretende imitar y suplantar al de las ciencias, no se obtiene de la aplicación del método científico, sino de supersticiones, prejuicios o invenciones no contrastadas. Las pseudociencias suelen caracterizarse por su incapacidad para predecir hechos concretos y contrastables. Cuando se producen esas predicciones, son demasiado vagas como

para poder rechazarlas de manera contundente en caso de discrepancia con los hechos observados. Recordemos que en el método científico moderno, especialmente tras las aportaciones de Popper, resulta primordial construir hipótesis muy específicas que se puedan rechazar fácilmente tras su comparación con las observaciones, es decir, todo lo contrario a lo que ocurre en las pseudociencias. Estas hacen además un empleo abusivo de las hipótesis *ad hoc*, aunque no usen esta nomenclatura; cuando un pseudocientífico se ve forzado a elaborar una predicción concreta y esta contradice las observaciones, recurre a excusas en forma de condiciones, cada vez en mayor número y más enrevesadas, que se deberían haber cumplido para que su predicción fuese correcta, pero que no se han dado en esa ocasión.



Licencia: dominio público. Fuente: Pixabay.

Desgraciadamente no es fácil luchar contra las pseudociencias desde la propia ciencia. Si bien las pseudociencias no son capaces de sostener científicamente la veracidad de sus enunciados, la ciencia auténtica tampoco está capacitada para demostrar definitivamente su falsedad o su imposibilidad. Recordemos que el proceder de la ciencia, según el punto de vista moderno desde Popper, no consiste en la búsqueda de la certeza, sino en la permanente crítica y revisión de lo que se da por sabido. Eso implica que no se puede rechazar *a priori* de manera dogmática ningún postulado de las pseudociencias, a pesar de que estas sí se caracterizan por el dogmatismo. Si además se intenta remplazar el contenido de una pseudociencia con el de la ciencia, se tiene que el de esta última no es necesariamente el verdadero y definitivo. Como sabemos, el contraste de las teorías científicas con la observación no conlleva nunca la certeza absoluta sobre su verdad, sino solamente su confirmación provisional tras continuas verificaciones, o su aceptación también provisional tras superar continuas pruebas de falsación. Dicho de otro modo, en un debate entre un científico y un pseudocientífico, el primero no siempre tiene las de ganar, principalmente porque su estrategia es la duda y la autocrítica, mientras que la del segundo es la defensa vehemente, aunque sea de sinsentidos, generalidades o medias verdades. Estas conclusiones pueden resultar decepcionantes, pero es así como funciona la ciencia, y es esencial conocer sus limitaciones cuanto antes y con la mayor claridad posible, para poder emplear las armas con las que sí cuenta con la mayor eficacia posible.

Un ejemplo de pseudociencia es la astrología, que afirma que la personalidad y los acontecimientos futuros de cada ser humano vienen determinados por la posición de los cuerpos celestes el día de su nacimiento, que es lo que define su signo del zodiaco. Evidentemente, no aporta ninguna prueba experimental de cómo esos astros influyen en nosotros: ¿lo hacen a través de la fuerza gravitatoria o a través de otra fuerza desconocida para la ciencia? ¿cómo afectan sus posiciones a nuestro organismo el día de nuestro nacimiento? ¿no tendría más sentido que lo hicieran sobre las reacciones bioquímicas en el momento de la concepción o durante el desarrollo fetal? Como se puede comprobar, se generan más preguntas que las que se responden. Por otro lado, cuando se leen las predicciones del horóscopo semanal, ¿son concretas y específicas? Resulta difícil que así sea, cuando supuestamente han de aplicarse a los cientos de millones de personas que comparten signo del zodiaco. De hecho, suelen ser tan vagas que, ocurra lo que nos ocurra durante la semana, muchas encajarán con lo vaticinado. Una ciencia, por el contrario, se esfuerza en hacer predicciones muy concretas con la (trágica) esperanza de que fallen estrepitosamente, para poder falsar cuanto antes la hipótesis incorrecta.

En definitiva, la mejor manera de luchar contra las pseudociencias y las supersticiones es a través de la cultura científica. Esta no consiste solamente en poseer una cierta visión de conjunto y actualizada sobre el conocimiento científico (y no únicamente el tecnológico), sino también en conocer los fundamentos, métodos y limitaciones de la ciencia, que es lo que se ha intentado transmitir en esta primera parte.

## HISTORIAS DE LA CIENCIA. Qué es la ciencia según un juez de Arkansas

El caso judicial McLean v. Arkansas Board of Education trata sobre una denuncia presentada ante la Justicia del Estado de Arkansas en Estados Unidos por padres de alumnos, grupos religiosos, científicos, etc. contra una ley estatal (Act 590) que obligaba a las escuelas públicas del estado a enseñar en igualdad de condiciones el creacionismo y la teoría de la evolución, o como se denominan en esa ley, la «ciencia de la creación» y la «ciencia de la evolución». La primera de ellas recoge lo que dice la Biblia, interpretado al pie de la letra, acerca del origen de los seres vivos en general y de los seres humanos en particular. La clave del asunto es que ese contenido se considera científico, al igual que el de la teoría de la evolución descrita en el capítulo anterior, y por tanto ambas deben ser transmitidas en las mismas condiciones en las escuelas.

Por extraño que parezca, buena parte del juicio se dedicó a establecer qué se entiende por ciencia, para poder sustentar la decisión de si el creacionismo es o no científico. Así se refleja en la sentencia, de enero de 1982, de la que se reproduce a continuación un fragmento traducido (sección IV(C)), que trata sobre la definición de ciencia y que encaja muy bien con lo que hemos visto hasta ahora, además de aportar otros matices interesantes:

Varios testigos sugirieron definiciones de ciencia. Una definición descriptiva fue que la ciencia es aquello que es «aceptado por la comunidad científica» y que es «lo que hacen los científicos». La implicación obvia de esta descripción es que, en una sociedad libre, el conocimiento no requiere del visto bueno de la legislación para convertirse en ciencia.

Más en concreto, las características esenciales de la ciencia son:

- (1) Está guiada por la ley natural.
- (2) Tiene que ser explicativa refiriéndose a la ley natural.
- (3) Puede ser contrastada con el mundo empírico.
- (4) Sus conclusiones son provisionales, es decir, no son necesariamente la última palabra.
- (5) Es falsable.



Licencia: dominio público. Fuente: Pixabay.

Para que nadie se quede con la intriga: el juez dictaminó que el creacionismo no es ciencia, sino religión, y que, según la Constitución de Estados Unidos, debe estar separada del Estado. Por tanto, la ley Act 590 era anticonstitucional y no podía aplicarse.

## PARTE II



# Capítulo 8.

## Metodologías de investigación

La idea de partida de una investigación científica puede tener orígenes muy diversos: investigaciones anteriores realizadas por otras personas o por nosotros mismos que pueden ser ampliadas o profundizadas, conversaciones entre investigadores o en la vida cotidiana, observaciones del entorno natural o de las relaciones sociales que no concuerdan con las teorías vigentes, etc. Cualquier tema que centre una investigación científica debe ir asociado a hechos problemáticos, sobre los que interesa conocer las causas, o a relaciones dudosas entre varios hechos, que pueden deberse o no a nexos causales. A menudo, el tema de la investigación puede plantearse como una pregunta.

De los innumerables hechos desconocidos o problemáticos que pueden surgir, no todos tienen el mismo interés. Realizar una investigación requiere tiempo, esfuerzo y, en ocasiones, también recursos económicos, por lo que será necesario plantearse la motivación o justificación para llevar a cabo el estudio. La motivación que encuentren los investigadores para llevar a cabo el estudio podrá ser tenida en cuenta también por los receptores potenciales, que deben decidir si dedican su tiempo a leer la memoria final o la publicación resultante. Por esa razón es importante reflejar las motivaciones y justificaciones de la investigación al publicar los resultados.

La motivación puede ser puramente teórica, basada en el avance del conocimiento científicamente establecido sobre un tema, o bien de carácter práctico, basada en las posibles aplicaciones tecnológicas o sociales de la investigación. Cuanto más amplios, diversos, profundos o novedosos sean los resultados esperados de una investigación, mejor motivada estará.

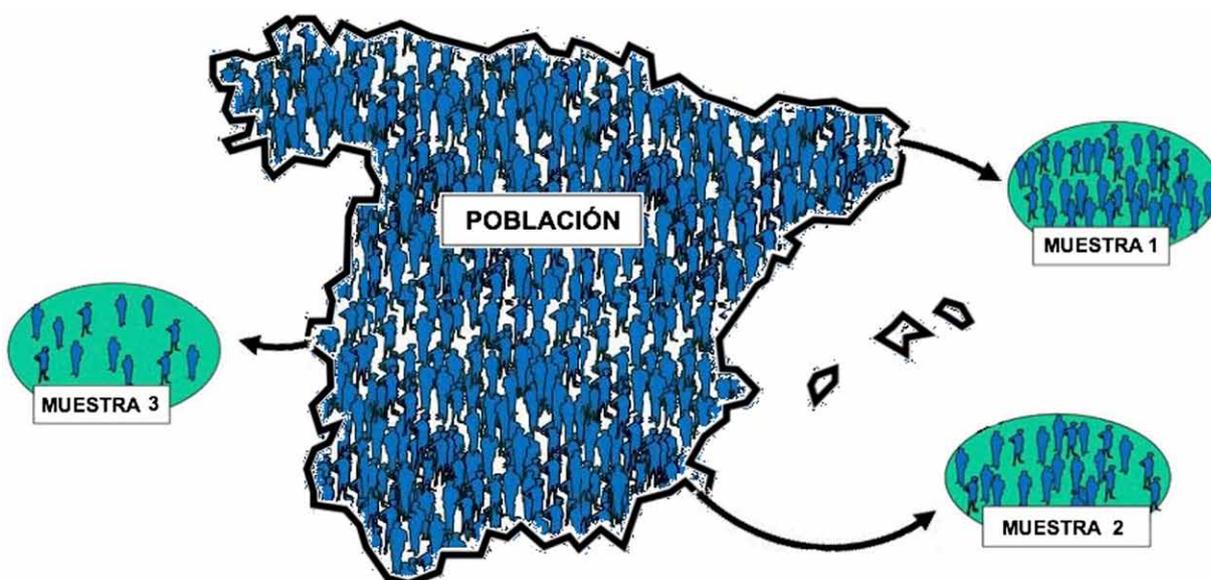
Por muy interesante que sea el tema que tenemos en mente, no tiene sentido iniciar una investigación sobre él si no se tiene claro cómo podemos contribuir a su avance. Para ello, el primer paso es familiarizarse con lo que se ha investigado al respecto en el pasado, llevando a cabo una revisión bibliográfica o mediante conversaciones con investigadores del área. En este punto aún no ha comenzado la investigación en sí, pues estamos valorando en primer lugar si tendrá sentido. Si finalmente se inicia la investigación, la revisión bibliográfica llevada a cabo resultará imprescindible

para elaborar el estado de la cuestión del tema (*state-of-the-art*), es decir, la recopilación de todo lo que se sabe sobre él. Sin esa información, se corre el riesgo de repetir una investigación ya realizada en el pasado, o de incurrir en los mismos errores cometidos por estudios anteriores.

A partir de la idea o problema de partida, debidamente motivado, y una vez establecido el estado de la cuestión, deben definirse los objetivos de la investigación, tanto generales como específicos. Entre estos últimos usualmente figura la comprobación de una hipótesis de trabajo, que es la forma en que se define y concreta el tema de la investigación en el método científico. La hipótesis es el enunciado que se quiere contrastar con los datos recogidos y analizados en la investigación. Si el tema de la investigación se formula como pregunta, la hipótesis constituye una respuesta provisional.

La propia elaboración de la hipótesis conlleva emplear conceptos bien definidos, que a su vez han de asociarse a variables, que son propiedades o características de los elementos de estudio y que pueden tomar distintos valores (al menos dos). Algunos ejemplos de variables y sus posibles valores son el color, que puede tomar valores como «rojo», «verde» o «azul»; la edad, que puede tomar valores como «1 año», «10 años» o «50 años»; o el nivel cultural, con valores como «alto», «medio» o «bajo», cuya definición es *a priori* más arbitraria que en los ejemplos anteriores.

Los elementos de estudio pueden ser objetos materiales o eventos físicos, como ocurre en las ciencias naturales, o bien acontecimientos, procesos, estados u otros entes abstractos, que aparecen habitualmente en ciencias sociales (por ejemplo, «delincuencia», «bienestar», «nivel cultural», «aprendizaje», etc.). Los elementos a los que tenemos acceso y podemos medir en el transcurso de una investigación constituyen la muestra. Sin embargo, el objetivo de muchas investigaciones es conocer las propiedades del conjunto de elementos, mucho más amplio, del que se ha extraído la muestra, y que constituye la población. Las muestras son, por tanto, subconjuntos de la población, generalmente mucho más pequeños que esta. Gracias a ese tamaño reducido, se pueden medir los valores que toman las variables en todos los elementos de la muestra, algo que suele ser inviable para todos los elementos de la población, ya sea porque su cantidad es inabarcable o porque muchos de los elementos pueden no estar ni siquiera accesibles a los investigadores. Los términos «muestra» y «población» cobran especial relevancia en la metodología cuantitativa y en los procedimientos estadísticos asociados a ella, como veremos posteriormente.



Fuente: elaboración propia.

## EJEMPLO

En una investigación se seleccionan cien personas y se miden sus estaturas. Ese conjunto de datos, y todas las operaciones que se puedan hacer con ellos, solamente aportan información sobre esa muestra concreta de personas. La investigación tratará posiblemente sobre un conjunto de individuos mucho más amplio que la muestra, por ejemplo, sobre la talla de los europeos en la actualidad. Esa sería la población de la que habría que extraer la muestra, que debería contener personas de diferentes países de Europa. Si la investigación tratase sobre la talla de mujeres adolescentes españolas, la población ya no incluiría varones, ni mujeres no adolescentes, ni personas de otros países europeos. La muestra de cien personas debería ser extraída de esta otra población, más restringida que la anterior, pero todavía demasiado numerosa como para ser estudiada al completo.

Los conceptos y procedimientos anteriores forman parte de la metodología científica en sentido general, que implementa el método científico que se describió en la parte I. Profundizando algo más, pueden distinguirse dos tipos de metodologías distintas: la cuantitativa y la cualitativa. La primera de ellas es típica de las ciencias naturales, mientras que la segunda lo es de las ciencias sociales, por sí sola o en combinación con la cuantitativa.

### 8.1. Metodología cuantitativa

La metodología cuantitativa se basa en la descripción de conceptos mediante variables que se asocian a valores numéricos. La medida u observación de las variables ha de ser sistemática y estructurada según protocolos bien definidos, y debe establecerse una relación estricta entre los conceptos que se quiere analizar y las variables que han de medirse para ello. El procedimiento completo, que toma el nombre de operacionalización, sigue la siguiente secuencia:

conceptos → traducción a variables medibles → asignación de valores a las variables

Los datos numéricos obtenidos en las medidas pueden ser sometidos a cálculos estadísticos para ordenarlos, para identificar las relaciones entre ellos o para describirlos mediante un conjunto reducido de valores numéricos. También se puede aplicar la estadística para obtener, a partir de una cierta cantidad de medidas tomadas en una muestra, propiedades de la población al completo de la que se ha extraído.

La finalidad de los métodos estadísticos es formular explicaciones de tipo causa-efecto o de tipo nomológico (leyes que relacionan variables), generalizar por inducción observaciones concretas, contrastar hipótesis, etc., procedimientos todos ellos que forman parte del método científico descrito en la parte I.

La metodología cuantitativa tiende a una mayor objetividad, reduciendo el margen para las opiniones o las interpretaciones. Eso no quiere decir que conduzca rigurosamente a la verdad, porque algunos de sus procedimientos son de tipo inductivo, que genera conocimiento general a partir de un conocimiento particular observado. Como vimos en la parte I, la inducción no produce certezas científicas, pero las técnicas de la estadística inferencial sí permiten asignar probabilidades al conocimiento general extraído de un conjunto limitado de observaciones.



Licencia: dominio público. Fuente: Pixabay.

## 8.2. Metodología cualitativa

La metodología cualitativa se refiere generalmente a la descripción, la comprensión y la interpretación de las ideas y conductas humanas, tanto a nivel individual como a nivel colectivo (social). La comprensión y la interpretación solo pueden ser llevadas a cabo, claro está, por la mente humana, que es capaz de ponerse en el lugar de otro. La recogida de información es mucho más flexible que en el caso cuantitativo, porque no está restringida a variables cuantificables (asociadas a valores numéricos) y porque puede irse adaptando al propio desarrollo de la investigación, en el que pueden intervenir los sujetos en estudio además del propio investigador. Ejemplos de recogida de información en esta metodología es la entrevista en profundidad a un sujeto o la recopilación de textos de los escritores de un determinado grupo o generación literaria.

Esta metodología busca los motivos, las intenciones, las significaciones de los sujetos o sociedades en estudio, y lo hace observando sus acciones o analizando sus discursos. En este último caso, se atiende no solo a lo que se dice sino también a cómo se dice, en el marco de las siguientes disciplinas:

- Sociolingüística: estudia los aspectos sociales que influyen en el uso de la lengua;
- Semántica: analiza las relaciones entre los conceptos y los signos lingüísticos asociados a ellos;
- Pragmática: contempla las condiciones de uso de los signos lingüísticos, es decir, el contexto en el que se producen.

Las investigaciones cualitativas pueden fundamentarse en la comparación (por ejemplo, entre los sistemas legales de diferentes países, o entre los ritos de diferentes tribus), en la perspectiva histórica (por ejemplo, el origen y la evolución del derecho español), en la consideración de la finalidad y utilidad de las acciones o en el valor y significado que tienen para los propios sujetos que las realizan. Cada uno de estos puntos de vista se puede identificar con variantes del método cualitativo: el método comparativo, el método histórico, el método crítico-racional o el método etnográfico.



Licencia: dominio público. Fuente: Pixabay.

### ▶ AMPLIACIÓN. Historia, comparación, crítica y etnografía

Aunque se podría definir un método cualitativo específico para cada ciencia, conviene clasificarlos en un conjunto reducido como el siguiente:

El método histórico comprende la recopilación y análisis de fuentes primarias documentales y de elementos materiales conservados en el tiempo (de tipo arqueológico) para estudiar los hechos ocurridos en el pasado que tuvieron mayor influencia en el desarrollo de los acontecimientos posteriores. Este método parte, como cualquier otro, de la formulación de un problema y una

hipótesis de trabajo, a partir de lo cual se inicia la búsqueda de las fuentes o elementos materiales adecuados para contrastarla. A continuación, se procede al análisis y crítica de todas las evidencias recopiladas. La crítica de documentos históricos conlleva valorar su autenticidad, autoría y procedencia (lo que se conoce como crítica externa), y también su fiabilidad histórica (crítica interna): imparcialidad, formación e intenciones del autor o autores, así como coherencia y factibilidad del contenido del documento. Por último, se lleva a cabo una síntesis de toda la información recopilada que permita sustentar las conclusiones del estudio, que serán finalmente difundidas.

El método comparativo consiste en buscar similitudes y analogías entre sistemas o procesos que tienen lugar en diferentes lugares o en diferentes épocas. La aplicación de este método a distintas áreas de conocimiento da lugar a la sociología comparada, etnografía comparada, lingüística comparada, derecho comparado, educación comparada, etc. (también se usa el adjetivo comparativo/a para estas disciplinas). Empleando este método pueden extraerse conclusiones como, por ejemplo, que los diversos sistemas comparados hayan tenido o tengan algún tipo de contacto o intercambio, o que provengan todos de la evolución de un único sistema anterior en el tiempo. Por ejemplo, si la aplicación del método comparativo evidencia un gran número de similitudes entre dos culturas muy separadas geográficamente, podría deducirse la existencia en el pasado de migraciones o relaciones culturales o comerciales entre ellas. Lo mismo podría concluirse a partir de la lingüística comparada, que en un ámbito más general permite establecer árboles genealógicos entre lenguas.

El método crítico-racional introduce el análisis de los fines y objetivos de los fenómenos que se investigan. En primer lugar, se trata de identificar esos objetivos, recurriendo a la interpretación de las acciones humanas. En segundo lugar, se critica racionalmente si la manera en que se desarrollan los fenómenos se ajusta o no a la consecución de los objetivos definidos. Por ejemplo, este método puede emplearse para estudiar (solo o en combinación con otros) el papel actual de las tecnologías de la información y la comunicación: cuáles son los fines, supuestamente beneficiosos para las personas y la sociedad en su conjunto, de su cada vez más extendido uso, cuál es el grado en el que se alcanzan, y qué otras consecuencias inesperadas y quizá contraproducentes aparecen.



Licencia: dominio público. Fuente: Pixabay.

El método etnográfico se caracteriza por la exploración in situ, mediante trabajo de campo, de los fenómenos sociales y culturales de un grupo humano, en particular a través de una observación participante, en la que el propio investigador interactúa con los miembros de la sociedad en estudio. Además de la observación, se realizan entrevistas en profundidad a los individuos para elaborar detalladas biografías, de modo que se recoge información tanto de lo que hacen como de lo que dicen y piensan. Todo ello permite comprender e interpretar con mayor profundidad las acciones humanas en el contexto social.

Otra de las virtudes de este método es que permite contrastar dos puntos de vista generalmente diferentes de los fenómenos socioculturales: el emic y el etic. La perspectiva emic ofrece una explicación desde el punto de vista de las personas que participan en él, y se sustenta por tanto en su sistema de creencias compartidas y en el marco de los conceptos que tienen significado para ellos. La perspectiva etic, por su parte, explica desde el punto de vista del investigador, que aplica un marco científico racional. En la parte I vimos un ejemplo de explicaciones emic y etic de un mismo fenómeno, el de la veneración hacia las vacas en la India: mientras la perspectiva emic explica ese comportamiento en el marco de un complejo sistema de creencias, la perspectiva etic lo hace en base a las posibles ventajas económicas de mantener una abundante cabaña bovina en lugar de consumir su carne (esta sería, en particular, la explicación del antropólogo Marvin Harris desde el punto de vista del materialismo cultural).

### 8.3. Metodologías y tipos de ciencias

La metodología cuantitativa resulta en general más adecuada para las ciencias naturales (física, química, biología, etc.) y la cualitativa, para las ciencias sociales. Pero la metodología cuantitativa también puede aplicarse en investigaciones en ciencias sociales. En estos casos, es necesario expresar los conceptos propios de estas ciencias en términos operacionales, es decir, asociándolos a variables con procedimientos de medida bien definidos que toman valores numéricos.

Como ejemplos de datos numéricos en una investigación social cuantitativa se pueden citar los resultados de exámenes, pruebas o tests diseñados para medir ciertas capacidades de los individuos. Esos datos pueden tratarse a continuación estadísticamente para contrastar la hipótesis de partida en el estudio.

Las metodologías cuantitativa y cualitativa no son mutuamente excluyentes, sino que pueden ser complementarias en una misma investigación, especialmente en el ámbito ya señalado de las ciencias sociales. En ocasiones será imprescindible interpretar los resultados estadísticos obtenidos de los datos numéricos de acuerdo con las motivaciones, intenciones o significaciones de los sujetos que los han generado.

Dada la importancia de la metodología cuantitativa tanto para las ciencias naturales como para las sociales, la parte III se dedicará a profundizar en sus procedimientos y particularidades, incluyendo la estadística descriptiva, la estadística inferencial, el cálculo de incertidumbres en las medidas y la presentación de resultados numéricos.

### 8.4. Triangulación

La combinación de distintas metodologías en una investigación, como la cuantitativa y la cualitativa, se denomina triangulación metodológica. Existen otros tipos de triangulación, como la que consiste en combinar distintas fuentes de datos, investigadores de áreas o formación académica diferente, marcos teóricos diversos, etc.

En general, la triangulación en una investigación es deseable porque amplía el ámbito de estudio y a la vez profundiza en él, además de dotar de mayor validez a las conclusiones obtenidas. Si estas no cambian cuando se emplean distintas fuentes de datos, marcos teóricos, métodos de análisis, investigadores, etc., significa que son muy robustas. Las principales desventajas de la triangulación son que requiere más tiempo y recursos para la investigación y una mayor formación de los investigadores en diferentes metodologías. Además, puede ocurrir que los resultados obtenidos con los diferentes elementos combinados en una triangulación no sean fácilmente comparables entre sí, lo que complicaría ofrecer unas conclusiones unificadas y coherentes.

Esquema del capítulo:



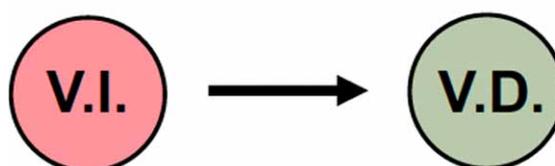
Fuente: elaboración propia con Genially.

<sup>1</sup> Enlace relacionado: <http://view.genial.ly/5c607dd64e10564933e0ff3a/metodologias>

# Capítulo 9.

## VARIABLES DE UNA INVESTIGACIÓN

Las variables que aparecen en la formulación de la hipótesis de partida en una investigación tienen dos funciones principales. Por un lado, la variable independiente es la que ejerce su influencia o se puede considerar la causa de otra variable. Puede ser manipulada por el experimentador o directamente medida, dependiendo del diseño de la investigación. Por otro lado, la variable dependiente es la que resulta influida o se puede considerar efecto o consecuencia de otra variable, y su valor directamente se mide.

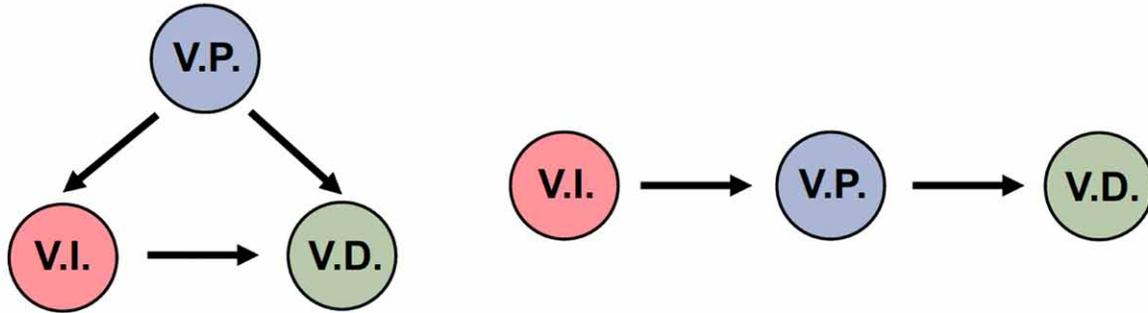


Fuente: elaboración propia.

Además, pueden existir otras variables que afecten, interfieran o confundan la relación entre la independiente y la dependiente, que se denominan variables perturbadoras o variables extrañas. Dentro de ellas, una variable antecedente es aquella que influye a la vez en la variable independiente y en la dependiente; puede ser el origen de una correlación espuria entre ambas, es decir, que no se deba a que la independiente influya necesariamente en la dependiente. Otro tipo es la variable interviniente, que depende de la independiente y a su vez influye en la dependiente, es decir, interviene o media en el efecto (ahora indirecto) de la independiente sobre la dependiente. También puede aparecer una variable contextual, cuyos valores modifican la naturaleza o modulan la intensidad de la influencia de la independiente sobre la dependiente; a menudo están relacionadas con el contexto espacio-temporal al que se circunscribe la investigación.

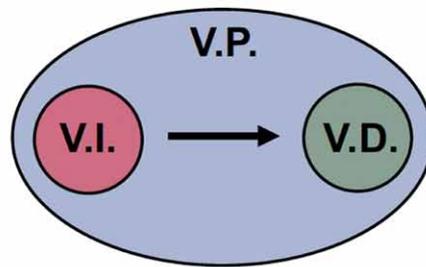
La actuación de una variable perturbadora podría provocar que la conclusión obtenida en el contraste de la hipótesis de la investigación no fuese válida, es decir, podríamos afirmar una relación entre la variable independiente y la dependiente que no fuese real. Para evitarlo, puede ser necesario aplicar alguna técnica de control de las variables perturbadoras, tras lo cual pasan a ser variables

controladas. La identificación y medida de variables perturbadoras no siempre resulta sencilla, pero aun así pueden ser controladas con algunas técnicas, como la aleatorización. Todas ellas se explican en un capítulo posterior.



Variable perturbadora antecedente

Variable perturbadora interviniente

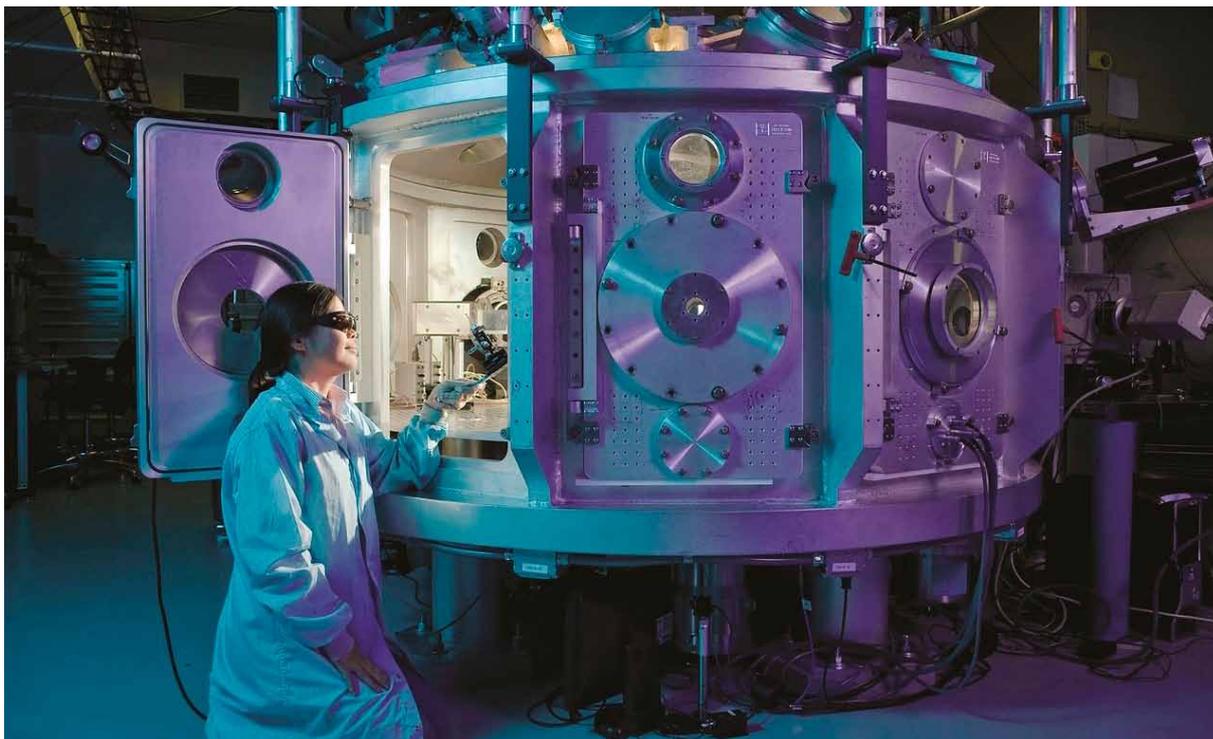


Variable perturbadora contextual

Fuente: elaboración propia.

### EJEMPLO

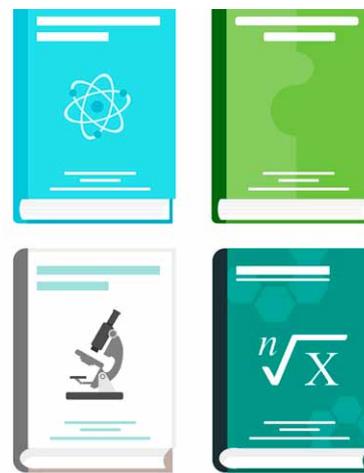
En una investigación se quiere establecer si la incidencia de luz (variable independiente) influye en la conductividad eléctrica de un cierto material (variable dependiente). Pero el investigador sospecha que la temperatura también puede influir en la conductividad, actuando como variable perturbadora, por lo que decide controlarla. Para ello, puede tomar todas sus medidas siempre a la misma temperatura, o bien puede medir y anotar la temperatura cada vez que mide las otras variables, y analizar los datos por separado para distintas temperaturas. En ambos casos, la variable temperatura se ha convertido en variable controlada. Para entender cómo la variable perturbadora podría interferir en el contraste de la hipótesis de la investigación, se puede suponer que el investigador realiza medidas cada diez minutos aumentando en cada una de ellas la incidencia de luz sobre el material, empezando a las ocho de la mañana y terminando a las tres de la tarde. Si los datos recogidos muestran que la conductividad aumenta en cada medida, puede deberse al incremento en la incidencia de luz (lo que confirmaría la hipótesis de la investigación), pero también puede deberse a que la temperatura también iba aumentando conforme avanzaba el día, o a ambas causas a la vez. En este ejemplo, la teoría física nos hace sospechar muy seriamente sobre la influencia de la variable temperatura, por lo que es razonable controlarla desde el principio. Pero si se sospecha de otra variable menos evidente, quizá lo más conveniente no sea controlarla, sino introducirla como una nueva variable independiente para analizar en detalle su influencia, ampliando así la investigación.



Licencia: dominio público. Fuente: Pixabay.

## EJEMPLO

Se quiere analizar el efecto de un nuevo método didáctico sobre los resultados académicos de los alumnos a los que se aplica. La variable independiente es el método didáctico, que tiene solo dos categorías: nuevo o tradicional. La variable dependiente es el rendimiento académico, que será medido con un examen. Una posible variable perturbadora es el nivel educativo de los alumnos. Imaginemos que el nuevo método se aplica a estudiantes de segundo de Bachillerato y el método tradicional a estudiantes de primero de la ESO. La diferencia en los resultados se debería muy probablemente a la disparidad de conocimientos previos y madurez intelectual entre los dos grupos, y el efecto del método didáctico aplicado, si es que existe, se vería difuminado.



Licencia: dominio público. Fuente: Pixabay.

## 9.1. Escalas de medida

Los valores que toman las variables pueden ser cualitativos (categorías), como sucede con la variable color, o cuantitativos (números), como ocurre con la variable edad. Los valores de una variable deben ser mutuamente excluyentes, para que no haya duda de cuál asociar con cada medida. Conviene además que el rango de valores sea lo suficientemente amplio para que ninguna medida quede sin clasificar, y que esos rangos contengan un gran número de valores distintos para distinguir con la mayor precisión posible unas medidas de otras. Así, por ejemplo, será preferible en general que la variable color pueda ser asociada a veinte categorías distintas en lugar de a siete, o que las edades se midan en años y meses en lugar de en años enteros, pero siempre teniendo en cuenta el aumento de complejidad que se introduce en la investigación.

Las variables cualitativas son discretas, es decir, no pueden tomar valores intermedios entre las categorías establecidas. También pueden ser discretas algunas variables cuantitativas, como la población o el número de votos recibidos por un partido político, porque no existen fracciones de personas o de votos (sus posibles valores numéricos van de uno en uno). En contraste, las variables continuas sí admiten valores intermedios entre dos valores dados, de hecho una infinidad de ellos. En la práctica, el número de valores posibles depende de la precisión de las herramientas empleadas en la recogida de datos. Un ejemplo de variable continua es la temperatura: entre uno y dos grados existen en principio infinitos valores posibles, como 1,14569... grados, pero el termómetro usado no será capaz de determinar más allá de un cierto decimal.

Según el tipo de valores que toman las variables, se clasifican en los siguientes tipos:

– Variables nominales. Son variables cualitativas cuyos valores son categorías entre las que no puede establecerse un orden. Simplemente se pueden distinguir unas categorías de otras, es decir, establecer relaciones de igualdad o desigualdad entre ellas. La variable «color» es de este tipo, porque sus valores son categorías y no existe un criterio para ordenarlas (excepto en una investigación en física, donde cada color se asocia a un rango de longitudes de onda, que son valores numéricos). Otros ejemplos en ciencias naturales son el carácter ácido o base, u oxidante o reductor de una sustancia química, el tipo de magnetismo que presenta, o la especie a la que pertenece un ser vivo. Ejemplos en ciencias sociales son el sexo, la nacionalidad o el partido político al que se vota. Este tipo de variables están muy ligadas a la clasificación, procedimiento habitual en ciencia que vimos en la parte I.



Licencia: dominio público. Fuente: Pixabay.

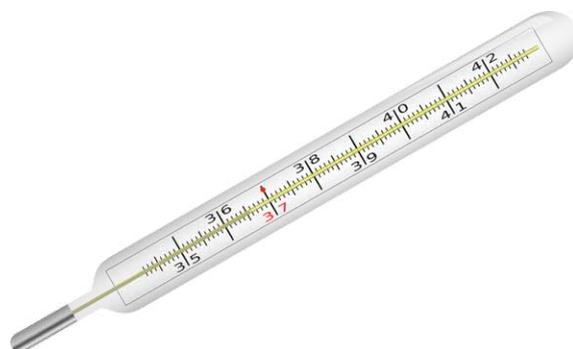
– Variables ordinales. Como las anteriores, son variables cualitativas cuyos valores son categorías, pero estas sí pueden ordenarse siguiendo algún criterio. Ese orden puede resultar claro, como en la variable nivel educativo: estudios primarios / secundarios / universitarios. O puede ser algo más subjetivo, como en la variable partido político al que se tiene intención de votar, que antes habíamos considerado como variable nominal, pero en cuyas cate-



Licencia: dominio público. Fuente: Pixabay.

gorías sí se podría establecer un orden desde la extrema derecha hasta la extrema izquierda pasando por el centro; esto puede resultar interesante para algunos tipos de investigación, por ejemplo, para estudiar la relación entre el nivel educativo o económico de una persona y su posición ideológica.

- Variables de intervalo. Son variables cuantitativas, cuyos valores numéricos permiten definir intervalos o diferencias entre ellos. Por ejemplo, en la variable longitud existe el mismo intervalo entre los valores «2 metros» y «3 metros» que entre los valores «3 metros» y «4 metros»: en ambos casos la diferencia en los valores numéricos es de 1 metro, que representa la misma distancia física. Otro ejemplo es el de la variable temperatura medida en la escala centígrada o Celsius (°C): existe la misma diferencia de temperatura física entre los valores 0 °C y 100 °C que entre los valores 100 °C y 200 °C. Las variables de intervalo permiten realizar operaciones de suma y resta sobre sus valores.



Licencia: dominio público. Fuente: Pixabay.

- Variables de razón. Son variables cuantitativas, cuyos valores numéricos permiten definir, además de intervalos como en las anteriores, relaciones de proporción o razón. La variable longitud del ejemplo anterior es en realidad una variable de razón, porque tiene sentido decir, por ejemplo, que la proporción, razón o cociente entre los valores «2 metros» y «1 metro» es igual a la proporción existente entre los valores «4 metros» y «2 metros»: en ambos casos hay un factor de proporción 2 entre los valores, y físicamente es cierto que una de las longitudes es el doble que la otra. Las variables de razón permiten realizar operaciones de multiplicación y división sobre sus valores, además de las de suma y resta. La mayoría de variables cuantitativas son de razón, es decir, admiten el cálculo tanto de intervalos como de proporciones, aunque existen excepciones.

En las investigaciones donde es posible elegir el tipo de variable será necesario plantearse el grado de precisión que se quiere alcanzar en el estudio, siendo mayor con las variables de intervalo o razón (cuantitativas) que con las nominales u ordinales (cualitativas). Esta elección afecta al procedimiento de medida de las variables y al tipo de análisis estadístico (paramétrico o no paramétrico) que se aplica a los datos recogidos, y dependerá de la complejidad que se quiera o se pueda alcanzar en la investigación y de la calidad de los resultados que se desee obtener.



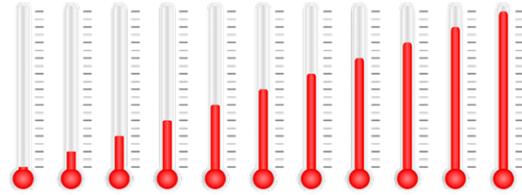
Licencia: dominio público. Fuente: Pixabay.

En la siguiente tabla se resumen las relaciones que se pueden establecer entre los valores de los diferentes tipos de variable según su escala, así como el tipo de análisis que se puede llevar a cabo con ellas:

Tipo de variable	Relaciones que permite establecer	Tipo de análisis
Nominal	Igualdad	Cualitativo, estadística no paramétrica
Ordinal	Igualdad y orden	
De intervalo	Igualdad, orden e intervalo	Cuantitativo, estadística paramétrica
De razón	Igualdad, orden, intervalo y proporción	

**AMPLIACIÓN. Escalas de intervalo y de razón**

La diferencia entre las escalas de razón y las que son exclusivamente de intervalo reside en que en las primeras se puede definir un origen de la escala, es decir, un punto cero con verdadero significado. Por ejemplo, en la escala de longitudes, cero metros significa tamaño nulo, o en la escala de edades, cero años significa recién nacido.



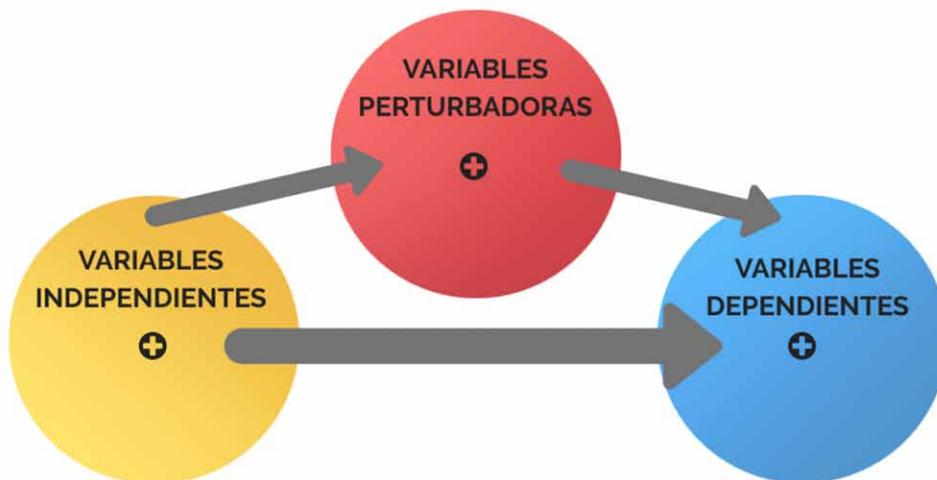
Licencia: dominio público. Fuente: Pixabay.

Esto no ocurre, por ejemplo, con la temperatura en escala centígrada, donde el valor cero (cero grados centígrados, 0 °C) no significa ausencia de temperatura (ausencia de agitación térmica); pueden darse de hecho temperaturas más bajas, por ejemplo -10 °C (diez grados bajo cero). Por esta razón no tiene sentido decir que 20 °C es el doble de temperatura que 10 °C, aunque los números así lo indiquen, porque no es cierto físicamente: la primera temperatura no implica el doble de agitación térmica que la segunda. Así, la temperatura en escala centígrada es una variable exclusivamente de intervalo, no de razón.

Sin embargo, la temperatura medida en escala absoluta o Kelvin sí es una variable de razón, porque el cero de esa escala (cero grados Kelvin, 0 K) sí tiene el significado de ausencia de agitación térmica. No existen grados Kelvin negativos porque no se puede estar más frío que absolutamente frío. En esta escala, sí se puede decir que 300 K implica el doble de agitación térmica que 150 K. En escala centígrada, esas temperaturas corresponden a valores de 26,85 °C y -123,15 °C respectivamente, que no son uno el doble del otro. En ambas escalas el intervalo sí es el mismo, 150 grados.

Esquema del capítulo<sup>2</sup>

**FUNCIONES DE LAS VARIABLES EN UNA INVESTIGACIÓN**



Fuente: elaboración propia con Genially.

<sup>2</sup> Enlaces relacionados:  
<http://view.genial.ly/5c71ade1a159c7762902b00d/interactive-content-funciones-de-las-variables>  
<http://view.genial.ly/5c7195eb08403f02612637f4/tipos-de-variables>



Fuente: elaboración propia con Genially.

# Capítulo 10.

## Diseños de investigación

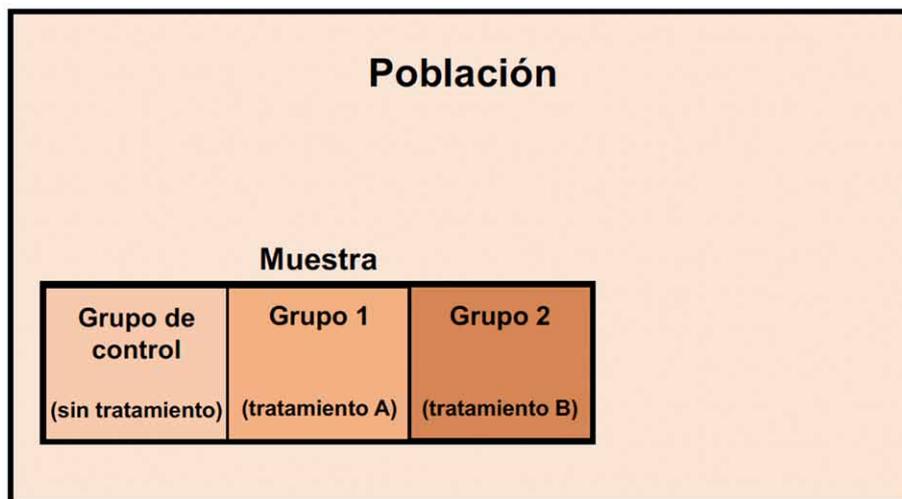
El diseño de una investigación recoge las estrategias, técnicas y métodos necesarios para desarrollarla y alcanzar los objetivos propuestos. Existen varios tipos de diseños, que se clasifican respecto a diversas características.

En general, una investigación pretende estudiar el modo en que los valores que toman las variables independientes influyen en los valores de las variables dependientes. Cuando se contempla una única variable independiente, se trata de un diseño simple, mientras que si se introducen dos o más, se trata de un diseño factorial, que resulta mucho más complejo. Por otro lado, cuando se contempla una única variable dependiente se trata de un diseño univariado, mientras que si se incluyen dos o más, es un diseño multivariado, que también es más complicado. En un proyecto de introducción a la investigación científica en la Enseñanza Secundaria Obligatoria y Bachillerato conviene limitarse, en general, a diseños simples y univariados.

Los diseños también se pueden clasificar según su grado de experimentalidad. Un diseño experimental se caracteriza por controlar y manipular las condiciones en las que se recogen los datos, que es lo que distingue la experimentación de la mera observación, como vimos en la parte I. Se puede definir un estudio experimental como aquel en el que el investigador manipula algunas variables (las independientes), para establecer con mayor precisión y fiabilidad su efecto sobre otras variables (las dependientes). Un estudio plenamente experimental mantiene además bajo control las variables perturbadoras, que pueden influir sobre la dependiente, pero que no son objeto de análisis detallado en el estudio en cuestión. Si este control no se efectúa de forma completa, el diseño se denomina cuasi-experimental.

La manipulación y el control de variables se consiguen típicamente en el laboratorio, donde es habitual realizar las medidas para investigaciones en física, química, biología, etc. En muchos de estos casos la variable independiente se mide en escala de intervalo o de razón y se puede manipular de manera continua, dando lugar en general a cambios también continuos en la variable dependiente.

En otros casos, los valores de la variable independiente se agrupan en intervalos o vienen dados ya directamente en categorías. La muestra de la investigación se divide entonces en grupos, cada uno de ellos asociado a una categoría distinta de la variable independiente que se suele denominar tratamiento. Es habitual que alguna de esas categorías corresponda a lo que se considera un valor nulo de la variable independiente, y el grupo al que se le asigna se denomina entonces grupo de control.



Fuente: elaboración propia.

Por otro lado, los valores de las variables perturbadoras que pueden influir en la variable dependiente han de distribuirse por igual en todos los grupos de la muestra, de manera que estos resulten equivalentes entre sí. Esto se consigue aplicando técnicas de control de variables, que veremos posteriormente. La más importante de todas ellas es la aleatorización, que en estos estudios plenamente experimentales debería aplicarse en varios niveles: selección aleatoria de los elementos de la muestra de entre los de la población, asignación aleatoria de esos elementos a cada grupo formado en la muestra, y asignación aleatoria de los valores de la variable independiente (tratamientos) a cada grupo.

Este diseño experimental basado en la división de la muestra en grupos según el valor de la variable independiente es habitual en investigaciones sociales, psicológicas o biomédicas (epidemiológicas). En ellas, los elementos de la muestra son seres humanos, que pueden reaccionar consciente o inconscientemente a la manipulación de la variable independiente, es decir, a la administración de los tratamientos. Para evitarlo, se puede ocultar a los sujetos en estudio la información sobre qué tratamiento están recibiendo; esto implica que el grupo de control, al que en teoría no se le debería administrar nada, ha de recibir un placebo de características similares al resto de tratamientos. La información del tratamiento asignado a cada grupo puede permanecer oculta incluso para los propios investigadores para evitar que ellos también influyan en los resultados, lo que se denomina técnica de doble ciego.

### EJEMPLO

En una investigación se quiere estudiar la eficacia de dos nuevos fármacos para reducir el nivel de colesterol. Se procede a seleccionar un conjunto de sujetos con el nivel de colesterol alto, que constituyen la muestra del estudio. La variable independiente será el fármaco con el que ha sido tratado cada individuo de la muestra (de aquí procede el término «tratamiento» para referirse en general a los valores de la variable independiente manipulada).

En esta investigación la variable independiente puede tomar tres valores categóricos: fármaco A, fármaco B y ausencia de tratamiento. Se divide la muestra en tres grupos, a cada uno de los cuales se le asigna una de esas tres categorías: a un grupo se le suministra el fármaco A, a otro grupo se le suministra el fármaco B, y a otro grupo no se le suministra ningún fármaco. Este último caso se asocia con el valor nulo de la variable independiente, y constituye el grupo de control; para que sus miembros no sepan que no se les está suministrando un fármaco real, reciben un placebo de aspecto similar a los fármacos A o B, pero sin efecto sobre el colesterol. La variable dependiente de esta investigación es el nivel de colesterol, que se mide después de aplicar los tratamientos a los grupos.

En cuanto a las variables perturbadoras de esta investigación, se nos pueden ocurrir muchas, como la edad, el grado de sedentarismo o el grado de obesidad. Todas ellas pueden influir directamente en el nivel de colesterol de los individuos, o bien influir en el modo en que los fármacos actúan sobre el nivel de colesterol. Es decir, pueden interferir en el efecto de la variable independiente sobre la dependiente. Por ejemplo, si la media de edad de los individuos de un grupo es bastante mayor que en los otros grupos, es posible que sus niveles de colesterol tras aplicar los tratamientos sean mayores, pero no porque el fármaco suministrado a ese grupo sea en general menos efectivo que el otro, sino por la influencia de la edad en el efecto. Se trata tan solo de una suposición, pero lo más razonable es controlar esa variable para neutralizar su posible influencia. Una manera de hacerlo es eligiendo al azar los individuos de la muestra y su reparto en grupos, de modo que sus edades queden distribuidas de manera equilibrada, o bien eligiendo (si fuera posible) todos los individuos dentro de un mismo rango de edades. En esta investigación también resultaría conveniente mantener oculto el tratamiento que recibe cada grupo tanto a los pacientes como a los investigadores, para evitar que unos u otros desvíen de manera consciente o inconsciente las medidas o el análisis de los datos (técnica de doble ciego).

Cuando la variable independiente de la investigación no se manipula, se trata de un diseño no experimental (o preexperimental), que no debe confundirse con no científico, pues estos diseños lo son plenamente. En ellos, los valores de las variables han sido adquiridos por los elementos de la muestra en su contexto natural, sin manipulación por parte del investigador. El hecho de no manipular la variable independiente puede deberse a varios motivos: porque es difícil y no se cuenta con los recursos necesarios, porque podría influir en los resultados alejándolos de lo que ocurriría sin intervención humana, porque no es ético (en humanos), o directamente porque no es posible.

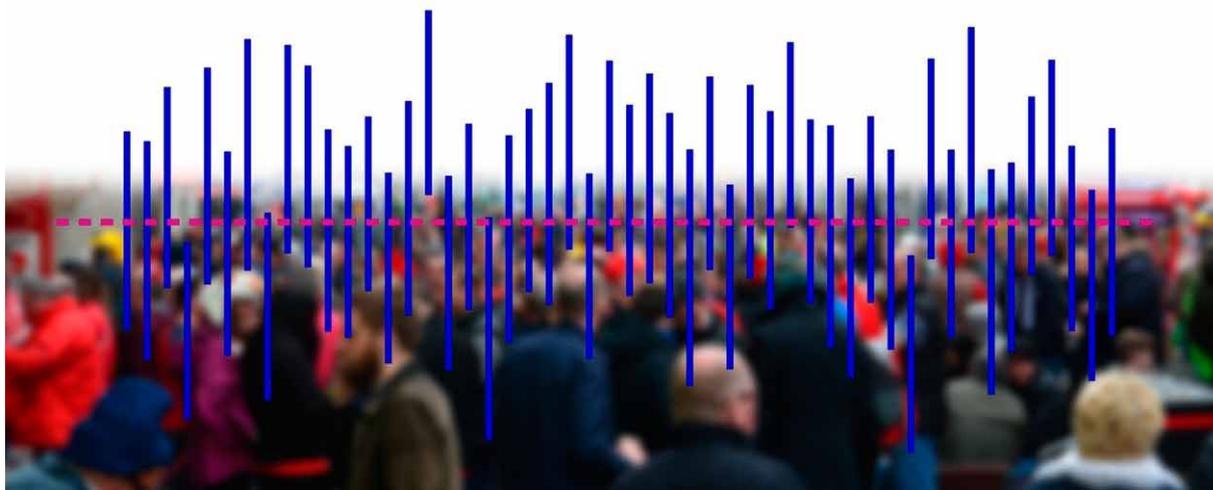
Un ejemplo donde no hay más remedio que aplicar un diseño no experimental es el de investigaciones históricas, donde los hechos ya han ocurrido y no existe la opción de manipular variables independientes, de controlar variables perturbadoras *a priori* ni de tomar nuevas medidas de las variables dependientes. Otro ejemplo es un estudio de campo en geología, en el que se realizan observaciones sobre el terreno de los efectos de la erosión o la sedimentación, cuyas causas a grandes escalas espaciales y temporales no son, evidentemente, reproducibles ni manipulables. Se podría complementar este estudio, aunque no remplazar, con una investigación experimental en laboratorio, por ejemplo sometiendo las rocas y minerales recogidos en el terreno a diversos procesos químicos y físicos relacionados con los procesos geológicos, y midiendo los efectos.

#### ▶ AMPLIACIÓN. Diseños transversales y longitudinales

En una investigación con diseño sincrónico o transversal se estudian todos los grupos de la muestra a la vez durante un mismo periodo de tiempo relativamente corto. Si, por el contrario, se estudia un único grupo durante un periodo de tiempo relativamente largo, se trata de un diseño diacrónico o longitudinal.

Dentro de este último tipo se encuentran los diseños de cohorte y los diseños de panel. Una cohorte es un conjunto de elementos que comparten alguna característica, que a menudo se refiere a que hayan adquirido alguna cualidad en un momento pasado. Por ejemplo, forman una cohorte los individuos que han nacido el mismo año, los que se han graduado en ciertos estudios (ya sea en un año específico o en cualquier momento del pasado), o los que han estado expuestos a un determinado factor de riesgo para su salud (durante un mismo periodo concreto o en cualquier momento del pasado). La composición de algunas cohortes puede cambiar con el tiempo: unos elementos entran y otros salen, según adquieran o pierdan la característica que define la pertenencia a ella.

En un diseño longitudinal de cohorte, cada cierto tiempo se extrae una muestra distinta de entre los elementos de la cohorte y se mide en ella la variable dependiente. A veces la cohorte se identifica con la población total (es decir, la característica compartida por todos los elementos es la pertenencia a esa población), y el diseño se llama entonces de tendencia. En un diseño longitudinal de panel también se mide la variable dependiente cada cierto tiempo, pero siempre a los mismos elementos de la cohorte. Así, la diferencia entre ambos diseños es que en el de cohorte se extraen muestras distintas cada vez que se va a realizar una medida, mientras que en el de panel siempre es la misma muestra. En ambos casos se estudia la evolución en el tiempo de la cohorte, pero el de panel permite además estudiar la evolución de cada uno de los elementos individualmente y analizar cuáles de ellos son los responsables principales de los cambios en la cohorte en general. Sin embargo, el diseño de panel es más difícil de llevar a cabo que el de cohorte porque exige realizar medidas a lo largo del tiempo a un mismo conjunto de elementos, que podría ir reduciéndose progresivamente y dejar de ser representativo.



Licencia: dominio público. Fuente: Pixabay.

#### ▶ AMPLIACIÓN. Algunos tipos de diseños no experimentales

Existen diversos tipos de investigaciones no experimentales. El diseño *ex post facto* se aplica a las investigaciones de hechos que ya han sucedido. Se miden los valores de la variable independiente, que no se puede manipular porque sus valores ya fueron fijados en el pasado, y de la variable dependiente, así como de las posibles variables perturbadoras. Cuando los valores de la variable dependiente también han sido fijados con anterioridad, el estudio es de tipo retrospectivo, y no permite establecer una relación de causalidad: no se puede saber si los valores de la variable independiente quedaron fijados antes que los de la dependiente y pudieron por tanto influir en ellos; por esta razón, este tipo de diseño se emplea para investigar correlaciones entre variables en lugar de relaciones causa-efecto.

Si la variable dependiente aún no se ha fijado o no ha ocurrido, el estudio es de tipo prospectivo, y en él sí está asegurado el orden temporal entre variable independiente y dependiente que podría hacer sospechar de una relación casual. En un diseño *ex post facto*, el análisis que se lleva a cabo con las medidas realizadas, aplicando en particular métodos estadísticos, es similar al de los diseños experimentales.

El término *ex post facto* es típico de investigaciones psicológicas o sociales, pero un diseño similar puede darse también en las investigaciones de carácter histórico ya mencionadas, o en las que se basan únicamente en fuentes documentales (no necesariamente antiguas, sino también actuales, como bases de datos estadísticas) o en estudios de ciencias naturales sobre hechos pasados o no manipulables: geología, biología (evolución, ecología), astronomía, etc.

Otro tipo de diseño no experimental es el estudio de caso, que es una investigación profunda, detallada y sistemática sobre un objeto, hecho, proceso o individuo. A menudo es de tipo cualitativo, aunque nada impide que se recojan también datos cuantitativos. Los estudios de caso tienen un importante componente descriptivo, pero también pueden orientarse a la explicación, explorando el cómo y el porqué del objeto de análisis. En ciencias sociales los estudios de caso se refieren a hechos o actividades relacionadas con personas, ya sea individualmente (estudios de caso biográficos) o en grupo (estudios de caso etnográficos).

## EJEMPLO

En una investigación *ex post facto* retrospectiva sobre la relación entre los hábitos alimentarios y el riesgo de infarto, se forman dos grupos de personas según hayan sufrido o no un infarto (variable dependiente), y se estudia en cada grupo los hábitos alimentarios que han seguido los sujetos en el pasado (variable independiente). El hecho de organizar los grupos según el valor de la variable dependiente es especialmente útil cuando la incidencia de alguno de sus valores es baja; en este ejemplo, permite asegurarnos de contar con un grupo suficientemente grande de personas que han sufrido un infarto (si se formasen los grupos con otro criterio, quizá el número de infartos en esos grupos sería pequeño e impediría obtener conclusiones sólidas).

Se pueden medir también variables que puedan resultar perturbadoras, e intentar equilibrarlas en los grupos formados para que no afecten a la posible correlación entre la variable dependiente y la independiente. Un ejemplo de variable perturbadora en este estudio podría ser el nivel de estrés que han sufrido los sujetos en el pasado, que puede influir en el riesgo de infarto, en los hábitos alimentarios por efecto de la ansiedad, o en ambas variables a la vez. En un estudio análogo, pero prospectivo, se podrían formar grupos de individuos según la variable independiente, es decir, según los hábitos alimentarios llevados en el pasado, y hacer un seguimiento durante un largo tiempo posterior para estudiar la incidencia de infarto en cada grupo; en este caso la variable dependiente aún no ha ocurrido al inicio de la investigación y, como se señaló antes, la incidencia de uno de sus valores (el hecho de sufrir un infarto) podría ser muy baja y afectar a la validez de los resultados.

Un estudio similar a los descritos, pero de carácter experimental, exigiría que los grupos se formasen al azar y que a cada uno de ellos se le asignase también al azar un tratamiento en forma de diferente alimentación durante un largo periodo de tiempo. En esos grupos el nivel de estrés, o cualquier otra variable potencialmente perturbadora, estaría repartido al azar, sin predominar significativamente en un grupo o en otro. Durante un periodo de tiempo aún más largo se vigilaría la ocurrencia de la variable dependiente, es decir, los sujetos de cada grupo que sufren un infarto. En consecuencia, el resultado que se obtendría para la relación causal entre alimentación y riesgo de infarto sería mucho más claro. Un estudio de estas características presenta, como resulta evidente, dos graves problemas: por un lado, el de

mantener una dieta controlada en los sujetos durante los larguísimos plazos en los que se supone que esta variable puede influir en el riesgo de infarto; por otro lado, las implicaciones éticas que conlleva: ¿a qué grupo se le asigna la dieta potencialmente más perjudicial para la salud?



Licencia: dominio público. Fuente: Pixabay.

**AMPLIACIÓN. Diseños factoriales y cuadrado latino**

Como hemos visto, en algunos diseños de investigación la muestra se divide en grupos según las categorías de la variable independiente. Si se trata de un diseño factorial, que contempla varias variables independientes, el número de grupos necesarios crece muy deprisa. Supongamos, por ejemplo, que se pretende investigar la influencia del libro de texto (variable independiente) en los resultados académicos (variable dependiente), teniendo también en cuenta el efecto del tipo de profesor y del horario de las clases, que son otras dos variables independientes. Se estudian tres libros (impreso tradicional, impreso innovador y *e-book* multimedia), tres tipos de profesores (joven, mediana edad y muy experimentado) y tres horarios de clase (primera hora, media mañana y última hora). Existen 27 combinaciones distintas de valores (categorías) de las tres variables (por ejemplo: profesor joven, clase a primera hora, uso de *e-book* multimedia), lo que exigiría formar 27 grupos de alumnos para analizar cada combinación por separado. Probablemente este diseño resultaría inviable, sobre todo si se quiere que todos los grupos sean lo suficientemente numerosos como para que los resultados sean significativos.

Existe la posibilidad de formar menos grupos, pero es muy importante que en el conjunto aparezcan todas las variables y todos sus posibles valores en la misma proporción, evitando que haya algún valor de una variable que se combina más de una vez con un mismo valor de otra variable. En la práctica, las combinaciones adecuadas pueden construirse a través del denominado cuadrado latino. Consiste en una cuadrícula en la que cada fila corresponde a una categoría de una de las variables independientes, y cada columna corresponde a una categoría de otra de las variables independientes (se suponen iguales el número de categorías de ambas, para que resulte un cuadrado). A continuación, en cada una de las casillas se introduce una categoría de la tercera variable independiente, de tal modo que no se repitan en la misma fila ni en la misma columna.

En el ejemplo anterior, donde hay tres variables independientes cada una con tres categorías, el cuadrado latino, de tamaño 3×3, quedaría de la siguiente manera:

Profesor → Horario ↓	Joven	Mediana edad	Experimentado
primera hora	<b>Texto:</b> impreso tradicional	<b>Texto:</b> impreso innovador	<b>Texto:</b> <i>e-book</i> multimedia
media mañana	<b>Texto:</b> impreso innovador	<b>Texto:</b> <i>e-book</i> multimedia	<b>Texto:</b> impreso tradicional
última hora	<b>Texto:</b> <i>e-book</i> multimedia	<b>Texto:</b> impreso tradicional	<b>Texto:</b> impreso innovador

Como se puede observar, se han formado nueve combinaciones, que corresponderían a nueve grupos de sujetos. Hay combinaciones que faltan (hasta llegar a 27), pero las que se han formado cumplen las condiciones expuestas antes: aparecen todas las variables y todas las categorías en la misma proporción. Así, hay grupos de primera hora con los tres tipos posibles de profesores y también con los tres tipos de libros de texto diferentes, y lo mismo ocurre con las otras horas; o bien, hay grupos de profesor joven con los tres tipos posibles de horario y también con los tres tipos de libros de texto diferentes, y lo mismo ocurre con los otros profesores. La investigación podría consistir entonces en medir el rendimiento académico de los estudiantes en cada uno de los nueve grupos, calcular en cada uno de ellos la media aritmética, y compararlas entre sí para extraer el posible efecto de cada una de las variables independientes.

Es posible añadir al diseño otra variable independiente más, con tres categorías repartidas entre las casillas con los mismos criterios de antes, dando lugar al denominado cuadrado grecolatino. En las casillas hay que distribuir entonces las categorías de dos variables independientes distintas, sin repetirse en la misma fila o columna y sin repetirse la misma pareja en todo el cuadrado.

### HISTORIAS DE LA CIENCIA. Fisher y los cuadrados latinos en agricultura

El estadístico inglés Ronald Fisher (1890-1962) usó la técnica del cuadrado latino para diseñar una investigación sobre la eficacia de fertilizantes en el rendimiento agrícola. En su estudio, el cuadrado representaba una porción de terreno real dividido en parcelas. Conforme se avanzaba en cada fila y en cada columna de parcelas, la calidad de la tierra disminuía (en este caso no había dos variables independientes distintas, una para cada lado del cuadrado, sino solo una). En las parcelas del cuadrado se distribuían una serie de tratamientos distintos, en forma de diferentes cantidades de fertilizantes, tantos como parcelas había en cada fila o columna, cumpliendo las condiciones del cuadrado latino. Si además se quería estudiar el efecto de alguna otra variable, como por ejemplo la cantidad de riego, debía diseñarse un cuadrado grecolatino. Una vez asignados y efectuados los distintos tratamientos con fertilizantes y los distintos niveles de riego en cada parcela, al cabo de un tiempo se medía el rendimiento de la cosecha en cada una de ellas, y se comparaban entre sí bien a ojo o bien aplicando métodos de inferencia estadística desarrollados por el propio Fisher.

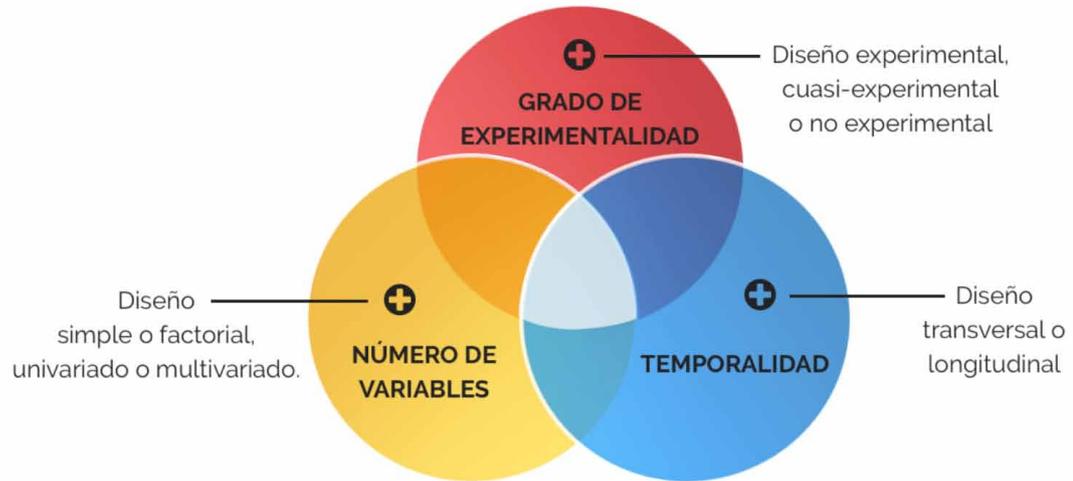
Fisher explicó los diseños de experimentos con cuadrados latinos y grecolatinos en su obra *The Design of Experiments* (1935), en la que se introducen muchos otros elementos básicos de la estadística inferencial, como el contraste de hipótesis, uno de los pilares de la ciencia moderna que trataremos más adelante.



Ronald Fisher. Licencia: dominio público.  
Fuente: Wikimedia Commons.

Esquema del capítulo<sup>3</sup>

## TIPOS DE **DISEÑOS** DE INVESTIGACIÓN



Fuente: elaboración propia con Genially.

---

<sup>3</sup> Enlace relacionado: <http://view.genial.ly/5c67d887a159c77629f5159d/tipos-de-disenos>

# Capítulo 11.

## Control de variables

Una variable perturbadora es aquella que interfiere en la relación que existe entre la variable independiente y la dependiente de una investigación. Su efecto puede estudiarse considerándola como una variable independiente adicional, pero eso incrementa notablemente la dificultad del estudio, convirtiéndolo en un diseño factorial. Si no interesa hacerlo así o no se cuenta con los recursos suficientes, será necesario al menos minimizar el efecto de la variable perturbadora a través de técnicas de control. Estas pueden consistir en procedimientos estadísticos de análisis de varianza, que se llevan a cabo tras medir todas las variables del estudio, incluyendo las perturbadoras. También existen técnicas de control de tipo experimental, que se ejecutan durante la fase de recogida de datos, y entre las que destacan las siguientes:

– Aleatorización:

Consiste en seleccionar aleatoriamente, es decir, al azar, los elementos que formarán parte de la muestra de estudio de entre el conjunto total, la población. Cuando la población no es infinita, se puede llevar a cabo la selección aleatoria asignando un número a cada elemento de la población y generando a continuación un conjunto de números aleatorios, que corresponderán a los elementos escogidos. Esto último es un procedimiento más complejo de lo que podría parecer, pero existen tablas y programas informáticos que permiten obtener números aleatorios con bastante buena aproximación (en realidad son pseudoaleatorios).

Por ejemplo, en una investigación sobre las propiedades físicas de un material, esta técnica podría consistir en realizar medidas en momentos o posiciones aleatorias. En una investigación sobre la eficacia de un método didáctico o de un tratamiento médico, la técnica de aleatorización implicaría que los sujetos que formen parte de la muestra se elegirían al azar de entre la población, se distribuirían al azar en los grupos formados según los valores de la variable independiente y se asignaría también

al azar el valor de esa variable (tratamiento) a cada grupo. A veces es difícil lograr la aleatorización completa. Por ejemplo, por razones logísticas quizá tengamos que realizar la investigación en el centro educativo u hospital en el que trabajamos, en lugar de elegirlo al azar, o tendremos que usar grupos de personas previamente establecidos, sin posibilidad de construir grupos nuevos al azar.

Se trata de una técnica de control porque los valores de las posibles variables perturbadoras se espera que queden representados de manera equilibrada, sin mayor proporción de unos valores sobre otros, ya sea en la muestra completa o en cada uno de los grupos formados dentro de ella.

– Igualación:

Consiste en seleccionar para la muestra únicamente los elementos en los que la posible variable perturbadora toma un determinado valor. De este modo, la variabilidad de la variable dependiente sería efecto de los distintos valores que toma la independiente, no la perturbadora.

En el ejemplo de la investigación sobre ciertas propiedades físicas de un material, esta técnica podría consistir en realizar todas las medidas a la misma temperatura, con la misma luminosidad, a la misma hora del día, etc. En el ejemplo de la investigación sobre métodos didácticos, habría que seleccionar estudiantes del mismo nivel educativo para que esa variable potencialmente perturbadora no influyese.

Para aplicar esta técnica es necesario tener claramente identificada la variable perturbadora. Además, hay que ser muy cuidadosos al tratar de generalizar las conclusiones obtenidas a la parte de la población que no tiene el valor seleccionado de la variable perturbadora. Así, si la investigación sobre el método didáctico se lleva a cabo en una muestra restringida por igualación a estudiantes de segundo de Bachillerato, habrá que valorar con cautela si las conclusiones obtenidas serán válidas en otros niveles educativos. A menos, claro está, que se lleve a cabo el estudio en distintos grupos de estudiantes de niveles diferentes, convirtiendo en la práctica esa variable perturbadora en una variable independiente adicional, y el estudio en uno de tipo factorial.

– Emparejamiento:

Como sabemos, es habitual hacer varios grupos en la muestra de estudio según los valores de una variable independiente de tipo categórico, para analizar por separado su efecto sobre la variable dependiente. Una variable perturbadora podría quedar desigualmente distribuida en esos grupos, con valores más altos en uno de ellos y más bajos en otro, lo que afectaría a los valores de la dependiente que se midan en cada grupo.

Supongamos una investigación en la que se necesitan dos grupos según los valores de la variable independiente; por ejemplo, el seguimiento de un método didáctico nuevo o de uno tradicional. Se miden los valores de la variable perturbadora en los elementos de la muestra (por ejemplo, el cociente intelectual de los estudiantes) y se ordenan de mayor a menor. En la técnica de emparejamiento, los elementos de esta lista ordenada se distribuyen en los dos grupos de la muestra de la siguiente manera: de los dos primeros elementos (los de valores más altos de la variable perturbadora) uno va al grupo 1 y otro al grupo 2, al azar; de los dos elementos siguientes, uno va al grupo 1 y otro al grupo 2, al azar; y así sucesivamente con todas las parejas de elementos contiguos en la lista ordenada. De este modo, se garantiza que ambos grupos contienen una distribución similar de valores de la variable perturbadora, que pasa a denominarse variable de emparejamiento y queda controlada. Si se forman más de dos grupos en la muestra, el procedimiento es análogo a este.

– Bloqueo:

Es una técnica muy similar a la de emparejamiento. Se mide la variable perturbadora en los elementos de la muestra y estos se distribuyen en bloques definidos por un rango de valores de la variable. Así, los elementos de un mismo bloque son muy similares entre sí en cuanto al valor de la variable perturbadora, porque se encuentran en el mismo rango de valores. A continuación, los elementos de cada bloque se distribuyen al azar entre los grupos creados según los valores de la variable independiente. Dado que el orden de los elementos es menos exhaustivo que en el emparejamiento (aquí se ordenan por bloques, no individualmente), el control alcanzado en la variable perturbadora, que pasa a denominarse variable de bloqueo, es menor.

– Medidas repetidas:

Esta técnica se puede usar en las investigaciones donde la variable dependiente se mide en varias ocasiones a lo largo del tiempo en los mismos elementos, para analizar su evolución. No requiere la medida de las variables perturbadoras, al contrario de lo que ocurre en las técnicas de igualación, emparejamiento o bloqueo. Como se trata de los mismos elementos, se supone que los valores de las variables perturbadoras permanecen constantes en ellos, por lo que no causan variaciones en las sucesivas medidas de la variable dependiente. Su efecto queda así controlado (se dice que los elementos ejercen su propio control), y las variaciones en la variable dependiente pueden entenderse debidas a los cambios en la independiente.

Esta técnica no sirve para controlar variables perturbadoras cuando se forman grupos en la muestra según los valores de la variable independiente, ya que no elimina posibles desequilibrios entre ellos. Una posible alternativa es aplicar los diversos tratamientos a toda la muestra, sin dividir en grupos, pero tiene sus inconvenientes. En el ejemplo de los métodos didácticos, se podría aplicar a toda la muestra el método didáctico tradicional en primer lugar, y medir los resultados académicos, y a continuación aplicar el método nuevo a toda la muestra y volver a medir los resultados, realizando así las medidas repetidas. La primera medida reflejaría el efecto en el rendimiento académico del método tradicional, pero la segunda medida recogería el efecto del método nuevo y también el del tradicional aplicado con anterioridad a los mismos estudiantes, especialmente si no ha transcurrido mucho tiempo entre ambas aplicaciones. En consecuencia, la comparación entre los resultados de ambos métodos no podría realizarse correctamente.

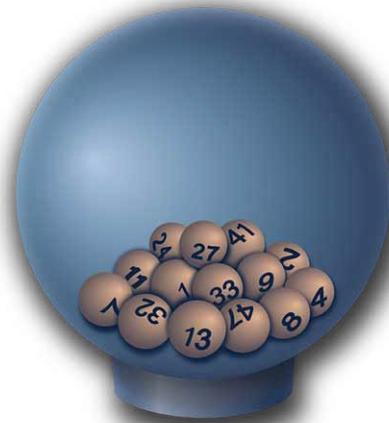
### AMPLIACIÓN. Métodos de aleatorización

Existen diversos métodos de selección aleatoria de elementos de una muestra:

En la aleatorización simple se selecciona al azar cada uno de los elementos de la muestra, de manera individual. En muestras grandes puede resultar una tarea muy compleja y costosa.

En la aleatorización sistemática se elige al azar un único elemento de una lista, y se escogen para la muestra ese elemento y los que se sitúan después de él en la lista cada cierto número de posiciones. Por ejemplo, si de una lista de 1000 elementos se quiere extraer una muestra de 20, se elige al azar un número  $x$  entre 1 y 50 (este último se obtiene de dividir 1000 entre 20), y se escogen los elementos con numeración  $x$ ,  $x+50$ ,  $x+100$ ,  $x+150$ , y así sucesivamente hasta acabar la lista, resultando la muestra deseada de 20 elementos. La ventaja de este método es que solo precisa la generación de un número aleatorio, pero puede no ser válido si la lista que se usa sigue algún tipo de orden, ya que la selección perdería el carácter aleatorio.

En la aleatorización estratificada se divide la población en estratos, que son conjuntos de elementos con valores iguales o similares de algunas variables potencialmente perturbadoras. A continuación, se elige al azar un número determinado de elementos de cada estrato y se reparten de manera equitativa entre los diversos grupos en que se quiere dividir la muestra. De este modo, la muestra al completo y cada uno de sus grupos reproducen mejor la variabilidad y estructura de la población. En ciencias sociales se suelen utilizar estratos respecto al sexo, edad, nivel educativo, nivel económico, ocupación, etc. La muestra puede contener el mismo número de elementos de cada estrato (si se quiere estudiar todos ellos en igualdad de condiciones), o bien un número proporcional al tamaño de cada estrato (que es más representativo de la estructura de la población).



Licencia: dominio público. Fuente: Pixabay.

En la aleatorización por conglomerados se escogen al azar grupos completos dentro de la población, que se suponen suficientemente homogéneos. Es especialmente útil cuando ya existen grupos previamente formados, como por ejemplo pueblos, barrios, centros educativos, centros de salud, etc. También puede aplicarse una selección por conglomerados de manera sucesiva, comenzando por los más amplios y reduciendo progresivamente el tamaño. Por ejemplo, se eligen al azar unas cuantas poblaciones del territorio, dentro de cada una de ellas se eligen al azar unos centros educativos, y dentro de cada uno de ellos se eligen al azar unos grupos de clase.

► **AMPLIACIÓN. Simbolización del control y el diseño de Solomon**

Es habitual, especialmente en ámbitos como la psicología, la educación o similares, resumir de forma simbólica el conjunto de procedimientos que se llevan a cabo en una investigación, incluyendo los relativos al control de variables perturbadoras que acabamos de ver.

Si la muestra se divide en grupos según los posibles valores de la variable independiente, cada uno de ellos se representa en una fila distinta. Cada fila se inicia con el símbolo R si los sujetos de ese grupo han sido escogidos de manera completamente aleatoria, con una R en el interior de un recuadro si la aleatorización ha sido solamente parcial, y en blanco si no ha habido aleatorización. Otros posibles métodos de asignación de elementos a grupos pueden también simbolizarse.

A continuación, se suceden símbolos O, que representan una medida (observación) de la variable dependiente, y símbolos X (o Y o Z), que representan aplicaciones de la variable independiente (tratamientos), como, por ejemplo, el empleo de un método didáctico en el aula. Los símbolos O y X pueden llevar un subíndice para diferenciar entre sí los procedimientos aplicados al mismo grupo (en la misma fila) o a grupos distintos. Todos ellos se sitúan de izquierda a derecha en el orden cronológico en el que se aplican.

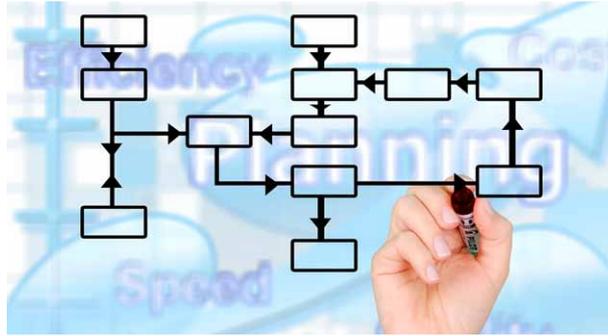
Un ejemplo de diseño de tipo experimental es el de cuatro grupos de Solomon (por el psicólogo estadounidense Richard Solomon, que lo desarrolló en 1949):

R	O <sub>1</sub>	X	O <sub>2</sub>
R	O <sub>3</sub>		O <sub>4</sub>
R		X	O <sub>5</sub>
R			O <sub>6</sub>

Esta simbolización indica que los elementos del primer grupo (primera fila) son asignados a él completamente al azar (R), y se les mide la variable dependiente (O<sub>1</sub>) antes de aplicar ningún tratamiento; esta medida suele llamarse pretest. A continuación, se les aplica un tratamiento (X), que constituye uno de los posibles valores de la variable independiente en el estudio en cuestión, y finalmente se vuelve a medir la variable dependiente (O<sub>2</sub>); esta medida se suele llamar posttest. De este modo, es posible estudiar el efecto que ha tenido la aplicación X en los elementos del grupo analizando los cambios en la variable dependiente deducidos de la comparación entre las medidas O<sub>1</sub> y O<sub>2</sub>.

El resto de grupos actúan como controles, de la manera descrita a continuación. En el segundo grupo se efectúan dos medidas de la variable dependiente, pero sin aplicar tratamiento entre ellas, y su comparación con el primer grupo permite deducir si el cambio en la variable dependiente podría deberse al mero paso del tiempo (desde que se inicia el estudio hasta que se realiza el último posttest), en lugar de deberse al efecto del tratamiento. El tercer grupo es como el primero, pero sin pretest, de modo que al compararlos se puede averiguar si el simple hecho de medir previamente la variable dependiente interacciona con el tratamiento X, modificando su efecto sobre los elementos. Por último, el cuarto grupo es como el segundo, pero sin pretest, y la comparación entre ambos permite concluir si el simple hecho de medir previamente la variable dependiente modifica los resultados obtenidos en la segunda medida realizada un tiempo después, sin haber tratamiento de por medio.

Los efectos controlados en este ejemplo de diseño están asociados a lo que se denominan fuentes de invalidez de una investigación, como son la interacción entre medidas (efecto memoria), interacción entre medidas y tratamientos, efecto maduración, etc. En el capítulo siguiente los describiremos en detalle y explicaremos por qué pueden ser tan perjudiciales para una investigación, hasta el punto de llegar a invalidar sus conclusiones.



Licencia: dominio público. Fuente: Pixabay.

Esquema del capítulo<sup>4</sup>



Fuente: elaboración propia con Genially.

<sup>4</sup> Enlace relacionado: <http://view.genial.ly/5c7119d87d15495764eb0b30/interactive-content-control>

# Capítulo 12.

## Validez de una investigación

El diseño de una investigación debe asegurar que las conclusiones que se obtengan sean plenamente válidas. Al reflexionar sobre las circunstancias que suelen afectar a la validez de las investigaciones podemos introducir mejoras, si contamos con los recursos suficientes, o al menos podemos identificar los puntos débiles, lo que puede ayudar a defender nuestro trabajo de manera más rigurosa y honesta.

Se pueden distinguir los siguientes tipos de validez:

– Validez interna:

Una investigación tiene que permitir contrastar una hipótesis de partida, que generalmente se refiere al efecto que tiene una variable independiente sobre una variable dependiente. Esto debe hacerse con la máxima garantía posible de que no estén actuando otras variables adicionales (perturbadoras) que influyan en la relación entre la independiente y la dependiente. Para ello se emplean procedimientos de control de variables perturbadoras como los introducidos en el capítulo anterior, mejorando así la validez interna del diseño de la investigación.

– Validez externa:

Una investigación analiza normalmente un conjunto reducido de elementos (hechos, situaciones, objetos, personas...), que son los que constituyen la muestra. Ese conjunto no suele coincidir con el total de elementos existentes o posibles, que corresponde a la población, ya que esta no sería manejable. Pero las conclusiones extraídas de la muestra sí que deben poder ser generalizables o extrapolables a la población completa, no con absoluta certeza, pero sí con una cierta probabilidad bien definida. Para conseguirlo, la muestra ha de ser lo más representativa posible de la población, siguiendo un cuidadoso procedimiento de selección de elementos, particularmente la selección aleatoria. Estas consideraciones en torno a la generalización de las conclusiones constituyen la validez externa de la investigación.

– Validez de constructo:

Las investigaciones tratan sobre conceptos, o relaciones entre ellos, que tienen que estar bien definidos para que otros investigadores entiendan a qué se refieren de manera precisa. Además, los procedimientos de la investigación implican registrar información acerca de esos conceptos, es decir, medirlos en los elementos de la muestra en estudio. Es esencial que la relación entre los conceptos y las variables que los miden sea lo más concreta y específica posible, y que los valores asignados a la variable sean adecuados. Estos procedimientos, que forman parte de la operacionalización de la investigación, afectan a su validez de constructo. En la mayoría de las investigaciones en ciencias naturales la operacionalización está bien establecida, pero resulta mucho más difícil en áreas como la psicología, donde los conceptos son constructos mentales sobre cuya definición o medida puede no haber consenso, como por ejemplo la creatividad o la inteligencia. En estos casos se podrían plantear cuestiones como: ¿están relacionadas las variables medidas a través de una cierta herramienta (test, encuesta...) con lo que se entiende por creatividad de una persona? ¿o esas variables miden solamente algo parecido, o quizá algo totalmente distinto?; si empleásemos otras variables, ¿los resultados serían similares? (es decir, ¿convergerían?).

– Validez estadística:

Es necesario que las técnicas estadísticas empleadas en la investigación sean adecuadas para generalizar los resultados de la muestra a toda la población. El procedimiento de generalización implica el uso de estadística inferencial, que proporciona resultados sobre la población a partir de los obtenidos en las muestras con un cierto grado conocido de probabilidad. Las técnicas de estadística inferencial dependen del tamaño y selección de la muestra y de la característica de la población que se quiera estudiar; elegir la técnica adecuada en cada caso dota a la investigación de validez estadística.

## 12.1. Efectos sobre la validez interna

Existen varias circunstancias que pueden influir en la validez interna de una investigación, como son los efectos de la historia, la maduración, el aprendizaje, la instrumentación, la selección de elementos, la mortalidad experimental o la regresión a la media. Si varios de ellos ocurren conjuntamente, pueden potenciar sus efectos negativos sobre la validez de la investigación.

– Efecto historia:

El efecto historia se refiere a los sucesos imprevistos que pueden ocurrir durante la manipulación experimental o durante la recogida de datos, y que pueden influirlos. Por ejemplo, la aparición de un molesto ruido de obras durante la aplicación de un test de inteligencia o un corte de luz cuando se están haciendo cuidadosas medidas de laboratorio. Cuanto más tiempo se extienda la experimentación y mayor variedad de localizaciones requiera, más probabilidades habrá de que sucedan imprevistos. Para ayudar a reducirlos conviene planificar detalladamente la investigación o realizar todos los procedimientos experimentales en las mismas circunstancias. Por ejemplo, si hay ruido ambiente muy molesto durante la aplicación de un test y no se puede evitar, que al menos afecte a todos los sujetos o grupos de la muestra por igual, y no a unos sí y a otros no dependiendo del día o lugar en que se realice.

– Efecto maduración:

Otro posible efecto es el de maduración, que puede ocurrir cuando en una investigación se mide una variable dependiente antes y después de que actúe una variable independiente por aplicación de un tratamiento. El objetivo sería determinar la influencia de la independiente en la dependiente comparando el valor final de esta con el estado de partida, pero si transcurre demasiado tiempo entre las dos medidas, quizá la variación se deba al simple paso del tiempo, y no directamente a la acción de la variable independiente. Un ejemplo clásico es el de un método didáctico aplicado a un grupo de estudiantes, en los que se miden los resultados académicos antes y después de la aplicación del método. Si transcurre demasiado tiempo entre ambas medidas, por ejemplo un curso académico, y se observa una notable mejoría, quizá se deba simplemente a la maduración intelectual natural de los jóvenes, y no tanto al método didáctico aplicado.

– Efecto memoria:

Cuando la variable dependiente se mide dos veces en los mismos elementos, puede actuar también el efecto memoria o aprendizaje, que consiste en que la primera medida afecta a los resultados de la segunda. En el ejemplo sobre el método didáctico, se medía el rendimiento académico antes y después de su aplicación. Por un lado, es conveniente emplear el mismo examen en las dos ocasiones para facilitar su comparación, pero, por otro lado, es posible que los estudiantes aprendan del primer examen y que los resultados del segundo sean mejores, sin que por ello pueda concluirse que el método didáctico aplicado entre ambos sea eficaz. Este problema puede darse siempre que exista el riesgo de aprender en la primera medida, aunque la herramienta no sea idéntica a la de la segunda, y su efecto puede reducirse si se deja transcurrir mucho tiempo entre ambas medidas. En investigaciones en ciencias naturales este efecto suele ser menos relevante (aunque resulta ineludible en física cuántica, donde el propio acto de medir una propiedad puede cambiarla o restringir la precisión con la que se miden otras).

– Efecto instrumentación:

Siempre que se toman datos en distintos grupos dentro de una muestra, o en un mismo grupo pero en distintos momentos, debe hacerse de la manera más homogénea posible, para evitar el efecto de la instrumentación. Este efecto consiste en obtener distintas medidas de la variable dependiente debido a las diferentes condiciones en la toma de datos, entre las que se encuentran los instrumentos o herramientas de recogida de datos, la persona que los usa o aplica, las condiciones del entorno, etc. Las variaciones de la variable dependiente podrían deberse a esas circunstancias cambiantes y no necesariamente al efecto de la variable independiente.

– Efecto selección:

En algunos tipos de investigación se divide la muestra en distintos grupos, cada uno de ellos con un valor distinto de la variable independiente para evidenciar si la variable dependiente difiere de un grupo a otro. Para que esa dependencia se vea claramente, los grupos tienen que ser homogéneos en todas aquellas variables perturbadoras que puedan interferir en la relación. En el ejemplo de dos grupos en los que se aplican dos métodos didácticos distintos (nuevo y tradicional), si uno de los grupos es de primero de la ESO y el otro de Bachillerato, los resultados serán muy distintos debido a la diferente madurez intelectual de los estudiantes y no necesariamente a la influencia del método. Este efecto de selección de los elementos puede reducirse escogiéndolos aleatoriamente.

– Efecto de mortalidad experimental:

Con el paso del tiempo algunos de los elementos de una muestra pueden dejar de estar accesibles para la investigación, o sus propiedades pueden cambiar. El efecto de mortalidad experimental implica que una muestra que comienza siendo suficientemente representativa de una población puede dejar de serlo porque se reduce mucho su tamaño o porque queda desequilibrada si cambian o desaparecen muchos elementos con alguna característica en común. Por ejemplo, si una investigación sobre un método didáctico para secundaria se alargase mucho en el tiempo, algunos de los estudiantes terminarían ese nivel educativo y ya no podrían seguir formando parte de la muestra, que se reduciría considerablemente.

– Efecto de regresión a la media:

Los valores extremos (los más altos y los más bajos) de la variable dependiente en una muestra se deben en general a la combinación de dos contribuciones. Por un lado, el efecto causal de la variable independiente u otras variables perturbadoras, y por otro lado, el efecto del azar, ya sea como variaciones aleatorias en las propiedades de los elementos o en el propio proceso de medida de esas propiedades (errores de medida).

Por ejemplo, en una prueba de conocimientos cabe esperar que los mejores resultados correspondan a los estudiantes que han seguido un mejor método didáctico, a los más inteligentes, a los más motivados, etc.; todas ellas son posibles variables independientes o perturbadoras. Las pun-

tuaciones extremas por arriba corresponderán a estudiantes que, además de lo anterior, estuvieron especialmente lúcidos durante el examen o habían estudiado particularmente bien las preguntas que cayeron. En otras palabras, tuvieron «buena suerte». Un razonamiento análogo se puede aplicar a los peores resultados de la prueba, en los que podrá haber influido un peor método didáctico, menor inteligencia de los sujetos, menor motivación, etc., pero también se habrán visto afectadas por una cierta dosis de «mala suerte». Lo que hemos denominado «buena» y «mala» suerte no es más que el inevitable azar actuando en una u otra dirección.

Si los sujetos vuelven a realizar la prueba en otro momento, el efecto de las variables independientes y perturbadoras será previsiblemente el mismo de antes, pero el azar, por su naturaleza, puede actuar en un sentido distinto. Es de esperar, por tanto, que el subconjunto de estudiantes con notas extremas superiores empeore sus resultados en la siguiente prueba, mientras que el subconjunto con notas extremas inferiores, las mejore. Es decir, ambos grupos de puntuaciones extremadamente alejadas de la media de la muestra se desplazan (regresan) hacia ella en la segunda medida, lo que se denomina efecto de regresión a la media. No existe una explicación causal para ello, sino el mero efecto del azar actuando en la segunda prueba en sentido opuesto a la primera.

Todo lo anterior influye en la validez cuando se selecciona una muestra con valores extremos de la variable dependiente o en alguna otra muy relacionada con ella. Por ejemplo, en una investigación sobre un nuevo tratamiento médico se seleccionan las personas más enfermas, pensando que serán las que mejor reflejen el posible beneficio. Una vez aplicado el tratamiento, se vuelve a estudiar la muestra, comprobándose que los enfermos presentan una mejoría. Este cambio puede deberse al efecto de regresión a la media, y no necesariamente a la eficacia del tratamiento.

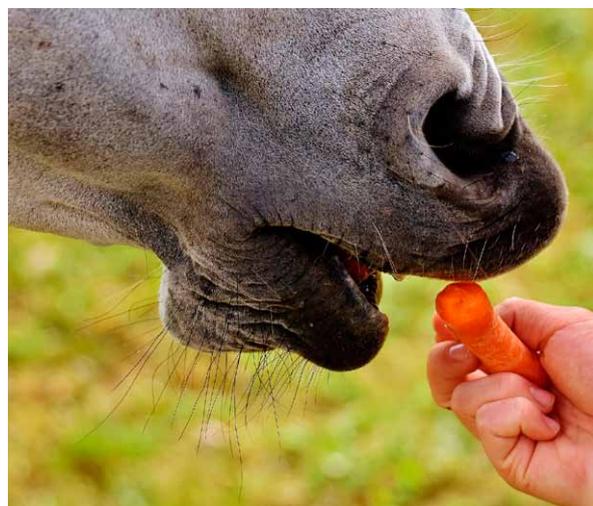
Una manera de eliminar esta fuente de invalidez en muestras extremas es llevar a cabo la investigación con dos grupos, uno de los cuales recibe el nuevo tratamiento y el otro (grupo de control), no. En el segundo examen médico ambos grupos habrán mejorado por efecto de la regresión a la media, pero el grupo que ha recibido el nuevo tratamiento puede haber mejorado más que el de control, de lo que se puede deducir sin confusión la eficacia del tratamiento.

### HISTORIAS DE LA CIENCIA. De la regresión a la media a la regresión en la educación

Existen numerosas evidencias, tanto en animales como en humanos, de que enseñar o entrenar en ciertas tareas resulta más efectivo cuando se recompensan los avances que cuando se castiga el estancamiento o el retroceso. Es decir, en una táctica de palo y zanahoria, es mucho más efectiva la zanahoria (la recompensa o refuerzo positivo) que el palo (el castigo).

El Premio Nobel de Economía Daniel Kahneman cuenta la anécdota de un oficial de la fuerza aérea israelí que le rebatió esta teoría basándose en su propia experiencia. El oficial aseguraba que los cadetes a los que felicitaba por haber realizado un excelente ejercicio aéreo casi siempre empeoraban en la siguiente ocasión. En cambio, los cadetes a los que reprendía airadamente por su nefasta ejecución mejoraban considerablemente en el siguiente ejercicio. Así pues, este oficial tenía muy claro que en el entrenamiento funciona mejor el castigo que la recompensa.

Kahneman reflexiona acerca de la aparente contradicción entre las evidencias acumuladas en psicología y ciencias del comportamiento y la experiencia del oficial. Lo que observaba



Licencia: dominio público. Fuente: Pixabay.

este último no tenía nada que ver con los efectos neurológicos de castigos o recompensas, sino que era un mero efecto estadístico de regresión a la media. La ejecución excepcionalmente buena de un ejercicio aéreo se deberá probablemente a la pericia del piloto y a un conjunto de efectos aleatorios que en esa ocasión han jugado todos a la vez a su favor (estado mental, condiciones meteorológicas, etc.). Con recompensa hacia el piloto o sin ella, parte de esos efectos aleatorios seguramente no actuarán en el mismo sentido en el siguiente ejercicio, con lo que la ejecución empeorará (sin dejar de ser buena, dado que la pericia del piloto, que no ha cambiado, será seguramente el factor más relevante). Análogamente, un ejercicio excepcionalmente deficiente se deberá probablemente a la peor formación del piloto junto con diversos efectos aleatorios actuando todos ellos como obstáculos. Parte de estos últimos influirán de manera distinta en la siguiente ocasión, dando lugar a un resultado mejor que el anterior, con o sin bronca del instructor.

En conclusión, es importante tener en cuenta el efecto de regresión a la media para evitar la regresión en los métodos de enseñanza (vuelta al castigo), así como para no afectar a la validez de una investigación.

Para saber más: Kahneman, D. (2011) *Thinking fast and slow*. New York: Farrar, Strauss and Giroux.

## 12.2. Efectos sobre la validez externa

Entre las circunstancias que pueden ser origen de invalidez externa en una investigación se pueden incluir las siguientes:

### – Efecto de interacción entre medidas y tratamientos:

En algunas investigaciones experimentales se realiza una medida previa de la variable dependiente (llamada a veces pretest) para conocer el punto de partida o estado inicial de los elementos de la muestra. A continuación, se introduce la variable independiente en forma de tratamiento sobre los elementos de la muestra. Finalmente, se mide de nuevo la variable dependiente en los elementos (postest) para poder analizar el efecto que ha tenido sobre ellos la aplicación de la variable independiente.

El problema de validez en este procedimiento puede surgir cuando la realización de medidas previas afecta al modo en que la variable independiente influye sobre los elementos de la muestra, haciéndolos más o menos sensibles, receptivos o reactivos a los tratamientos asociados a ella. Este efecto no tendría lugar, evidentemente, en los elementos de la población en general, que no han sido sometidos a la medida previa. Este es el origen de la invalidez externa, ya que los sujetos de la muestra, por efecto de las medidas previas, pueden dejar de ser representativos de la población, de modo que las conclusiones de la investigación pueden no ser generalizables.

### – Efecto de interacción entre tratamientos múltiples:

Es similar al efecto anterior, pero ocurre cuando a la misma muestra se le aplican diferentes tratamientos (distintos valores de la variable independiente) a lo largo del tiempo. Los efectos de tratamientos anteriores pueden persistir cuando se aplica el siguiente, aumentando o reduciendo su influencia. En la población donde se quieren extrapolar los resultados no se aplicarían varios tratamientos a la vez, sino solo uno de ellos, por lo que las condiciones diferirían de las de la muestra estudiada, dificultando la generalización de resultados.

Este efecto se puede evitar dividiendo la muestra en varios grupos y aplicando un tratamiento distinto en cada uno de ellos.

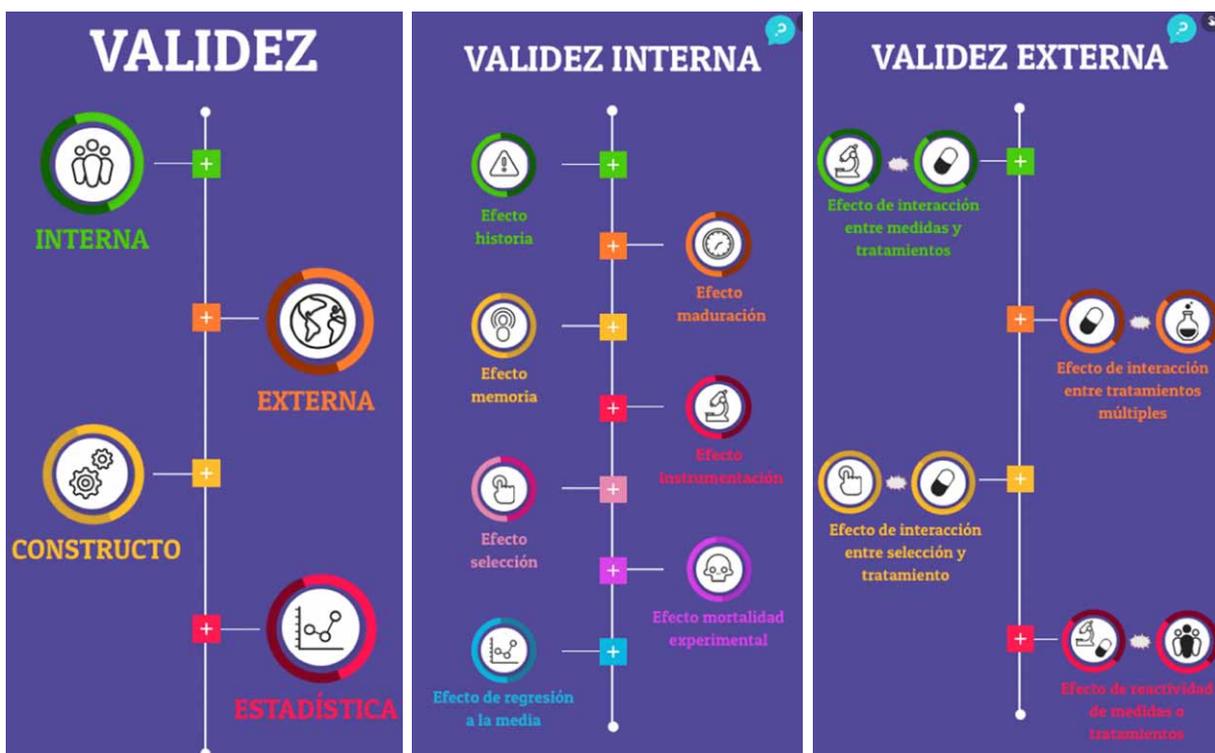
### – Efecto de interacción entre selección de la muestra y tratamientos:

La selección de los elementos que forman la muestra de estudio también puede afectar a la validez externa de la investigación. Si la variedad de los elementos de la muestra no refleja la variedad en la población, es posible que los efectos de los tratamientos difieran en ambos conjuntos.

La selección de los elementos de la muestra puede verse sesgada por el hecho de que el experimentador tenga más facilidad para acceder a unos elementos que a otros, quedando los segundos infrarrepresentados en su estudio. Cuando los elementos de la muestra son personas (por ejemplo, en investigaciones en ciencias sociales o ciencias de la salud), algunos sujetos o colectivos pueden ser más proclives o más reacios a participar en el estudio. Resulta especialmente importante reflexionar acerca de si la muestra está formada en gran medida por voluntarios, ya que ese simple hecho puede implicar que no sea en absoluto representativa de la población.

- Efecto de reactividad de tratamientos y medidas sobre los elementos de la muestra:  
Este efecto ocurre exclusivamente en investigaciones con personas, como estudios en ciencias sociales o ciencias de la salud y disciplinas afines. Se refiere a la reacción que tienen los individuos al ser conscientes de estar participando en una investigación, que suele ir acompañada de situaciones excepcionales. La aplicación de tratamientos o la toma de medidas suele darse en circunstancias desconocidas o poco habituales para los sujetos, que incluyen lugares, horarios, personas, procedimientos, etc. Todo ello puede afectar a la sensibilidad de los sujetos a los tratamientos, marcando una diferencia con la población general que impide una correcta extrapolación de los resultados.

Esquema del capítulo<sup>5</sup>



Fuente: elaboración propia con Genially.

<sup>5</sup> Enlaces relacionados:

<http://view.genially.ly/5c509adfa9e79373d7e85de8/validez>

<http://view.genially.ly/5c4f81661ef86d73dd8ae7f6/interactive-content-validez-interna>

<http://view.genially.ly/5c504965a9e79373d7e7d6da/interactive-content-validez-externa>

## **PARTE III**



# Capítulo 13.

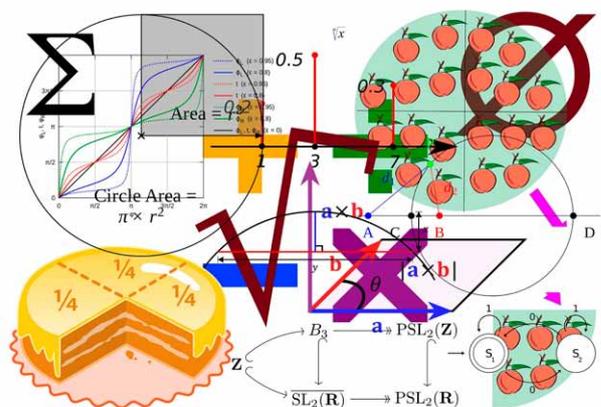
## Estadística descriptiva

Cuando se recoge una gran cantidad de datos cuantitativos en una investigación es necesario organizarlos y caracterizarlos de manera concisa para que el estudio resulte manejable. Esto se consigue mediante una serie de operaciones matemáticas sobre los datos extraídos de la muestra de la investigación. Evidentemente, la descripción resultante ignora muchos detalles de la muestra, pero conserva los rasgos más importantes y permite analizarla más fácilmente.

La estadística descriptiva es la rama de las matemáticas que define e interpreta los valores que describen de manera concisa un conjunto de datos, que son de varios tipos: medidas de centralización, medidas de dispersión, medidas de posición, medidas de correlación, etc.

### 13.1. Medidas de centralización

Las medidas de centralización o de tendencia central expresan el valor en torno al cual se sitúan los datos de una muestra. La medida más habitual de este tipo es la media aritmética, que se obtiene sumando todos los datos y dividiendo el resultado entre el número total de datos sumados. Se trata de un cálculo con el que todos estamos muy familiarizados, y en particular los estudiantes cuando quieren saber su nota final del curso a partir, por ejemplo, de las notas de varios exámenes parciales o de las evaluaciones trimestrales.



Licencia: dominio público. Fuente: Pixabay.

La media de un conjunto de  $n$  datos denotados  $x_i$  se suele representar por el símbolo  $\bar{x}$  con una barra horizontal encima, aunque aquí la llamaremos  $m$ . Su expresión matemática es:

$$m = \frac{\sum_{i=1}^n x_i}{n}$$

donde el numerador contiene la suma de todos los valores  $x_i$  cuyo subíndice va cambiando desde 1 hasta el último que haya, representado por  $n$ :  $x_1, x_2, x_3, \dots, x_n$ .

Cuando hay valores que se repiten en el conjunto de datos, en vez de sumarlos uno a uno se pueden multiplicar por el número de repeticiones, que es su frecuencia. Al dividir entre el número total de datos no hay que olvidar aquellos que se repiten, que cuentan tantas veces como indique su frecuencia.

Una medida relacionada con la anterior es la media ponderada, en la que cada uno de los datos se multiplica por un valor, denominado peso, que refleja cuánto contribuye a la media. Los pesos de todos los valores tienen que sumar 1 (o 100 si están expresados en porcentaje). Volviendo al ejemplo de las notas, algunos ejercicios o exámenes pueden contribuir más que otros para la nota final del curso, es decir, tener diferente peso. Por ejemplo, un 60 % el examen, un 20 % los ejercicios de clase y un 20 % un trabajo; para calcular la media ponderada se multiplica la nota del examen por 0,6, la de los ejercicios por 0,2 y la del trabajo por 0,2, y se suman.

Otra medida de centralización es la moda, que es el valor que más se repite dentro del conjunto de datos. Por ejemplo, en el conjunto siguiente: 9 2 5 7 5 4 1 7 3 9 9, la moda es el 9, que se repite tres veces o, dicho de otro modo, su frecuencia es tres. Si existen varios valores que se repiten el mismo número de veces, se dice que la muestra es multimodal (bimodal si hay en concreto dos modas).

Por último, la mediana es el valor que se sitúa justo en la mitad de un conjunto de datos dispuestos en orden creciente o decreciente. Respecto a la mediana, la mitad de los datos del conjunto (el 50 %) quedan por encima de ella, y la otra mitad (el otro 50 %) quedan por debajo. Si el conjunto tiene un número impar de datos, la mediana es el que queda en la mitad de la lista ordenada. Si el conjunto tiene un número par de datos, la mediana se calcula como la media aritmética entre los dos valores que quedan en la mitad de la lista ordenada. Siguiendo con el ejemplo anterior, los datos ordenados quedan: 1 2 3 5 5 7 7 9 9 9. Como hay un número par de datos, en la mitad de la lista quedan dos valores, 5 y 7, y la mediana se obtiene como su media aritmética:  $(5+7)/2 = 6$ . Se observa que hay cinco datos con valor menor que 6 y otros cinco con valor mayor que 6.

### EJEMPLO

Los datos del Instituto Nacional de Estadística (INE) sobre los sueldos anuales brutos en España en el año 2017 permiten obtener el salario medio, que fue de 23 646 euros, el salario mediano, 19 830 euros, y el salario más frecuente (moda), 17 482 euros. Es evidente la diferencia entre estos tres valores, que pueden dar lugar a malentendidos. En concreto, la media puede resultar alta porque los salarios muy elevados, aunque no sean muy frecuentes, tienen un fuerte impacto en su cálculo; lo más probable es percibir el sueldo indicado en la moda, precisamente porque es el más frecuente.

### EJEMPLO

Supongamos que nos piden preparar una actividad para un grupo de diez estudiantes. Para adaptar la actividad, necesitaríamos saber la edad de los participantes, pero en lugar de darnos la lista de edades completa nos proporcionan una medida estadística de centralización,

la media aritmética, que es 10 años. Puede parecer que este dato es todo lo que necesitamos saber, pero en realidad podría no ser suficiente. Para los tres grupos siguientes se listan las edades de sus miembros, en años:

Grupo A: 10 10 10 10 10 10 10 10 10 10

Grupo B: 2 4 6 8 10 10 12 14 16 18

Grupo C: 4 4 4 4 4 16 16 16 16 16

Se puede comprobar fácilmente que la media aritmética es la misma en los tres grupos, 10 años, aunque salta a la vista lo diferentes que son. El grupo A es homogéneo, similar a los que encontraríamos en un colegio. El grupo B incluye bebés, niños pequeños, mayores y adolescentes, lo que haría bastante difícil diseñar una actividad. El grupo C también incluye niños pequeños y adolescentes, y en este caso resulta evidente que convendría hacer dos actividades separadas, una para los niños de 4 años y otra para los adolescentes de 16. Otras medidas de centralización no mejoran mucho la caracterización de estos grupos. En los tres la mediana es 10, y en los grupos A y B la moda también es 10, puesto que ese es el valor que más se repite; en el grupo C la moda es 4 o 16 (ambos valores con igual número de repeticiones).



Licencia: dominio público. Fuente: Pixabay.

Lo que diferencia claramente esos tres grupos es la variabilidad o dispersión de sus datos. Por eso las medidas de dispersión, que veremos a continuación, también contribuyen a la caracterización de un conjunto de datos, complementando a las medidas de centralización.

## HISTORIAS DE LA CIENCIA. La ciencia del Estado

Etimológicamente, la palabra estadística significa ciencia del Estado, origen que se debe al uso que gobernantes y administradores han hecho de esta rama de las matemáticas a lo largo de la historia. La recogida de datos y su manipulación estadística, más o menos desarrollada, se ha empleado en el ámbito estatal especialmente a partir del siglo xvii para organizar los diversos aspectos de una sociedad, como son los demográficos (censos, estadísticas de nacimientos, fallecimientos, etc.), comerciales y económicos, etc.

Uno de los primeros análisis de este tipo fue publicado en 1662 por el inglés John Graunt (1620-1674), que basándose en los datos semanales de fallecimientos en Londres estableció tendencias de mortalidad anual y estacional y las causas más frecuentes de enfermedad. Graunt es considerado por algunos como el primer demógrafo y epidemiólogo, pero su verdadera profesión era mercero.

A principios del siglo xix comenzó a emplearse la estadística en ámbitos distintos a los estatales, y hoy es una de las herramientas más importantes de las ciencias, tanto sociales como naturales. Sin embargo, para que la estadística arroje claridad sobre la ciencia, y no confusión, es necesario interpretar correctamente los resultados que ofrece. En relación con esto, el escritor norteamericano Mark Twain popularizó la siguiente expresión: «*There are three kinds of lies: lies, damned lies, and statistics*» («Existen tres tipos de mentiras: mentiras, mentiras absolutas y estadísticas»). Twain atribuyó la frase precisamente a un político, el primer ministro británico Benjamin Disraeli, aunque no está claro que realmente fuera suya. Otra crítica humorística en la misma dirección se debe al escritor irlandés Bernard Shaw:

«La estadística es una ciencia que demuestra que, si mi vecino tiene dos coches y yo ninguno, los dos tenemos uno» (en referencia a la media aritmética).

En realidad, nunca ha sido necesario saber de estadística para ser considerado un buen estadista, pero la confusión derivada de la inadecuada interpretación de resultados estadísticos sí que es empleada por muchos políticos. Un buen ejemplo es el de los dos coches del vecino, o la situación análoga en relación a los salarios medios: la existencia de unos pocos que reciben salarios desorbitados aumenta el salario medio, pero la cantidad que percibe la mayoría dista bastante de ese valor.



Licencia: dominio público. Fuente: Pixabay.

## 13.2. Medidas de dispersión

Las medidas de dispersión, o de variabilidad, expresan cómo se distribuyen los datos en torno a alguna de las medidas de centralización definidas antes, y son un complemento a estas últimas para describir más fielmente un conjunto de datos.

La medida de dispersión más sencilla que puede definirse es la distancia entre el mayor y el menor de los datos de un conjunto, denominada amplitud total, pero ignora mucha información al contar solo con dos valores para su cálculo.

Una medida de dispersión que depende de todos los datos de un conjunto consiste en calcular la distancia de cada uno de ellos a la media aritmética, es decir, su diferencia. Los datos que sean menores que la media darán una diferencia negativa, pero como únicamente interesa la distancia a la media, se ignora el signo menos empleando el valor absoluto. A continuación, se suman las diferencias obtenidas (todas ellas con signo positivo) y se divide el resultado entre el número de datos. Con ello se obtiene la desviación media respecto a la media aritmética, aunque también se puede calcular respecto a la mediana.

El hecho de ignorar el signo de las restas es importante, porque de otro modo las que son negativas compensarían en gran medida a las positivas, y su suma daría un valor pequeño que no reflejaría la verdadera dispersión de los datos. Esta medida no resulta práctica matemáticamente debido a la dificultad que presenta el manejo de valores absolutos. Por eso, y por otras muchas propiedades matemáticas importantes, es mucho más habitual elevar al cuadrado cada una de las diferencias calculadas, que también elimina los signos menos. A continuación, se suman los cuadrados de las diferencias y el resultado se divide entre el número de datos. La medida de dispersión así obtenida es la varianza, que se suele representar como  $s^2$  y que se puede calcular de la siguiente manera:

$$s^2 = \frac{\sum_{i=1}^n (x_i - m)^2}{n} = \frac{\sum_{i=1}^n x_i^2}{n} - m^2$$

La varianza refleja la dispersión de los datos de una manera un tanto distorsionada, porque las distancias a la media han sido elevadas al cuadrado. Por eso se define otra medida, la desviación típica o desviación estándar, representada por  $s$ , que es la raíz cuadrada de la varianza:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - m)^2}{n}}$$

Esta operación no deshace por completo la distorsión (la raíz de una suma no es la suma de las raíces de cada término), pero la atenúa. Otra ventaja de la desviación típica es que tiene las mismas unidades que los datos. Por ejemplo, en un conjunto de estaturas dadas en centímetros, la desviación típica, al igual que la media (o cualquier otra medida de centralización), se mide también en centímetros; la varianza, por el contrario, vendría dada en centímetros al cuadrado.

Una vez calculado el valor de la desviación típica es necesario interpretarlo, es decir, establecer si se trata de una dispersión grande o pequeña. Lo más apropiado es compararlo con el valor de los propios datos, y en concreto de su media aritmética. Así, no es lo mismo una desviación típica de 200 euros en un conjunto de salarios de 800 euros de media que en un conjunto de 5000 euros de media. Esa desviación implica grandes desigualdades en el primer caso, mientras que en el segundo caso es poco relevante. Para tener en cuenta estas circunstancias se define el coeficiente de variación (CV), que es la desviación típica dividida entre la media aritmética, es decir, proporciona el número de medias aritméticas que «cabén» en la desviación típica.

### EJEMPLO

Los tres grupos de personas del ejemplo anterior tenían la misma media de edad (10 años), pero la dispersión de los datos era evidentemente muy distinta, como se puede comprobar a través de la desviación típica: en el grupo A es 0, ya que no hay dispersión, en el grupo B es 4,9 años (varianza 24) y en el grupo C es 6 años (varianza 36).

La amplitud total (diferencia entre la edad más alta y la más baja) es mayor en el grupo B, 16 años, que en el grupo C, 12 años. Sin embargo, en este último todos los datos se encuentran alejados de la media, mientras que en el B los hay en la propia media y el resto se alejan progresivamente; por eso la desviación típica en el grupo B es menor que en el C.

### 13.3. Medidas de forma

Para caracterizar aún con más detalle un conjunto de datos se pueden calcular otras medidas relacionadas con su distribución, o medidas de forma. La curtosis expresa cuántos datos se sitúan muy lejos de la media (los denominados valores atípicos, *outliers* en inglés) y cómo de alejados están, en comparación con los datos que se encuentran próximos a ella. El nombre proviene del griego *kurtos*, que significa curvado o arqueado, y a veces se relaciona esta medida con la manera en que los datos se concentran en torno a la media (es decir, con la forma del pico de datos a su alrededor en una representación gráfica), pero en realidad, como se ha descrito, depende principalmente de los datos más alejados.

Otra medida de forma es la asimetría, que expresa la diferencia en la cantidad de datos que se sitúan a un lado de la media con respecto al otro, es decir, refleja si existe un exceso de datos por encima o por debajo de la media. Puede ocurrir que haya unos pocos datos mucho mayores que la media y muchos datos poco menores que ella, situaciones que se compensan entre sí y pueden dar una medida de asimetría pequeña, a pesar de que los datos se distribuyen de forma claramente asimétrica. La

curtosis puede reflejar este hecho, aportando así una información complementaria. En cualquier caso, ni la curtosis ni la asimetría son medidas especialmente eficientes a la hora de distinguir entre diferentes formas en que se distribuyen los datos de un conjunto, y siempre es preferible observarlo en una representación gráfica.

### EJEMPLO

El conjunto 2 3 4 7 10 1000 es claramente asimétrico: su media es 171, pero la mayoría de los datos (todos menos uno) se sitúan por debajo de ella; de hecho, la media es mucho mayor que la mayoría de los datos porque uno de ellos, 1000, está muy alejado de los demás: se trata de uno de los valores atípicos, que influye en la curtosis del conjunto.

## 13.4. Medidas de posición

Veámos antes que un conjunto ordenado de datos queda dividido por la mediana en dos mitades con el mismo número de datos cada una. De manera análoga se definen los cuartiles, que dividen el conjunto ordenado de datos en cuatro partes iguales. El primer cuartil ( $Q_1$ ) tiene una cuarta parte de los datos (un 25 %) por debajo de él, y las tres cuartas partes por encima (un 75 %). El segundo cuartil ( $Q_2$ ) tiene dos cuartas partes de los datos, es decir, la mitad, por debajo de él y la otra mitad por encima, por lo que coincide con la mediana. Por último, el tercer cuartil ( $Q_3$ ) tiene tres cuartas partes de los datos por debajo de él y una cuarta parte por encima.

A veces también se llama cuartiles a los grupos en que queda dividido el conjunto de datos, y no solo a los puntos de corte. Así, cuando se dice que un valor está situado en el primer cuartil significa que es mayor que el 75 % de los valores del conjunto, o que se encuentra entre el 25 % de valores más altos, es decir, su valor es mayor que  $Q_3$ . Estas medidas son muy usadas, por ejemplo, para caracterizar conjuntos de datos biométricos o psicológicos (como los resultados de un test de inteligencia).

A partir de los cuartiles se puede definir una medida de dispersión llamada recorrido semiintercuartílico, que se calcula como la diferencia entre los valores del tercer y del primer cuartil, dividida entre dos:  $(Q_3 - Q_1)/2$ . Su interés reside en que es muy poco sensible a la presencia de valores extremos (atípicos), al contrario que otras medidas como la desviación típica.

Análogamente se definen los deciles, que dividen el conjunto de datos ordenados en diez partes con el mismo número de datos cada una, y los centiles, que lo dividen en cien partes. Por ejemplo, el 10 % de los datos de un conjunto es menor que el primer decil, que coincide con el centil diez, y el 90 % restante de los datos es mayor. La mediana coincide con el quinto decil y con el centil 50. Todas estas medidas de posición en un conjunto ordenado de datos (mediana, cuartiles, deciles, centiles...) reciben el nombre conjunto de cuantiles.

## 13.5. Medidas de correlación

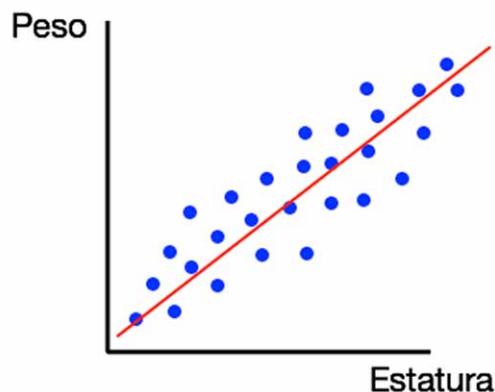
Las medidas de correlación se aplican a los valores que toman dos variables distintas en cada uno de los elementos de una muestra, y expresan si ambas variables varían a la vez, es decir, si los valores de una son más altos o bajos cuanto más altos o bajos son los de la otra. Si los valores de ambas variables aumentan o disminuyen a la vez, su correlación es positiva; en cambio, si los valores de una variable aumentan cuando los de la otra disminuyen, o viceversa, su correlación es negativa. Si las variables cambian de manera independiente una de la otra, no hay correlación entre ellas.

La pareja de valores de un elemento dado puede representarse como un punto en un plano cartesiano, en cuyo eje horizontal se sitúa el valor de una de las variables ( $x$ ) y en el eje vertical se sitúa

el valor de la otra variable ( $y$ ). El conjunto de elementos de la muestra queda representado entonces por una nube de puntos, y la correlación se evidencia si esa nube se aproxima a la forma de una línea recta, con distinta orientación según se trate de una correlación positiva o negativa. Si los puntos se sitúan exactamente sobre una recta, la correlación es perfecta, pero lo habitual es que tengan una cierta dispersión en torno a ella.

### EJEMPLO

Se miden las variables estatura y peso de un conjunto de personas. Es de esperar que entre ambas variables exista una correlación positiva, porque las personas con mayor estatura suelen tener también mayor peso, y las de menos estatura, menor peso. Sin embargo, esta relación no es estricta, porque puede ocurrir que una persona más baja que otra tenga mayor peso. La correlación, por tanto, no será perfecta, pero seguirá siendo bastante alta. En general, la correlación describe una tendencia, no una dependencia estricta.



Fuente: elaboración propia.

Si las parejas de valores de estatura y peso se representan en un diagrama cartesiano, la nube de puntos que aparece será alargada, tendiendo a seguir la forma de una línea recta con pendiente positiva, es decir, conforme aumenta la estatura aumenta también el peso, o viceversa. En esa nube se puede trazar gráficamente la recta de tendencia.

Para cuantificar el grado de correlación entre dos variables se puede calcular su covarianza, que expresa la dispersión de los valores (al igual que lo hace la varianza), pero de las dos variables a la vez:

$$s_{xy} = \frac{\sum_{i=1}^n (x_i - m_x) (y_i - m_y)}{n}$$

La covarianza es más alta cuanto mayor es la dispersión en torno a las medias de cada una de las variables ( $m_x$  y  $m_y$ ), pero eso es algo que no está estrictamente relacionado con su variación conjunta, es decir, con su correlación. Para eliminar el efecto de la dispersión de cada variable por separado, se puede dividir la covarianza entre la desviación típica de cada variable ( $s_x$  y  $s_y$ ), dando lugar al coeficiente de correlación de Pearson:

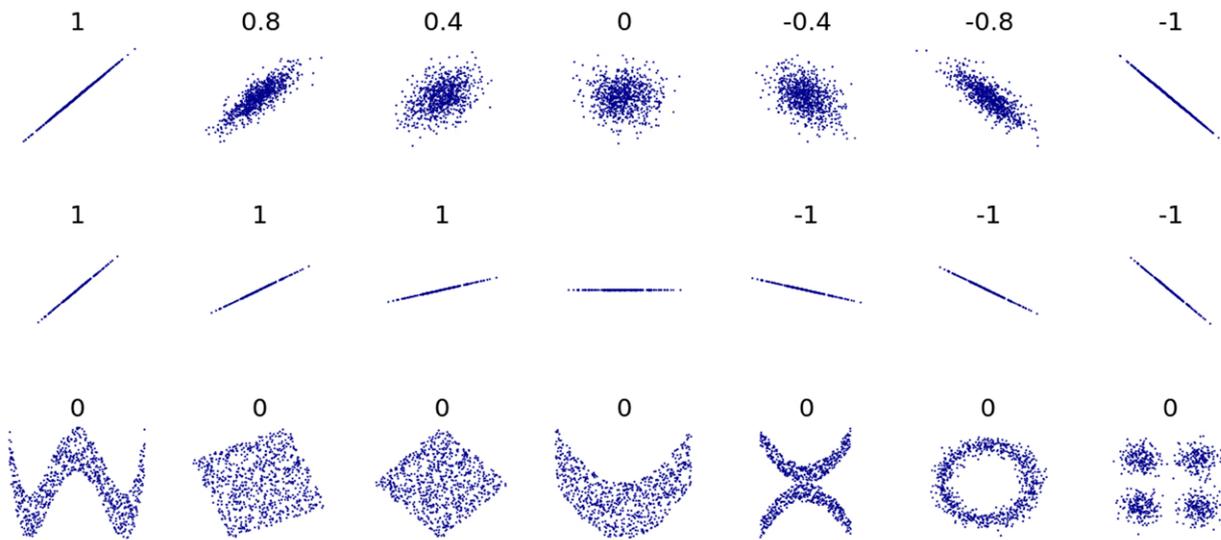
$$r_{xy} = \frac{s_{xy}}{s_x s_y}$$

Esta medida de correlación cuantifica el grado en que dos variables cambian conjuntamente, es decir, el grado en que se agrupan en torno a una cierta recta de tendencia.

El coeficiente de correlación varía entre -1 y +1, límites que expresan una correlación perfecta negativa o positiva, respectivamente; gráficamente, los puntos que representan los pares de valores se situarían exactamente sobre una recta. Los valores del coeficiente de correlación próximos, pero no

iguales, a -1 o +1, corresponden a una correlación alta, ya sea negativa o positiva. Valores del coeficiente cercanos a 0 indican ausencia de correlación, es decir, los valores de ambas variables varían de manera independiente.

En la siguiente figura aparecen distintas nubes de puntos para mostrar gráficamente los diversos grados de correlación, indicando además en cada caso el valor del coeficiente  $r_{xy}$ . En la primera fila se observan distribuciones con diferentes grados de correlación. En la segunda fila, la correlación en todos los casos es perfecta; el signo del coeficiente, como en la fila superior, depende de la orientación de la recta de tendencia; en la figura central una de las variables (la del eje vertical) no varía y no se puede establecer la correlación. En la tercera fila aparecen nubes de puntos variadas con correlaciones muy bajas o nulas.



Licencia: dominio público. Fuente: Wikimedia Commons.

Para una nube de puntos que muestra cierta correlación puede obtenerse la recta de tendencia que mejor se ajusta a todos los puntos. Matemáticamente, se calcula minimizando las distancias al cuadrado de cada uno de los puntos a la recta, de donde se obtiene la ecuación de la recta de regresión, que relaciona los valores de una de las variables ( $x$ ) con los valores de la otra ( $y$ ):

$$y = ax + b$$

El parámetro  $a$  es la pendiente de la recta (tangente del ángulo que forma con el eje horizontal, es decir, expresa su inclinación), y se calcula como:

$$a = \frac{n \sum_{i=1}^n x_i y_i - \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}$$

Por su parte, el parámetro  $b$  es la ordenada en el origen de la recta (valor de la variable  $y$ , representada en el eje de ordenadas del gráfico cartesiano, cuando la variable  $x$  toma el valor 0), y se calcula como:

$$b = \frac{\sum_{i=1}^n y_i - a \sum_{i=1}^n x_i}{n}$$

La recta de regresión se puede calcular para cualquier conjunto de pares de datos, pero no tiene sentido hacerlo cuando la correlación es muy pequeña o nula, es decir, cuando la nube de puntos es amorfa. El cálculo del coeficiente de correlación o de la pendiente y la ordenada en el origen de la recta de regresión resultan algo tediosos a partir de las fórmulas dadas aquí. Lo habitual es calcularlas en un ordenador usando un programa que tenga almacenadas funciones estadísticas, en el que únicamente hay que introducir los datos de la muestra.

#### ▶ AMPLIACIÓN. Otros coeficientes de correlación

El coeficiente de correlación de Pearson se emplea para dos variables continuas (medidas en escala de intervalo o de razón) y toma su nombre del matemático inglés Karl Pearson (1857-1936).

Para variables que se miden en otro tipo de escalas también se pueden definir coeficientes de correlación, como los siguientes:

- Coeficiente de Spearman, por el psicólogo inglés Charles Spearman (1863-1945), para variables ordinales, en las que solo se pueden establecer relaciones de orden entre sus categorías. A cada una de ellas se le asigna un valor que expresa su número de orden, y se calcula el coeficiente de correlación de Pearson usando esos valores.
- Coeficiente biserial-puntual, cuando una de las variables es continua y la otra es nominal formada por dos categorías (dicotómica).
- Coeficiente biserial, cuando una de las variables es continua y la otra también lo es inicialmente, pero sus valores se han agrupado en dos categorías, es decir, ha sido dicotomizada (por ejemplo, separando los valores según sean mayores o menores que la media aritmética del conjunto).
- Coeficiente cuádruple, cuando las dos variables son dicotómicas.
- Coeficiente tetracórico, cuando las dos variables eran inicialmente continuas, pero han sido dicotomizadas.



Karl Pearson. Licencia: dominio público.  
Fuente: Wikimedia Commons.

Esquema del capítulo<sup>6</sup>



Fuente: elaboración propia con Genially.

<sup>6</sup> Enlace relacionado: <http://view.genial.ly/5c76dd0f5d0e4e575e1ebd3f/estadistica-descriptiva>

# Capítulo 14.

## Estadística inferencial: distribución muestral de estadísticos

En el capítulo anterior se introdujeron una serie de valores numéricos, como la media y la desviación típica, que servían para describir el conjunto de datos obtenidos de la muestra de una investigación. En realidad, muchas investigaciones pretenden obtener información sobre la población de la que se ha extraído esa muestra, que es un conjunto mucho más amplio.

Para describir una población se emplean valores numéricos análogos a los de las muestras. En el primer caso reciben el nombre de parámetros poblacionales, y en el segundo, estadísticos muestrales. Así, la media o la desviación típica pueden ser estadísticos cuando se refieren a los datos de una muestra, o parámetros cuando se refieren a una población. Habitualmente, los estadísticos se simbolizan con caracteres latinos, y los parámetros correspondientes con los caracteres griegos equivalentes; por ejemplo, la media de una muestra se puede simbolizar como  $m$  y la de una población como  $\mu$ , o la desviación típica de una muestra como  $s$  y la de una población como  $\sigma$ .

Dado que suele ser inviable medir una población al completo, cualquier descripción sobre ella ha de ser inferida a partir de la muestra a la que sí se tiene acceso, procedimiento matemático del que se ocupa la estadística inferencial. Se pretende adquirir un conocimiento amplio y general (poblacional) a partir de medidas en un grupo particular (muestral), y se trata por tanto de una aplicación del método inductivo, que va de lo particular a lo general, al contrario que el deductivo. Como vimos en la parte I, la inducción conlleva incertidumbre. Si todo cisne que se observa es blanco y el tamaño de la muestra observada es cada vez mayor, se induce o infiere naturalmente que el color de los cisnes en general (es decir, de todos los cisnes que existen, la población), es blanco. Pero ya vimos que eso no tiene por qué ser cierto (de hecho, no lo es), y en todo caso solo podría confirmarse con certeza si realmente se observase toda la población.

La estadística inferencial no pretende ofrecer información sobre cada uno de los elementos de la población individualmente, sino que se emplea para obtener valores de ciertos parámetros poblacionales, como medias o proporciones, y en ese sentido representa una inducción más suave o flexible que en el ejemplo de los cisnes.

**EJEMPLO**

Para conocer la intención de voto de la población de un país no suele ser posible realizar encuestas a todos sus ciudadanos. Sería inviable acceder a todos ellos y, en todo caso, entrevistarlos requeriría una ingente cantidad de tiempo y recursos. Lo que se hace es seleccionar una muestra, por ejemplo de mil personas, entrevistarlas e inferir de sus respuestas ciertos parámetros de la población al completo, como la proporción de ciudadanos que votará a cada partido, pero, evidentemente, no lo que votará cada persona en particular. En un ejemplo como este, los parámetros poblacionales reales se conocerán el día de las elecciones, y se podrá comparar la inferencia proporcionada previamente a partir del análisis estadístico de las encuestas. Sin embargo, en la mayoría de los casos los parámetros poblacionales reales nunca llegan a conocerse, y en contextos más abstractos la población es de hecho infinita.

Incluso para ese cometido más restringido, que involucra únicamente la obtención de parámetros poblacionales, la estadística inferencial no puede ofrecer certeza sobre su valor. Para que los parámetros puedan inferirse correctamente a partir de los estadísticos muestrales es necesario que los elementos de la muestra estén tomados al azar de entre los de la población. La estadística inferencial permite cuantificar el efecto del azar, proporcionando la probabilidad de que el parámetro poblacional se encuentre en un determinado rango de valores.

Si los elementos de la muestra no se extraen al azar, se corre el riesgo de que no resulte representativa de la población, es decir, de que esté sesgada. Por ejemplo, si queremos conocer la estatura media de la población española y escogemos nuestra muestra en un club de baloncesto, es evidente que su estatura media va a diferir considerablemente de la estatura media de la población, debido al sesgo en la selección.

**EJEMPLO**

Se lleva a cabo una investigación que involucra la estatura media de una población. Por simplicidad, supongamos que la población completa consta solamente de diez personas, cuyas estaturas en metros son:

1,55 1,63 1,68 1,72 1,75 1,79 1,82 1,83 1,94 2,09

Como esta población es pequeña, se pueden calcular directamente parámetros como la media, que es de 1,78 metros. Supongamos ahora que, como ocurre habitualmente, no podemos manejar todos los datos de la población sino solo los de muestras más reducidas extraídas de ella. Supondremos aquí que solo tenemos recursos para estudiar una muestra de cuatro elementos, que podría ser alguna de las siguientes:

Muestra A: 1,82 1,83 1,94 2,09  
 Muestra B: 1,55 1,63 1,68 1,72  
 Muestra C: 1,55 1,72 1,75 1,94  
 Muestra D: 1,63 1,72 1,82 1,83

La muestra A sería el equivalente a acudir a un club de baloncesto para escoger los elementos de la muestra, porque resulta estar formada por las personas más altas de la población. La media de estatura en esta muestra es de 1,92 metros.

La muestra B también es muy particular, porque contiene los individuos de menor estatura, con una media de 1,65 metros.

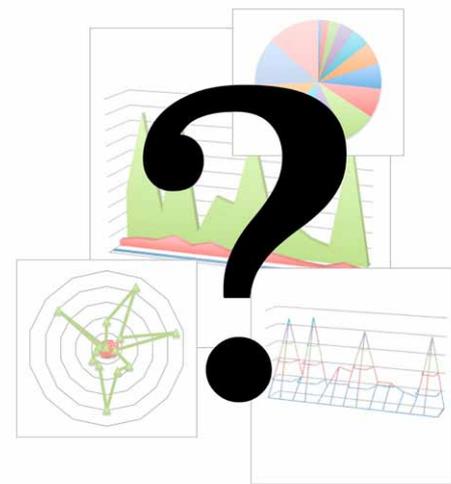
Las muestras C y D contienen una mayor variabilidad de estaturas, y sus medias son 1,74 y 1,75 metros, respectivamente.

Si escogemos al azar cuatro elementos de la población para formar nuestra muestra, podríamos haber obtenido con la misma probabilidad cualquiera de las muestras anteriores, o cualquier otra combinación de cuatro elementos (hay 210 posibles). Pero también es cierto que será mucho más probable que nuestra muestra sea del tipo C o D, porque hay muchas más combinaciones posibles en las que se mezclan estaturas muy distintas. Las muestras de tipo A o B son muy poco probables porque, como ya dijimos, son muy especiales: reúnen a los individuos de mayor y de menor altura, respectivamente.

En conclusión, es más probable que la muestra que extraigamos al azar para nuestra investigación tenga una media similar a la de las muestras C y D, en torno a los 1,75 metros, que resulta ser muy próxima a la media de la población, que en este ejemplo sí conocemos y es de 1,78 metros.

Recordemos que, en general, la media poblacional será desconocida, pero se puede inferir que, con alta probabilidad, será muy similar a la media obtenida en muestras aleatorias. Si la muestra obtenida al azar fuese la A o la B, y tomásemos su media como la de la población, nos estaríamos equivocando considerablemente, pero resulta poco probable (aunque no imposible) que nuestra muestra al azar sea de ese tipo.

En una investigación real no sabemos de qué tipo es la muestra que estamos usando, puesto que no conocemos toda la población de la que ha sido extraída. Desconocemos si nuestra muestra contiene un exceso de personas muy altas o de personas muy bajas, pero lo esperable, por ser lo más probable, es que resulte una buena representación de la población total.



Licencia: dominio público. Fuente: Pixabay.

## 14.1. Probabilidad

Todos tenemos cierta intuición de lo que significa la probabilidad, aunque es difícil definirla. La probabilidad puede entenderse como la facilidad o dificultad de que se produzca cada uno de los posibles resultados de un proceso aleatorio, es decir, con qué grado cabe esperar que suceda o no suceda cada uno de ellos. Si el proceso se repite muchas veces, la probabilidad de un cierto resultado se puede asociar al número de veces que se ha producido dividido entre el número total de repeticiones, que es la frecuencia relativa.

En el contexto de la probabilidad, los posibles resultados de un proceso se denominan sucesos. La estadística inferencial trata principalmente del suceso que consiste en obtener un determinado valor de un estadístico en una muestra extraída al azar de una población, conociendo cierta información sobre esa población (los valores de ciertos parámetros). También trata del suceso inverso, que consiste en que un parámetro de una población tome un determinado valor, conociendo cierta información de una muestra extraída al azar de esa población (los valores de ciertos estadísticos).

Matemáticamente, la probabilidad de un suceso se asocia a un valor entre 0 y 1, donde el 0 corresponde a un suceso imposible y el 1 corresponde a un suceso seguro. Si la probabilidad de un suceso  $A$  se denota como  $p(A)$ , la probabilidad del suceso contrario o complementario es  $1-p(A)$ .

### EJEMPLO

La probabilidad de sacar al azar un as de una baraja de 48 cartas es 0,083, que es un valor entre 0 y 1. La probabilidad del suceso complementario, que consiste en sacar cualquier carta que no sea un as, es  $1 - 0,083 = 0,917$  (también entre 0 y 1).

El suceso considerado (sacar un as) es compuesto, porque consiste en la unión de cuatro sucesos elementales: sacar el as de oros, el as de copas, el as de espadas o el as de bastos.

La probabilidad se ha obtenido dividiendo el número de sucesos individuales considerados, que son 4, entre el número de sucesos posibles, que son 48 (el número de cartas de la baraja):  $4/48 = 0,083$ . Este cálculo se ha podido realizar de esta manera porque la probabilidad de sacar al azar cualquiera de las cartas de la baraja se supone que es la misma. En general, no todos los sucesos posibles han de tener la misma probabilidad, como ocurriría en una baraja trucada. En esos casos es necesario conocer la probabilidad de cada uno de los sucesos elementales, que viene dada por una función de probabilidad.



Licencia: dominio público. Fuente: Pixabay.

### AMPLIACIÓN. Las matemáticas de la probabilidad

La probabilidad matemática tiene algunas otras propiedades importantes que están implícitas en los resultados de la estadística inferencial, y que se mencionan aquí por completitud.

La probabilidad de la unión de dos sucesos,  $p(A \cup B)$ , es la probabilidad de que ocurra alguno de los dos y se calcula como la suma de la probabilidad de ambos menos la probabilidad de su intersección,  $p(A \cap B)$ :

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

La intersección de dos sucesos se refiere al conjunto de sucesos elementales que forman parte de ambos a la vez, y es necesario restar su probabilidad para no contarla dos veces. Si los dos sucesos no tienen elementos en común se dice que son incompatibles, y la probabilidad de su intersección es 0. Cuando sí tienen elementos en común se dice que son compatibles, y la probabilidad de su intersección viene dada por el producto de la probabilidad de cada uno de ellos, siempre que ambos sean independientes entre sí, es decir, que la probabilidad de uno no dependa de que el otro suceso ocurra o no:

$$p(A \cap B) = p(A) \cdot p(B)$$

Si los sucesos son dependientes entre sí, entonces la probabilidad de su intersección es la probabilidad del primero multiplicada por la probabilidad del segundo condicionada a que ocurra el primero, que se simboliza como  $p(B/A)$ ; también es cierto si se intercambian ambos sucesos: la probabilidad de la intersección es la probabilidad del segundo por la del primero condicionada a que ocurra el segundo,  $p(A/B)$ :

$$p(A \cap B) = p(A) \cdot p(B/A) = p(B) \cdot p(A/B)$$

De esta igualdad se deduce una expresión para la probabilidad condicionada que se conoce como teorema de Bayes, por el matemático y filósofo inglés Thomas Bayes (1701-1761):

$$p(A/B) = \frac{p(A) \cdot p(B/A)}{p(B)}$$

Este teorema es fundamental en probabilidad y en estadística inferencial y está relacionado, por ejemplo, con los errores en el contraste de hipótesis y en particular con algunos ejemplos que se darán posteriormente sobre los conceptos de falso positivo y falso negativo.

### EJEMPLO

Si el suceso  $A$  consiste en sacar un as de una baraja y el suceso  $B$  consiste en sacar una carta de oros, la intersección entre ambos consiste en sacar el as de oros. Suponiendo que todas las cartas tienen la misma probabilidad de salir, la probabilidad de sacar un as es  $p(A) = 4/48$  (porque hay 4 ases en la baraja) y la de sacar una carta de oros es  $p(B) = 12/48$  (porque hay 12 oros en la baraja). Como son sucesos independientes, la probabilidad de su intersección es el producto de las probabilidades individuales:  $p(A \cap B) = p(A) p(B) = 1/48$ . Este resultado se puede obtener también directamente: la intersección corresponde a sacar un as de oros, y como solo hay uno en la baraja, la probabilidad es  $1/48$ . La unión de ambos sucesos consiste en sacar un as o una carta de oros, y teniendo en cuenta que son compatibles (su intersección no es cero), su probabilidad viene dada por:  $p(A \cup B) = p(A) + p(B) - p(A \cap B) = 15/48$ .



Licencia: dominio público.  
Fuente: Wikimedia Commons.

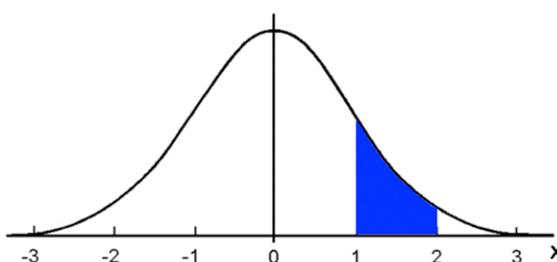
## 14.2. Distribución muestral de estadísticos

En cada una de las distintas muestras que pueden ser extraídas de una población se pueden calcular estadísticos como la media aritmética o la proporción de elementos que presentan cierta característica; por ejemplo, la media de estaturas o la proporción de licenciados universitarios. Cuando los elementos son escogidos de manera aleatoria, los estadísticos pueden tomar distintos valores en cada una de las muestras, cada uno de ellos con distinta probabilidad. En los ejemplos del inicio de este capítulo ya vimos que los valores de la media en diferentes muestras aleatorias se encontraban con mayor probabilidad cerca del valor de la media poblacional, y que era menos probable que se encontrasen muy alejados de ella.

La probabilidad de cada uno de los posibles valores que puede tomar un estadístico en muestras extraídas al azar viene dada por una función matemática denominada distribución muestral, que depende del estadístico en cuestión. Se habla así, por ejemplo, de la distribución muestral de la media aritmética o de la distribución muestral de la proporción.

Una distribución muestral es una función de probabilidad, ya que asigna a cada posible valor de un estadístico su probabilidad de aparecer en una muestra extraída al azar. En realidad, esta definición es estrictamente cierta solo cuando la variable toma valores discretos; por ejemplo, cuando procede de un conteo y sus posibles valores son 0, 1, 2, 3, etc. Cuando el valor del estadístico muestral es una variable continua, la distribución muestral correspondiente se denomina función de densidad de probabilidad. La probabilidad en este caso corresponde gráficamente a un área bajo la curva de esa función, delimitada por un cierto intervalo de la variable. Analíticamente, esa área se calcula como la integral de la función entre los límites del intervalo de la variable, que en la práctica se obtiene con un ordenador o se consulta en una tabla. El área total bajo la curva, que se extiende a todos los posibles valores de la variable, es siempre uno, que corresponde a la probabilidad de un suceso seguro.

La siguiente gráfica muestra la curva de una función de densidad de probabilidad para una variable  $x$ , y en ella se señala la probabilidad de que esa variable se encuentre entre los valores 1 y 2, que corresponde al área bajo la curva marcada en azul:



Fuente: elaboración propia.

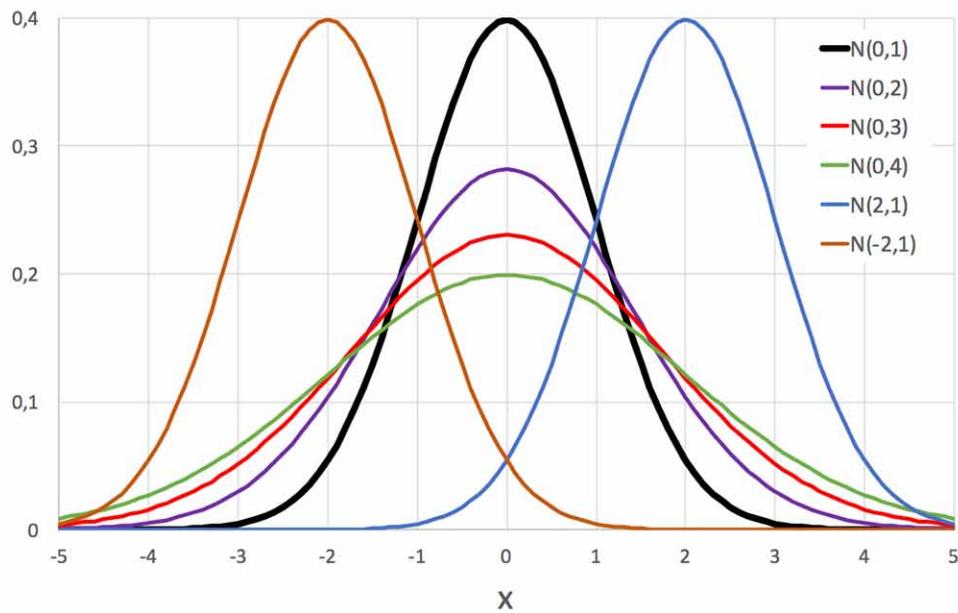
La función de densidad de probabilidad más importante en estadística se llama distribución normal o distribución gaussiana, o también campana de Gauss, por la forma que toma cuando es representada gráficamente (como aparece en la figura anterior). Su forma concreta depende de dos parámetros, la media y la varianza. La curva alcanza un máximo cuando la variable toma el valor de la media, y es simétrica respecto a ese valor, aproximándose a cero indefinidamente conforme la variable se aleja de la media por ambos lados. La desviación típica, que es la raíz cuadrada de la varianza, está relacionada con la anchura de la campana: a mitad de altura del máximo, la anchura de la campana es aproximadamente  $2,36 \sigma$ . Una distribución normal de media  $\mu$  y varianza  $\sigma^2$  se puede simbolizar como  $N(\mu, \sigma^2)$ , y así lo usaremos aquí.

Para una variable aleatoria  $x$  se define su variable tipificada como:

$$z = \frac{x - \mu}{\sigma}$$

Si la variable  $x$  sigue una distribución  $N(\mu, \sigma^2)$ , su variable tipificada correspondiente sigue una distribución  $N(0,1)$ , que se denomina distribución normal estándar, con media 0 y varianza 1 (es la representada en la figura de arriba). Existen infinitas distribuciones normales distintas, tantas como posibles valores de la media y la varianza, pero las áreas bajo la curva normal estándar son las únicas que se pueden encontrar en tablas para su consulta.

La siguiente figura recoge la representación gráfica de distribuciones normales con distintos valores de media y de varianza, según la notación  $N(\mu, \sigma^2)$ ; la curva negra más gruesa es la distribución normal estándar. Se puede observar cómo la posición del máximo de cada curva coincide con su media y cómo la anchura de cada curva es proporcional a su desviación típica; el área total bajo cualquiera de las curvas es siempre uno.



Fuente: elaboración propia.

### EJEMPLO

Se tiene una variable aleatoria  $x$  con distribución normal de media 2 y varianza 9,  $N(2,9)$ . Se quiere calcular la probabilidad de que la variable tome valores entre 1 y 2.

La probabilidad pedida corresponde al área bajo la curva normal especificada y limitada por los valores dados de la variable. Para poder consultar una tabla de áreas bajo la curva normal estándar, se tipifican los valores límite de la variable:

$$z_A = (x_A - \mu) / \sigma = (1 - 2) / \sqrt{9} = -0,333$$

$$z_B = (x_B - \mu) / \sigma = (2 - 2) / \sqrt{9} = 0$$

El área bajo la curva normal  $N(2,9)$  delimitada por los valores de la variable 1 y 2 es la misma que el área bajo la curva normal  $N(0,1)$  delimitada por los valores de la variable  $-0,333$  y 0. Si se busca en una tabla de áreas bajo la curva normal estándar (como se describe en los ejemplos posteriores), se obtiene 0,13. Así, la probabilidad de que una variable aleatoria distribuida según  $N(2,9)$  tome valores entre 1 y 2 es de 0,13, es decir, un 13 %.

A continuación se describen, a modo de ejemplo, las distribuciones muestrales de tres estadísticos habituales en investigaciones cuantitativas: la media aritmética, la diferencia de medias aritméticas de dos grupos y la proporción de un cierto tipo de elementos.

#### Distribución muestral de la media

La media  $m$  de las muestras extraídas al azar de una población con media  $\mu$  y varianza  $\sigma^2$  es una variable aleatoria que sigue una distribución de probabilidad normal caracterizada por:

- Su media, que coincide con la media poblacional  $\mu$ ;
- Su varianza, que viene dada por la varianza de la población  $\sigma^2$  dividida entre el tamaño de la muestra  $n$ , es decir,  $\sigma^2/n$ .

Así pues, la distribución muestral de la media  $m$  es  $N(\mu, \sigma^2/n)$ . La media muestral tipificada sigue una distribución normal estándar,  $N(0,1)$ , y se calcula como:

$$m^* = \frac{m - \mu}{\sigma/\sqrt{n}}$$

Con estos resultados se puede obtener fácilmente la probabilidad de que la media de una muestra extraída al azar se encuentre en un cierto intervalo, conociendo la media y la varianza de la población de la que se ha extraído la muestra. Para ello, se calcula el área bajo la curva de la distribución normal  $N(\mu, \sigma^2/n)$  entre los límites del intervalo, o bien se consulta en una tabla el área bajo la curva de la distribución normal estándar  $N(0,1)$  entre los límites tipificados del intervalo (ver tabla incluida al final de este capítulo).

### EJEMPLO

Si los individuos de una población tienen un peso medio de 70 kg, con una desviación típica de 10 kg, ¿cuál es la probabilidad de que la media de peso de los 121 pasajeros de un avión, que se supone que representan una muestra al azar, esté entre 72 y 73 kg?

Se puede suponer que la media muestral sigue una distribución normal de media 70 kg y varianza  $10^2/121 = 0,826$  kg<sup>2</sup>, es decir,  $N(70, 0,826)$ . Los valores tipificados de los límites del intervalo de la variable son:

$$m^*_A = (72 - 70)/\sqrt{0,826} = 2,20$$

$$m^*_B = (73 - 70)/\sqrt{0,826} = 3,30$$

La variable tipificada sigue una distribución  $N(0,1)$  cuyas áreas bajo la curva se encuentran tabuladas. En la tabla incluida al final de este capítulo se pueden encontrar las áreas bajo la curva normal estándar que quedan a la izquierda de un cierto valor de la variable. El área bajo la curva comprendida entre  $m^*_A$  y  $m^*_B$  es entonces el área a la izquierda de  $m^*_B$  menos el área a la izquierda de  $m^*_A$ .

En esa tabla la cifra de las unidades de la variable se encuentra en la primera fila y las cifras del primer y segundo decimal se encuentran en la primera columna. Una vez localizada la fila y la columna que corresponden al valor buscado, la casilla en la que se cruzan contiene el área bajo la curva normal estándar que queda a la izquierda de ese valor. Se procede entonces de la siguiente manera:

- Para  $m^*_A = 2,20$  se localiza la columna 2,00 y la fila 0,20, en cuya intersección aparece el área 0,986096.
- Para  $m^*_B = 3,30$  se localiza la columna 3,00 y la fila 0,30, en cuya intersección aparece el área 0,999516.

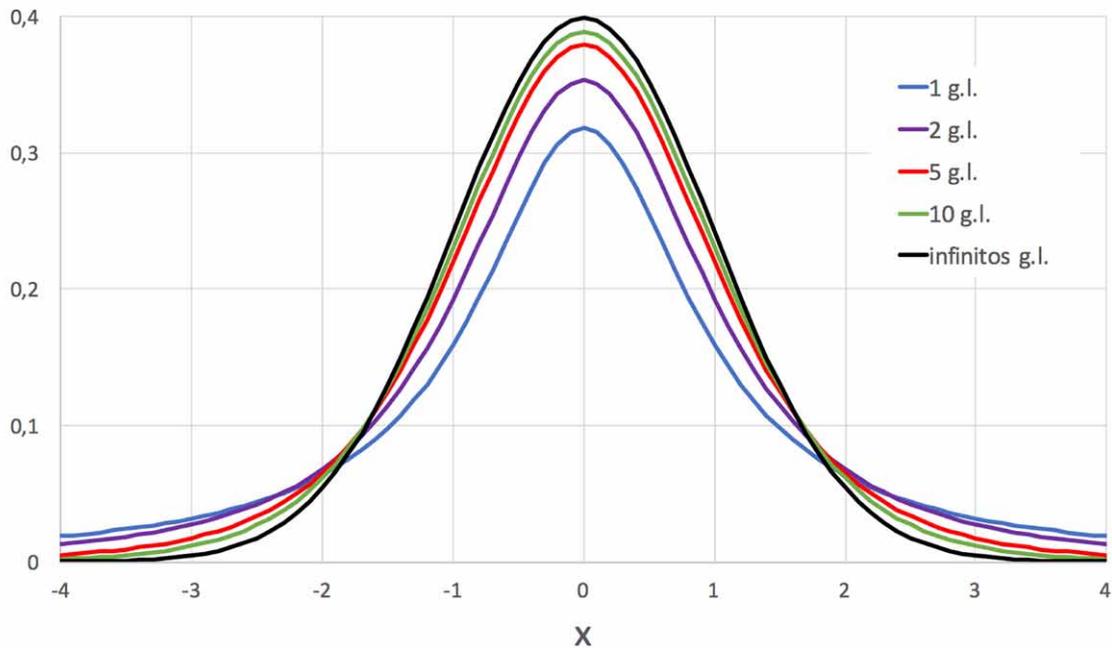
La diferencia entre ambas áreas es 0,01342, que corresponde al área bajo la curva delimitada entre  $m^*_A$  y  $m^*_B$ . Por tanto, la probabilidad de que la media de peso de los pasajeros del avión se encuentre entre 72 y 73 kg es de 0,013 aproximadamente, o de 1,3 %.

Cuando la varianza de la población  $\sigma^2$  no se conoce, que es lo más habitual, ha de estimarse a partir de la varianza de la propia muestra que se ha extraído,  $s^2$ . La media muestral tipificada se obtiene entonces sustituyendo la varianza poblacional por la muestral:

$$m_t^* = \frac{m - \mu}{s/\sqrt{n-1}}$$

Este nuevo estadístico  $t$  contiene la media muestral  $m$  y la desviación típica muestral  $s$ , esta última con su propia distribución muestral. El cociente entre ambas ya no sigue la distribución normal estándar, sino otra distribución denominada  $t$  de Student, que depende de los grados de libertad de la muestra (número de elementos que contiene menos uno,  $n-1$ ). Cuanto mayor es el tamaño de la muestra, más se parece esta distribución a la normal estándar, por lo que a menudo se emplea esta última en muestras grandes incluso si la varianza poblacional es desconocida.

La siguiente gráfica muestra varias distribuciones  $t$  de Student con distintos grados de libertad (g.l.); la curva negra, que corresponde a infinitos grados de libertad, coincide con la distribución normal estándar  $N(0,1)$ :



Fuente: elaboración propia.

### Distribución muestral de la diferencia de medias

También se puede analizar la distribución muestral de la diferencia entre dos medias,  $m_1 - m_2$ , obtenidas de muestras extraídas al azar de dos poblaciones distintas, una de ellas con media  $\mu_1$  y varianza  $\sigma_1^2$  y otra con media  $\mu_2$  y varianza  $\sigma_2^2$ . La distribución muestral de la diferencia de medias sigue una distribución normal caracterizada por:

- Su media, que coincide con la diferencia de las medias poblacionales  $\mu_1 - \mu_2$ .
- Su varianza, que depende de las varianzas de ambas poblaciones y del tamaño de ambas muestras,  $n_1$  y  $n_2$ , y que viene dada por  $\sigma_1^2/n_1 + \sigma_2^2/n_2$ .

Así pues, la distribución muestral de la diferencia de medias  $m_1 - m_2$  es  $N(\mu_1 - \mu_2, \sigma_1^2/n_1 + \sigma_2^2/n_2)$ . La diferencia de medias muestrales tipificada sigue una distribución normal estándar,  $N(0,1)$ , y se calcula como:

$$(m_1 - m_2)^* = \frac{(m_1 - m_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

### EJEMPLO

Se tienen dos poblaciones distintas, una con media de edad de 20 años y desviación típica 1,5 años y la otra con media 25 años y desviación típica 1,7 años. Se reúnen en un aula 50 individuos escogidos al azar de la primera población y en otra aula, 60 individuos de la segunda población. ¿Cuál es la probabilidad de que la diferencia entre las medias de edad de ambas aulas esté entre 4,5 y 5,5 años?

Se puede suponer que la diferencia de medias muestrales sigue una distribución normal de media  $25 - 20 = 5$  años y varianza  $1,5^2/50 + 1,7^2/60 = 0,0932$  años<sup>2</sup>, es decir,  $N(5, 0,0932)$ . Los valores tipificados de los límites del intervalo de la variable son:

$$(m_1 - m_2)^*_{\text{A}} = (4,5 - 5)/\sqrt{0,0932} = -1,64 \text{ años}$$

$$(m_1 - m_2)^*_{\text{B}} = (5,5 - 5)/\sqrt{0,0932} = 1,64 \text{ años}$$

Procediendo de manera análoga al ejemplo anterior, para  $(m_1 - m_2)^*_{\text{B}} = 1,64$  se localiza en la tabla de áreas bajo la curva normal estándar (ver páginas 132-133) la columna 1,00 y la fila 0,64, en cuya intersección aparece el área 0,949497. En la tabla no aparecen valores negativos de la variable, pero como la curva normal estándar es simétrica respecto al valor 0, se deduce que el área a la izquierda de  $(m_1 - m_2)^*_{\text{A}} = -1,64$  es igual al área a la derecha del valor  $+1,64$ ; esta última se puede obtener como el área total bajo la curva, que es 1, menos el área a la izquierda de  $+1,64$ , es decir,  $1 - 0,949497 = 0,050503$ .

Por tanto, el área bajo la curva delimitada por  $(m_1 - m_2)^*_{\text{A}}$  y  $(m_1 - m_2)^*_{\text{B}}$  corresponde a la diferencia  $0,949497 - 0,050503 = 0,898994$ . La probabilidad de que la diferencia entre las medias de edad de las personas de ambas aulas se encuentre entre 4,5 y 5,5 años es entonces de casi 0,9, o del 90 %.

### Distribución muestral de la proporción

La proporción es el número de elementos de un cierto tipo presentes en un conjunto dividido entre el número total de elementos. En una muestra se suele denotar por  $p$ , y la proporción de elementos que no presentan la característica en cuestión es entonces  $1 - p$ . En una población, se suelen denotar  $\pi$  y  $1 - \pi$ , respectivamente.

La proporción  $p$  en muestras grandes ( $n > 30$ ) extraídas al azar de una población con proporción  $\pi$  tiende a seguir una distribución de probabilidad normal, caracterizada por:

- Su media, que coincide con la proporción poblacional  $\pi$ .
- Su varianza, que viene dada por el producto de las proporciones poblacionales de los elementos que presentan y que no presentan la característica en cuestión, dividido entre el tamaño de la muestra  $n$ , es decir,  $\pi(1 - \pi)/n$ .

Así pues, la distribución muestral de la proporción  $p$  es  $N(\pi, \pi(1-\pi)/n)$ . La proporción tipificada sigue una distribución normal estándar,  $N(0,1)$ , y se calcula como:

$$p^* = \frac{p - \pi}{\sqrt{\pi(1 - \pi)/n}}$$

### EJEMPLO

El porcentaje de población española que fuma es del 22 %. En un grupo de 100 personas escogidas al azar, ¿cuál es la probabilidad de que fumen entre 25 y 30 personas?

La proporción de fumadores en la población es 0,22. Se puede suponer entonces que la proporción muestral sigue una distribución normal de media 0,22 y varianza  $0,22 \times (1 - 0,22)/100 = 0,001716$ , es decir,  $N(0,22, 0,001716)$ . Los límites del intervalo de la variable, expresados en proporciones, son 0,25 y 0,30, cuyos valores tipificados son:

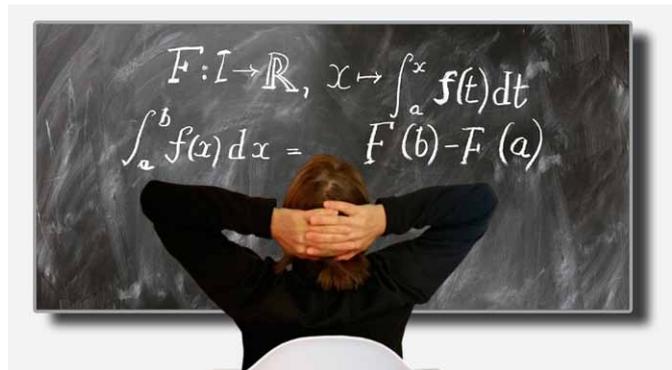
$$p^*_A = (0,25 - 0,22) / \sqrt{0,001716} = 0,72$$

$$p^*_B = (0,30 - 0,22) / \sqrt{0,001716} = 1,93$$

Procediendo de manera análoga a los ejemplos anteriores, para  $p^*_A = 0,72$  se localiza en la tabla de áreas bajo la curva normal estándar (ver páginas 132-133) la columna 0,00 y la fila 0,72, en cuya intersección aparece el área 0,764237, y para  $p^*_B = 1,93$  se localiza en la tabla la columna 1,00 y la fila 0,93, en cuya intersección aparece el área 0,973196. El área bajo la curva delimitada entre  $p^*_A$  y  $p^*_B$  corresponde entonces a la diferencia entre ambas, 0,208959. Por tanto, la probabilidad de que en esa muestra fume entre un 25 % y un 30 % de las personas es de casi 0,21, o del 21 %.

### AMPLIACIÓN. Probabilidad en variables continuas

El cálculo de probabilidades como áreas bajo curvas implica que la probabilidad de que una variable continua tome un cierto valor exacto es nula. Por ejemplo, la probabilidad de que una variable tome el valor exacto 1,52 sería el área bajo su función de densidad de probabilidad entre los valores 1,52 y 1,52, que es cero (esos límites, al ser iguales, definen una línea, no una superficie). Dicho de otro modo, la integral de la función de densidad en el intervalo que va entre  $a = 1,52$  y  $b = 1,52$  es nula.



Licencia: dominio público. Fuente: Pixabay.

En realidad, ninguna medida proporciona valores exactos de una variable, como sería el caso de 1,52000000... con infinitos ceros; se sobrentiende que se trata siempre de intervalos, en este caso desde 1,515 hasta 1,525, pero esto no se explicita porque ambos valores se redondean a 1,52. Así, la probabilidad de obtener el valor 1,52, correctamente interpretado, se calcula como el área bajo la curva (o integral) de su función de densidad de probabilidad entre los valores  $a = 1,515$  y  $b = 1,525$ . Si tuviéramos un valor más preciso de la variable, como 1,52689, se sobrentendería que el intervalo que habría que usar iría entre  $a = 1,526885$  y  $b = 1,526895$ .

El área total bajo una función de densidad de probabilidad corresponde a la probabilidad de que ocurra un valor cualquiera de la variable, que es un suceso seguro, ya que algún valor tiene que ocurrir. Por tanto, el área total coincide con la probabilidad del suceso seguro, que es uno.

### AMPLIACIÓN. El teorema central del límite

La distribución gaussiana toma su nombre del matemático y físico alemán Carl Friedrich Gauss (1777-1855). Su expresión matemática, que viene caracterizada, como ya sabemos, por la media  $\mu$  (que da la posición del máximo de la distribución) y la desviación típica  $\sigma$  (relacionada con la anchura de la distribución), es la siguiente:

$$N(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

El hecho de que algunos estadísticos muestrales sigan esta distribución es absolutamente esencial en estadística, y por extensión, en el método científico, y su origen se encuentra en un teorema matemático que se conoce como teorema central del límite, desarrollado por el matemático y físico francés Pierre Simon de Laplace (1749-1827).



Carl Friedrich Gauss. Licencia: dominio público.  
Fuente: Wikimedia Commons.



Pierre Simon de Laplace. Licencia: dominio público.  
Fuente: Wikimedia Commons.

Supongamos una muestra de valores numéricos  $(x_1, x_2, \dots, x_n)$  extraídos aleatoriamente de una población caracterizada por una media  $\mu$  y una varianza  $\sigma^2$ . Cada uno de esos valores es una variable aleatoria, cuya media y cuya varianza son las de la población de la que ha sido extraído. Su suma,  $X = x_1 + x_2 + \dots + x_n$ ,

es también una variable aleatoria, y se puede demostrar que tiende a seguir una distribución normal cuya media es la suma de las medias de las  $n$  variables,  $n\mu$ , y cuya varianza es la suma de las varianzas de las  $n$  variables (independientes entre sí),  $n\sigma^2$ . Cuanto mayor es el número de variables aleatorias sumadas, es decir, al aumentar el tamaño de la muestra  $n$ , más se aproxima la distribución de la variable suma  $X$  a la distribución normal  $N(n\mu, n\sigma^2)$ . En la práctica, se suele considerar que muestras con  $n > 30$  son suficientemente grandes como para que la distribución de la suma sea normal con muy buena aproximación.

Si la suma de los elementos de la muestra sigue una distribución normal, entonces la media aritmética también ha de seguirla, ya que no es más que esa misma suma dividida entre el número de elementos,  $X/n$ . La media de esta nueva variable (es decir, la media de las medias muestrales) resulta  $n\mu/n = \mu$ , y su varianza (es decir, la varianza de las medias muestrales) resulta  $n\sigma^2/n^2 = \sigma^2/n$ . Así, el teorema central del límite establece que la distribución muestral de la media tiende a una distribución normal  $N(\mu, \sigma^2/n)$ , que es el resultado que hemos venido usando.

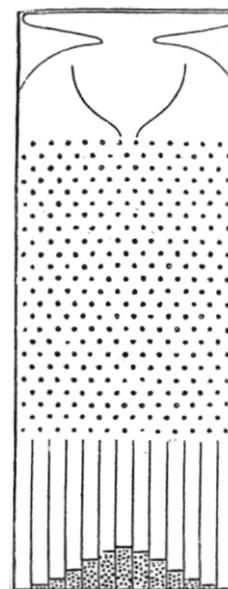
Este teorema es válido cuando las variables aleatorias pertenecen a una población muy general, que solo necesita estar caracterizada por una media y por una varianza, pero no tienen por qué seguir una distribución normal en esa población. Si lo hacen, la distribución muestral de la media es normal de manera exacta, incluso para muestras pequeñas. En cambio, si las variables no siguen una distribución normal en la población, la distribución muestral de la media es solo aproximadamente normal, pareciéndose más a ella conforme aumenta el tamaño de la muestra (y siendo casi igual para  $n > 30$ ).

Cuando la varianza poblacional no se conoce y se estima a través de la varianza muestral, la media tipificada sigue la distribución  $t$  de Student (la media en sí sigue tendiendo a una distribución normal, es su tipificación usando  $s^2$  la que sigue esta otra distribución). Esto solo es estrictamente cierto cuando las variables de la muestra siguen una distribución normal en la población. Si no la siguen, la distribución muestral de la media tipificada tiende a la  $t$  de Student solo en muestras grandes, pero no en muestras pequeñas, aunque tampoco es mala aproximación en este caso. Dado que la distribución  $t$  de Student tiende a la distribución normal conforme aumentan sus grados de libertad, para muestras grandes ( $n > 30$ ) es habitual usar la distribución normal incluso si la varianza poblacional es desconocida y se usa la de la muestra.

### HISTORIAS DE LA CIENCIA. La máquina de Galton

Francis Galton fue un estadístico inglés (1822-1911), primo de Charles Darwin. Diseñó un modelo mecánico para explorar y demostrar que la suma de variables aleatorias da lugar a una distribución normal, en relación con el teorema central del límite, uno de los más importantes en estadística inferencial y por tanto en la ciencia.

El modelo consiste en una tabla dispuesta verticalmente, por la que se dejan caer bolitas desde su parte central superior. En su caída, las bolitas van chocando con obstáculos iguales y equiespaciados, pero desplazados unos con respecto a otros en filas alternas, como se observa en el diseño del propio Galton mostrado en la figura. Cada vez que una bolita choca con un obstáculo solo hay dos opciones, desviarse hacia la izquierda o hacia la derecha. Idealmente, ambas presentan la misma probabilidad y el resultado es aleatorio. Al llegar a la parte de abajo de la máquina, tras numerosos choques aleatorios, las bolitas se acumulan en casilleros distribuidos horizontalmente. La altura que alcanzan las bolitas acumuladas en cada casillero está asociada a la probabilidad de acabar en él tras la sucesión (suma) de choques aleatorios. La distribución que se forma, como se ve en la figura, tiende a la curva normal o gaussiana, con mayor acumulación cerca del centro y disminuyendo en forma de campana de Gauss conforme los casilleros se alejan del centro.



Diseño original del modelo mecánico de Galton. Licencia: dominio público. Fuente: Wikimedia Commons.

Esta máquina, llamada de Galton, de las judías, o quincunx, es muy conocida porque ha servido como base a diversas máquinas recreativas y juegos de azar. En Japón, por ejemplo, son muy famosas las máquinas de pachinko, similares a la de Galton, con las que se realizan apuestas acerca de dónde caerá cada bola o si será atrapada en ciertos receptáculos. El mercado asociado al pachinko supera con creces todo el negocio de apuestas de Las Vegas.

#### HISTORIAS DE LA CIENCIA. El estudiante cervecero

William Sealy Gosset (1876-1937) fue un químico y matemático inglés que trabajó para la famosa compañía cervecera Guinness en Dublín a principios del siglo xx. Llevó a cabo una investigación sobre el efecto de diversas condiciones del suelo y del tipo de cebada sobre el rendimiento de los cultivos y la calidad de la cerveza, para lo que contaba con muestras pequeñas, de alrededor de diez datos. Además, la población de la que se extraían las muestras no se conocía demasiado bien, por lo que su varianza se estimaba a partir de la varianza de las muestras. Gosset se dio cuenta de que las medias de sus muestras no seguían una distribución de probabilidad normal, y desarrolló una distribución más adecuada. Como los datos con los que realizó las pruebas eran confidenciales y no debían caer en manos de la competencia, evitó que pudieran relacionarse con su compañía publicándolos con el seudónimo Student (estudiante). La variable que seguía su nueva distribución se comenzó a denotar como  $t$ , y con el tiempo tomó el nombre de  $t$  de Student.

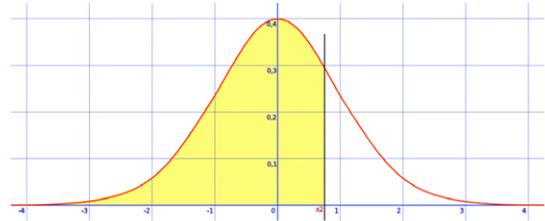


William Sealy Gosset. Licencia: dominio público.  
Fuente: Wikimedia Commons.

Tras ese primer trabajo, Gosset continuó estudiando las aplicaciones y las propiedades de su nueva distribución, analizando 750 muestras de cuatro elementos cada una extraídas aleatoriamente. Esta vez dejó a un lado los datos agrícolas y optó por una población un poco más extraña: la longitud de los dedos y la estatura de un conjunto de criminales encarcelados.

### Tabla de áreas bajo la curva normal estándar

Esta tabla contiene el área bajo la curva normal estándar situada a la izquierda de un cierto valor de la variable (región amarilla en la figura). La cifra de las unidades de la variable se encuentra en la primera fila y las cifras del primer y segundo decimal se encuentran en la primera columna. Una vez localizada la fila y la columna que corresponden al valor buscado de la variable, la casilla en la que se cruzan contiene el área bajo la curva normal estándar que queda a la izquierda de ese valor.



Licencia: dominio público. Fuente: Wikilibros.

Zp	0,00	1,00	2,00	3,00	4,00
0,01	0, 503 989	0, 843 752	0, 977 784	0, 998 693	0, 999 969
0,02	0, 507 978	0, 846 135	0, 978 308	0, 998 736	0, 999 970
0,03	0, 511 966	0, 848 494	0, 978 821	0, 998 777	0, 999 972
0,04	0, 515 953	0, 850 830	0, 979 324	0, 998 817	0, 999 973
0,05	0, 519 938	0, 853 140	0, 979 817	0, 998 855	0, 999 974
0,06	0, 523 922	0, 855 427	0, 980 300	0, 998 893	0, 999 975
0,07	0, 527 903	0, 857 690	0, 980 773	0, 998 929	0, 999 976
0,08	0, 531 881	0, 859 928	0, 981 237	0, 998 964	0, 999 977
0,09	0, 535 856	0, 862 143	0, 981 691	0, 998 999	0, 999 978
0,10	0, 539 827	0, 864 333	0, 982 135	0, 999 032	0, 999 979
0,11	0, 543 795	0, 866 500	0, 982 570	0, 999 064	0, 999 980
0,12	0, 547 758	0, 868 643	0, 982 997	0, 999 095	0, 999 981
0,13	0, 551 716	0, 870 761	0, 983 414	0, 999 125	0, 999 981
0,14	0, 555 670	0, 872 856	0, 983 822	0, 999 155	0, 999 982
0,15	0, 559 617	0, 874 928	0, 984 222	0, 999 183	0, 999 983
0,16	0, 563 559	0, 876 975	0, 984 613	0, 999 211	0, 999 984
0,17	0, 567 494	0, 878 999	0, 984 996	0, 999 237	0, 999 984
0,18	0, 571 423	0, 880 999	0, 985 371	0, 999 263	0, 999 985
0,19	0, 575 345	0, 882 976	0, 985 737	0, 999 288	0, 999 986
0,20	0, 579 259	0, 884 930	0, 986 096	0, 999 312	0, 999 986
0,21	0, 583 166	0, 886 860	0, 986 447	0, 999 336	0, 999 987
0,22	0, 587 064	0, 888 767	0, 986 790	0, 999 358	0, 999 987
0,23	0, 590 954	0, 890 651	0, 987 126	0, 999 380	0, 999 988
0,24	0, 594 834	0, 892 512	0, 987 454	0, 999 402	0, 999 988
0,25	0, 598 706	0, 894 350	0, 987 775	0, 999 422	0, 999 989
0,26	0, 602 568	0, 896 165	0, 988 089	0, 999 442	0, 999 989
0,27	0, 606 419	0, 897 957	0, 988 396	0, 999 462	0, 999 990
0,28	0, 610 261	0, 899 727	0, 988 696	0, 999 480	0, 999 990
0,29	0, 614 091	0, 901 474	0, 988 989	0, 999 499	0, 999 991

Zp	0,00	1,00	2,00	3,00	4,00
0,30	0, 617 911	0, 903 199	0, 989 275	0, 999 516	0, 999 991
0,31	0, 621 719	0, 904 902	0, 989 555	0, 999 533	0, 999 991
0,32	0, 625 515	0, 906 582	0, 989 829	0, 999 549	0, 999 992
0,33	0, 629 299	0, 908 240	0, 990 096	0, 999 565	0, 999 992
0,34	0, 633 071	0, 909 877	0, 990 358	0, 999 581	0, 999 992
0,35	0, 636 830	0, 911 491	0, 990 613	0, 999 595	0, 999 993
0,36	0, 640 576	0, 913 084	0, 990 862	0, 999 610	0, 999 993
0,37	0, 644 308	0, 914 656	0, 991 105	0, 999 624	0, 999 993
0,38	0, 648 027	0, 916 206	0, 991 343	0, 999 637	0, 999 994
0,39	0, 651 731	0, 917 735	0, 991 575	0, 999 650	0, 999 994
0,40	0, 655 421	0, 919 243	0, 991 802	0, 999 663	0, 999 994
0,41	0, 659 096	0, 920 730	0, 992 023	0, 999 675	0, 999 994
0,42	0, 662 757	0, 922 196	0, 992 239	0, 999 686	0, 999 995
0,43	0, 666 402	0, 923 641	0, 992 450	0, 999 698	0, 999 995
0,44	0, 670 031	0, 925 066	0, 992 656	0, 999 709	0, 999 995
0,45	0, 673 644	0, 926 470	0, 992 857	0, 999 719	0, 999 995
0,46	0, 677 241	0, 927 854	0, 993 053	0, 999 729	0, 999 995
0,47	0, 680 822	0, 929 219	0, 993 244	0, 999 739	0, 999 996
0,48	0, 684 386	0, 930 563	0, 993 430	0, 999 749	0, 999 996
0,49	0, 687 933	0, 931 887	0, 993 612	0, 999 758	0, 999 996
0,50	0, 691 462	0, 933 192	0, 993 790	0, 999 767	0, 999 996
0,51	0, 694 974	0, 934 478	0, 993 963	0, 999 775	0, 999 996
0,52	0, 698 468	0, 935 744	0, 994 132	0, 999 784	0, 999 996
0,53	0, 701 944	0, 936 991	0, 994 296	0, 999 792	0, 999 997
0,54	0, 705 401	0, 938 219	0, 994 457	0, 999 799	0, 999 997
0,55	0, 708 840	0, 939 429	0, 994 613	0, 999 807	0, 999 997
0,56	0, 712 260	0, 940 620	0, 994 766	0, 999 814	0, 999 997
0,57	0, 715 661	0, 941 792	0, 994 915	0, 999 821	0, 999 997
0,58	0, 719 042	0, 942 946	0, 995 059	0, 999 828	0, 999 997
0,59	0, 722 404	0, 944 082	0, 995 201	0, 999 834	0, 999 997

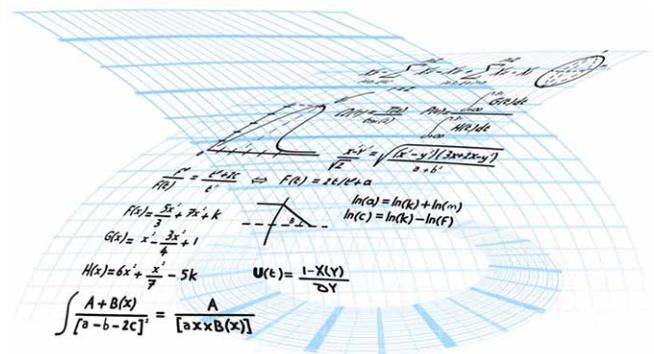
Zp	0,00	1,00	2,00	3,00	4,00
0,60	0,725 746	0,945 200	0,995 338	0,999 840	0,999 997
0,61	0,729 069	0,946 301	0,995 472	0,999 846	0,999 997
0,62	0,732 371	0,947 383	0,995 603	0,999 852	0,999 998
0,63	0,735 652	0,948 449	0,995 730	0,999 858	0,999 998
0,64	0,738 913	0,949 497	0,995 854	0,999 863	0,999 998
0,65	0,742 153	0,950 528	0,995 975	0,999 868	0,999 998
0,66	0,745 373	0,951 542	0,996 092	0,999 873	0,999 998
0,67	0,748 571	0,952 540	0,996 207	0,999 878	0,999 998
0,68	0,751 747	0,953 521	0,996 318	0,999 883	0,999 998
0,69	0,754 902	0,954 486	0,996 427	0,999 887	0,999 998
0,70	0,758 036	0,955 434	0,996 532	0,999 892	0,999 998
0,71	0,761 148	0,956 367	0,996 635	0,999 896	0,999 998
0,72	0,764 237	0,957 283	0,996 735	0,999 900	0,999 998
0,73	0,767 304	0,958 184	0,996 833	0,999 904	0,999 998
0,74	0,770 350	0,959 070	0,996 927	0,999 907	0,999 998
0,75	0,773 372	0,959 940	0,997 020	0,999 911	0,999 998
0,76	0,776 372	0,960 796	0,997 109	0,999 915	0,999 999
0,77	0,779 350	0,961 636	0,997 197	0,999 918	0,999 999
0,78	0,782 304	0,962 462	0,997 281	0,999 921	0,999 999
0,79	0,785 236	0,963 273	0,997 364	0,999 924	0,999 999
0,80	0,788 144	0,964 069	0,997 444	0,999 927	0,999 999
0,81	0,791 029	0,964 852	0,997 522	0,999 930	0,999 999
0,82	0,793 892	0,965 620	0,997 598	0,999 933	0,999 999
0,83	0,796 730	0,966 375	0,997 672	0,999 935	0,999 999
0,84	0,799 545	0,967 115	0,997 744	0,999 938	0,999 999
0,85	0,802 337	0,967 843	0,997 813	0,999 940	0,999 999
0,86	0,805 105	0,968 557	0,997 881	0,999 943	0,999 999
0,87	0,807 849	0,969 258	0,997 947	0,999 945	0,999 999
0,88	0,810 570	0,969 946	0,998 011	0,999 947	0,999 999
0,89	0,813 267	0,970 621	0,998 073	0,999 949	0,999 999

Zp	0,00	1,00	2,00	3,00	4,00
0,90	0,815 939	0,971 283	0,998 134	0,999 951	0,999 999
0,91	0,818 588	0,971 933	0,998 192	0,999 953	0,999 999
0,92	0,821 213	0,972 571	0,998 249	0,999 955	0,999 999
0,93	0,823 814	0,973 196	0,998 305	0,999 957	0,999 999
0,94	0,826 391	0,973 810	0,998 358	0,999 959	0,999 999
0,95	0,828 943	0,974 412	0,998 411	0,999 960	0,999 999
0,96	0,831 472	0,975 002	0,998 461	0,999 962	0,999 999
0,97	0,833 976	0,975 580	0,998 510	0,999 964	0,999 999
0,98	0,836 456	0,976 148	0,998 558	0,999 965	0,999 999
0,99	0,838 912	0,976 704	0,998 605	0,999 966	0,999 999

# Capítulo 15.

## Estadística inferencial: estimación de parámetros y contraste de hipótesis

El objetivo de la estadística inferencial es obtener información de una población a partir de los datos recogidos en una muestra extraída al azar; específicamente, se ocupa de estimar parámetros poblacionales a partir de estadísticos muestrales. En el capítulo anterior se introdujo el fundamento matemático que permite llevarlo a cabo, basado en las distribuciones muestrales de los estadísticos. A partir de ellas, la estadística inferencial logra cuantificar el efecto del azar asociado a la extracción de la muestra, proporcionando la probabilidad de que un parámetro poblacional se encuentre en un determinado rango de valores en torno al del estadístico obtenido en la muestra. Este es el primer procedimiento de la estadística inferencial que se tratará en este capítulo, el de la estimación de parámetros poblacionales. El segundo procedimiento, que está muy relacionado con el anterior, es el contraste de hipótesis acerca de parámetros poblacionales, que también parte de los datos recogidos en muestras extraídas al azar. Este último procedimiento suele tener un fuerte carácter dicotómico y por tanto un importante papel en la toma de decisiones: tras llevar a cabo el contraste, la hipótesis formulada sobre la población se rechaza o no se rechaza. Sin embargo, esto no tiene por qué ser siempre así, y de hecho la ciencia moderna tiende cada vez más a analizar y discutir las probabilidades calculadas en los contrastes de hipótesis (valor  $P$ ), más que centrarse en el simple rechazo o no rechazo de la hipótesis.



Licencia: dominio público. Fuente: Pixabay.

### 15.1. Estimación de parámetros

El valor de un parámetro poblacional suele ser desconocido, pero puede estimarse a partir de un estadístico obtenido en una muestra extraída al azar de la población; ese estadístico toma el nombre de estimador puntual del parámetro. El estimador puntual es insesgado cuando la media de sus valores en todas las posibles muestras coincide con el valor del parámetro poblacional. Por ejemplo, la media, la diferencia de medias o la proporción en una muestra son estimadores puntuales insesgados de la media, la diferencia de medias o la proporción en la población, respectivamente.

Dado que el valor del estimador puntual varía de una muestra a otra, usarlo para inferir un parámetro poblacional no es suficiente por sí solo. La estadística inferencial permite establecer un intervalo de confianza, que es el rango de valores en el que se encontrará el parámetro poblacional con una cierta probabilidad. Esa probabilidad, que el investigador elige libremente, recibe el nombre de nivel de confianza y se simboliza  $(1-\alpha)$ , donde la probabilidad  $\alpha$  recibe el nombre de nivel de significación. Ambos corresponden a áreas bajo la curva de una función de densidad de probabilidad, o en términos analíticos, a integrales definidas sobre ella. Esa función de densidad debe representar cómo se distribuye en las muestras el estimador puntual, es decir, se trata de su distribución muestral. Como vimos antes, la media, la diferencia de medias y la proporción siguen una distribución muestral normal, bajo ciertas condiciones.

El intervalo de confianza se extiende por encima y por debajo del valor del estimador puntual una cantidad dada por el error máximo,  $\varepsilon$ , llamado así porque expresa la mayor diferencia que puede haber entre el parámetro poblacional real y el estimado, para un nivel de confianza dado. Cuanto mayor es la confianza que se quiere tener en que el parámetro poblacional se encontrará en un intervalo dado, más amplio tendrá que ser este, es decir, mayor tendrá que ser el error máximo. Si quisiéramos alcanzar una confianza del 100 %, el error máximo sería infinito, lo cual no tiene ninguna utilidad: evidentemente, existe absoluta certeza de que la media poblacional tomará algún valor entre el infinito negativo y el infinito positivo. Lo que se hace es seleccionar una probabilidad alta, por ejemplo 95 %, es decir,  $(1-\alpha) = 0,95$  y  $\alpha = 0,05$ , que se logra con un error máximo de tamaño razonable (desde luego, no infinito). Otros niveles de confianza habituales son 99 %, 99,5 % o 99,9 %, y pueden acercarse aún más a 100 %, pero nunca alcanzarlo.

Para calcular el error máximo, y por tanto la amplitud del intervalo de confianza, se recurre a un estadístico de inferencia, que suele ser el estimador puntual tipificado. Este a menudo sigue una distribución muestral normal estándar, con media 0 y desviación típica 1 (aunque ciertos estadísticos siguen distribuciones distintas). En ese caso, los límites del intervalo de confianza se encuentran a la misma distancia del cero, uno de ellos negativo y el otro positivo; se denotan por  $-z_{\alpha/2}$  y  $+z_{\alpha/2}$ , respectivamente, donde la  $z$  hace referencia a la variable de la curva normal estándar y  $\alpha$  es el nivel de significación escogido, y se pueden obtener con un ordenador o consultando una tabla de áreas bajo la curva. A veces se fija primero el valor de  $z_{\alpha/2}$  en lugar de  $\alpha$ ; por ejemplo,  $z_{\alpha/2} = 1$  (que corresponde a confianza 68,5 % y se denomina  $1\sigma$  o «un sigma»),  $z_{\alpha/2} = 2$  (confianza 95,5 %,  $2\sigma$  o «dos sigmas»),  $z_{\alpha/2} = 3$  (confianza 99,75 %,  $3\sigma$  o «tres sigmas»), etc.

El error máximo también es proporcional a la desviación típica de la distribución muestral, y esta es inversamente proporcional a la raíz cuadrada del tamaño de la muestra,  $n$ . Así, para un nivel de confianza dado ( $z_{\alpha/2}$  fija), cuanto mayor sea el tamaño de la muestra, más pequeña será la amplitud del intervalo de confianza, es decir, mejor (más precisa) será la estimación del parámetro. Si inicialmente se fijan tanto el nivel de confianza como la amplitud deseada del intervalo de confianza, entonces se puede deducir a partir de ellos el tamaño que debe tener la muestra con la que se trabaje. Este es un dato muy relevante en una investigación, ya que influye notablemente en los recursos que hay que dedicar para llevarla a cabo.

#### Intervalo de confianza para la media poblacional

La media tipificada de una muestra extraída al azar de una población sigue una distribución normal estándar. Siguiendo el razonamiento explicado antes se deduce que la probabilidad de que esa media

tipificada se encuentre entre los valores  $-z_{\alpha/2}$  y  $+z_{\alpha/2}$  tiene que ser igual a  $(1-\alpha)$ , lo que matemáticamente se expresa como:

$$\text{prob} \left( -z_{\alpha/2} < \frac{m - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2} \right) = 1 - \alpha$$

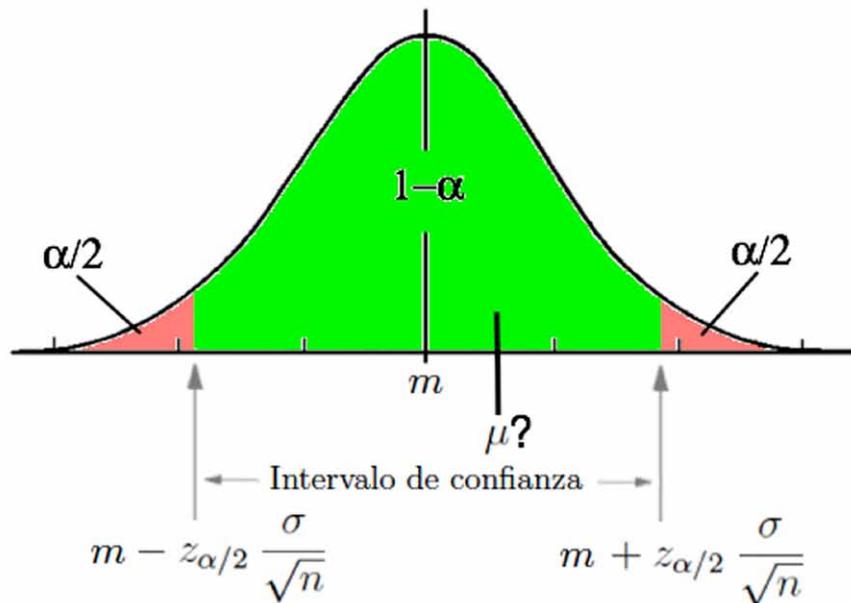
En la fórmula anterior se puede despejar el valor de la media poblacional  $\mu$ , resultando un intervalo para esta última:

$$\text{prob} \left( m - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < m + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right) = 1 - \alpha$$

Esta expresión significa que la media de la población,  $\mu$ , se esperaría encontrar con una probabilidad o confianza  $(1-\alpha)$  en el intervalo comprendido entre los límites:

$$\mu_{lim} = m \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

El intervalo de confianza resultante tiene su punto medio en  $m$  y se extiende por encima y por debajo de ese valor una longitud dada por el término que acompaña al símbolo  $\pm$ , que es el error máximo,  $\epsilon$ ; la amplitud total del intervalo es de dos veces el error máximo. El intervalo de confianza sobre la curva normal se puede visualizar en la siguiente figura<sup>7</sup>:



Fuente: elaboración propia.

<sup>7</sup> Enlace relacionado: <http://view.genial.ly/5d6baf6a961a560fa9cacb0c/interactive-content-estimacion-parametros>

Para un nivel de confianza  $(1-\alpha)$  previamente fijado y un error máximo  $\epsilon$  aceptable también fijado, el tamaño de la muestra se puede despejar de las expresiones anteriores, resultando:

$$n = \left( \frac{z_{\alpha/2} \sigma}{\epsilon} \right)^2$$

### EJEMPLO

Se lleva a cabo una investigación sobre el gasto medio mensual por español en alimentación, que sigue una distribución normal de desviación típica  $\sigma = 20$  euros. Se extrae una muestra al azar de 50 personas, cuyo gasto medio resulta ser de 290 euros. ¿En qué intervalo se encuentra la media en España para un nivel de confianza del 95 %? Si se quisiera reducir la amplitud del intervalo a 6 euros, ¿cuál debería ser el tamaño de la muestra estudiada?

El nivel de confianza es  $(1-\alpha) = 0,95$  y el de significación es  $\alpha = 0,05$ . El valor  $z_{\alpha/2} = z_{0,025}$  es el que deja a su derecha un área bajo la curva normal estándar de 0,025, y por tanto a su izquierda un área de 0,975. Buscando esta última área (o la más cercana a ella) en la tabla de áreas bajo la curva normal estándar (ver páginas 132-133), se encuentra que está situada en la columna 1,00 y en la fila 0,96, es decir, corresponde al valor  $z_{0,025} = 1,96$ . El error máximo es  $\epsilon = 1,96 \times 20/\sqrt{50} = 5,54$  euros, y por tanto los límites del intervalo de confianza para la media poblacional son  $290 \pm 5,54$  euros. Es decir, el gasto medio mensual en alimentación de la población española se encontrará entre 284,46 y 295,54 euros con una probabilidad del 95 %.

El intervalo obtenido para la media tiene una amplitud de dos veces el error máximo, es decir, 11,08 euros. Si se quisiera reducir a 6 euros, que corresponde a un error máximo de 3 euros, habría que estudiar una muestra de tamaño  $n = (1,96 \times 20 / 3)^2 = 171$  personas aproximadamente, en lugar de 50.

Cuando la varianza en la población no se conoce y se usa en su lugar la varianza en la muestra, el estadístico de inferencia de la media sigue una distribución muestral  $t$  de Student bajo ciertas condiciones. En ese caso, el intervalo de confianza para la media viene definido por los límites:

$$\mu_{lim} = m \pm t_{\alpha/2, n-1} \frac{s}{\sqrt{n-1}}$$

Es decir, los valores límite se refieren ahora a áreas bajo la distribución  $t$  de Student, no la normal, que se pueden consultar en otras tablas. No obstante, para muestras grandes ( $n > 30$ ) se puede seguir usando la distribución normal porque es muy buena aproximación.

### Intervalo de confianza para la diferencia de medias de dos poblaciones

De manera análoga a como se ha obtenido el intervalo de confianza para la media de una población, se puede establecer un intervalo de confianza para la diferencia de medias de dos poblaciones distintas, a partir de la diferencia de las medias de dos muestras extraídas al azar de esas poblaciones. Recordando la tipificación de la diferencia de medias vista en el capítulo anterior, y que su distribución muestral es la

normal estándar, un razonamiento análogo al de la media permite deducir que los límites del intervalo de confianza vienen dados en este caso por:

$$(\mu_1 - \mu_2)_{lim} = (m_1 - m_2) \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

### EJEMPLO

Se mide 40 veces el tiempo necesario para realizar un cierto trayecto en autobús, obteniendo una media de 46 minutos con desviación típica 8 minutos. Para el mismo trayecto se mide 30 veces el tiempo necesario para realizarlo en coche, y se obtiene una media de 36 minutos con desviación típica 6 minutos. Para un nivel de confianza del 95 %, obtener el intervalo en el que se encuentra la diferencia entre las medias de los tiempos de viaje en autobús y en coche.

Como en el ejemplo anterior, se tiene  $z_{\alpha/2} = z_{0,025} = 1,96$ . El error máximo es  $\varepsilon = 1,96 \times \sqrt{(8^2/40 + 6^2/30)} = 3,28$  minutos, y por tanto los límites del intervalo de confianza para la diferencia de medias en la población son  $10 \pm 3,28$  minutos. Es decir, se espera que el 95 % de las veces se tarde en hacer ese trayecto entre 6,72 y 13,28 minutos más en autobús que en coche.

### Intervalo de confianza para la proporción poblacional

De manera análoga se puede obtener el intervalo de confianza para la proporción en una población a partir de su valor en una muestra extraída al azar. Como la proporción poblacional no se conoce, ya que de hecho estamos tratando de estimarla, la varianza debe calcularse usando la proporción en la muestra, como  $p(1-p)/n$ . Empleando esa varianza para la tipificación de la proporción vista en el capítulo anterior, que sigue una distribución muestral normal estándar (para muestras grandes,  $n > 30$ ), los límites del intervalo de confianza vienen dados por:

$$\pi_{lim} = p \pm z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}$$

### EJEMPLO

Se entrevista a un grupo de 800 personas escogidas al azar, de las que 120 resultan encontrarse en situación de desempleo. Para un nivel de confianza del 99 %, ¿en qué intervalo se sitúa el porcentaje de desempleo en la población?

La proporción de parados en la muestra es  $p = 120/800 = 0,15$  (un 15 %). El nivel de confianza pedido es  $(1-\alpha) = 0,99$  y el de significación es  $\alpha = 0,01$ . Dado que el tamaño de la muestra es grande, se puede suponer una distribución normal estándar para la proporción muestral tipificada. El valor  $z_{\alpha/2} = z_{0,005}$  es el que deja a su derecha un área bajo la curva normal estándar de 0,005, y por tanto a su izquierda un área de 0,995. Buscando esta última área (o la más cercana a ella) en la tabla de áreas bajo la curva normal estándar (ver páginas 132-133), se encuentra que está situada en la columna 2,00 y en la fila 0,58, es decir, corresponde al valor  $z_{0,005} = 2,58$ .

El error máximo se obtiene como  $\varepsilon = 2,58 \times \sqrt{(0,15 \times (1-0,15) / 800)} = 0,033$ , y los límites del intervalo de confianza para la proporción poblacional resultantes son  $0,150 \pm 0,033$ . Es decir, el porcentaje de parados en la población se encontrará entre 11,7 % y 18,3 % con una probabilidad del 99 %.

### AMPLIACIÓN. Tamaño muestral

El tamaño de la muestra depende del nivel de confianza que se quiera alcanzar y del error máximo aceptable para el parámetro poblacional, es decir, del tamaño que se imponga al intervalo de confianza, pero no del tamaño de la población. Así, por ejemplo, en una investigación sobre intención de voto con confianza y error fijados *a priori* puede ser necesario encuestar a mil personas, resultado que será el mismo si el censo es de cien mil, de diez millones o de cien millones de personas. Esto puede resultar poco intuitivo, porque para poblaciones muy grandes esa muestra representa una proporción muy pequeña, que puede parecer poco representativa. Lo cierto es que da prácticamente igual qué proporción represente, lo importante es que la muestra se haya obtenido al azar. Se puede establecer una analogía con la medida del punto de sal en un caldo: bastará removerlo bien y probar una cucharadita, cuyo contenido será suficientemente representativo, independientemente de que el recipiente contenga caldo para una persona o para un regimiento; en este último caso, el cocinero no necesita beberse unos cuantos litros para comprobar el punto de sal.

En realidad, en la varianza de las distribuciones muestrales de los estimadores se puede introducir un factor de corrección para poblaciones finitas dado por  $1-n/N$ , que depende del número de elementos de la población,  $N$ . Pero este factor apenas modifica la varianza cuando el tamaño de la población es mucho mayor que el tamaño de la muestra, como suele ser lo habitual, porque se aproxima mucho a 1. Si ocurriera la situación opuesta, en la que el tamaño de la muestra fuese próximo al de la población, el factor de corrección prácticamente se hace cero y anula la varianza, lo cual resulta lógico: si la muestra es prácticamente toda la población, la distribución muestral del estadístico apenas tiene variabilidad, porque casi coincide con el parámetro poblacional. En ese caso no hay nada que inferir, porque ya se está trabajando con casi toda la población y se trata en realidad de un cálculo de estadística descriptiva.



Licencia: dominio público. Fuente: Pixabay.

## 15.2. Contraste de hipótesis

El contraste de hipótesis es otro de los procedimientos habituales en estadística inferencial y es una pieza clave en el método científico moderno. Comienza suponiendo un valor o rango de valores para el parámetro de una población, lo que constituye la hipótesis nula y se simboliza  $H_0$ . Esta es la hipótesis de trabajo que se contrastará con los datos empíricos que se obtengan de la muestra en estudio. Tras el contraste, la hipótesis nula puede resultar aceptada o rechazada; si ocurre lo segundo, entonces ha de ser cierta la hipótesis que supone justamente lo contrario, que se denominada hipótesis alternativa y se simboliza  $H_1$ .

La hipótesis nula suele expresar una ausencia de efecto en sentido general, tanto positivo como negativo. Por ejemplo, en una investigación acerca de un posible factor de riesgo para el desarrollo de una enfermedad, la hipótesis nula establecería que ese factor no tiene ningún efecto perjudicial. En el caso de una investigación acerca de un tratamiento médico para una enfermedad,

la hipótesis nula establecería que ese tratamiento no produce ningún efecto beneficioso. En ambos casos, la hipótesis nula debe asignarle al parámetro poblacional adecuado el valor correspondiente a una ausencia de efecto. Solo si la hipótesis nula es rechazada se podría considerar, con cierta probabilidad, que el factor realmente es de riesgo o que el tratamiento es verdaderamente efectivo, situaciones que vendrían descritas por la hipótesis alternativa.

El contraste estadístico de hipótesis se basa en un procedimiento opuesto en cierto modo al llevado a cabo para la estimación de parámetros, donde se calculaba el intervalo en el que se esperaría encontrar el parámetro poblacional conociendo el valor de su estadístico estimador en una muestra. En el contraste de hipótesis se trata de estimar el intervalo en el que se esperaría encontrar el valor del estadístico estimador en una muestra extraída al azar de una población cuyo valor del parámetro es el supuesto en la hipótesis nula.

Una vez que se ha obtenido ese intervalo, se calcula el valor que toma el estadístico estimador en la muestra que se está estudiando, que se llama valor empírico, y se comprueba si efectivamente se encuentra dentro del intervalo esperado. En caso afirmativo, se acepta la hipótesis nula, o, dicho con más rigor, no existen evidencias estadísticas que permitan rechazarla. En otras palabras, el valor del estadístico estimador en la muestra se encuentra suficientemente cercano al valor del parámetro poblacional dado en la hipótesis nula, de forma que este último puede considerarse correcto; el hecho de que ambos valores no sean exactamente iguales se debe únicamente al efecto del azar. Por esta razón, el intervalo de valores esperables se llama región o intervalo de aceptación.

Si, por el contrario, el estadístico estimador toma un valor empírico que se encuentra fuera del intervalo de aceptación, significa que es incompatible con que la población sea como supone la hipótesis nula, y por tanto se rechaza esta y se acepta en su lugar la hipótesis alternativa, que afirma lo opuesto de la nula.

Tanto en el caso de aceptación como de rechazo de la hipótesis nula, nunca se hace con certeza absoluta. El intervalo de aceptación se define para un cierto nivel de confianza,  $(1-\alpha)$ , que se fija al inicio del contraste. Un nivel de confianza de 0,95, por ejemplo, implica que, si la población fuese realmente como dice la hipótesis nula, el 95 % de las muestras extraídas al azar se encontrarían en el intervalo de aceptación, mientras que un 5 % se encontrarían fuera, por efecto del azar.

### Contraste de hipótesis para la media

El contraste de hipótesis comienza estableciendo la hipótesis nula, que supone que la media poblacional toma un cierto valor,  $\mu_0$ . La hipótesis alternativa afirma entonces que el valor de la media poblacional es distinto del que supone la hipótesis nula. Ambas se escriben simbólicamente de la siguiente manera:

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

Ya vimos anteriormente que el estadístico de inferencia asociado a la media poblacional es la media muestral tipificada, cuya distribución muestral, bajo ciertas condiciones, es la normal estándar. Por tanto, la probabilidad de encontrarla en el intervalo comprendido entre  $-z_{\alpha/2}$  y  $+z_{\alpha/2}$  es igual a  $(1-\alpha)$ , que se escribía como:

$$prob \left( -z_{\alpha/2} < \frac{m - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2} \right) = 1 - \alpha$$

Si en ese intervalo se despeja la media muestral para expresarla en función de la poblacional (al contrario de lo que se hizo para los intervalos de confianza), y se supone el valor  $\mu_0$  para esta última como indica la hipótesis nula, se obtiene:

$$prob \left( \mu_0 - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < m < \mu_0 + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right) = 1 - \alpha$$

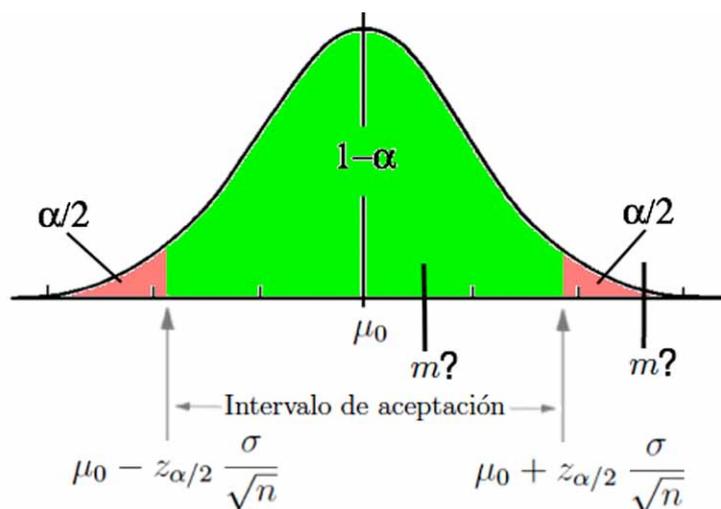
Es decir, la media de una muestra extraída al azar se esperaría encontrar con una probabilidad  $(1-\alpha)$  en el intervalo definido por los límites:

$$m_{lim} = \mu_0 \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Si el valor empírico de la media en la muestra en estudio,  $m$ , se encuentra efectivamente dentro de este intervalo, se acepta la hipótesis nula, es decir, no hay razones estadísticas para rechazar que la población tenga media  $\mu_0$ . Si, por el contrario, la media de la muestra estuviese fuera de este intervalo, se rechazaría la hipótesis nula y se aceptaría la alternativa.

Este contraste, donde la hipótesis nula supone que la media poblacional es igual a un cierto valor, se llama bilateral, porque el intervalo de aceptación se extiende hasta un cierto límite a ambos lados de la supuesta media, definiendo un área bajo la curva normal igual a  $(1-\alpha)$ . Por encima y por debajo de ese intervalo se sitúan dos intervalos de rechazo, cada uno de los cuales delimita un área bajo la curva normal igual a  $\alpha/2$ , que juntas corresponden al nivel de significación  $\alpha$ .

La siguiente gráfica<sup>8</sup> muestra el intervalo de aceptación para el contraste de hipótesis bilateral de la media poblacional, centrado en el valor supuesto en la hipótesis nula,  $\mu_0$ , junto con las áreas (probabilidades) bajo la curva normal asociadas a él. Se señalan dos posibles valores empíricos de la media muestral,  $m$ , uno que se encuentra dentro del intervalo de aceptación, con el que se aceptaría la hipótesis nula, y otro que se encuentra fuera, con el que se rechazaría la hipótesis nula.



Fuente: elaboración propia.

<sup>8</sup> Enlace relacionado: <http://view.genial.ly/5d6667984e0a540f4ef6aa88/interactive-content-interactive-image>

### EJEMPLO

La duración media de las bombillas de una determinada marca, según el fabricante, es de 10 000 horas, con una desviación típica de 800 horas. Se eligen al azar 80 bombillas de ese fabricante y se mide su duración, que en promedio resulta ser de 9720 horas. Para un nivel de confianza del 99 %, ¿es cierta la información que proporciona el fabricante?

El contraste de hipótesis (bilateral) que puede plantearse para responder esta pregunta es el siguiente:

$$H_0: \mu = 10000 \text{ horas}$$

$$H_1: \mu \neq 10000 \text{ horas}$$

A continuación, se obtiene el intervalo en el que cabría esperar que se encontrase la media muestral suponiendo cierta la hipótesis nula, con una confianza  $(1-\alpha) = 0,99$  (significación  $\alpha = 0,01$ ). El valor  $z_{\alpha/2} = z_{0,005}$  es el que deja a su derecha un área bajo la curva normal estándar de 0,005, y por tanto a su izquierda un área de 0,995. Buscando esta área (o la más cercana a ella) en la tabla de áreas bajo la curva normal estándar usada en problemas anteriores (ver páginas 132-133), se encuentra que está situada en la columna 2,00 y en la fila 0,58, que corresponde al valor  $z_{0,005} = 2,58$ . Los límites del intervalo de aceptación vienen dados entonces por  $10000 \pm 2,58 \times 800/\sqrt{80} = 10000 \pm 230,76$  horas.

Por otro lado, el valor empírico de la media en la muestra es  $m = 9720$  horas. Este valor se encuentra fuera del intervalo que va desde 9769,24 hasta 10230,76, es decir, está fuera del intervalo de aceptación de la hipótesis nula. La conclusión del contraste es por tanto que, con un nivel de confianza del 99 %, la información proporcionada por el fabricante es falsa, habiendo exagerado ligeramente la duración de sus bombillas.

Cuando la varianza en la población no se conoce y se usa en su lugar la varianza en la muestra, y el tamaño de esta es pequeño ( $n < 30$ ), se sigue un procedimiento análogo al anterior pero sustituyendo el estadístico  $z_{\alpha/2}$  por el  $t_{\alpha/2, n-1}$ , que sigue la distribución  $t$  de Student, como se hizo anteriormente para la estimación de la media poblacional.

También se puede establecer una hipótesis nula en la que la media poblacional se supone menor o mayor que un cierto valor  $\mu_0$ , en lugar de igual. Cuando la media de la muestra es mayor que ese valor supuesto ( $m > \mu_0$ ), las hipótesis se escriben como:

$$H_0 : \mu \leq \mu_0$$

$$H_1 : \mu > \mu_0$$

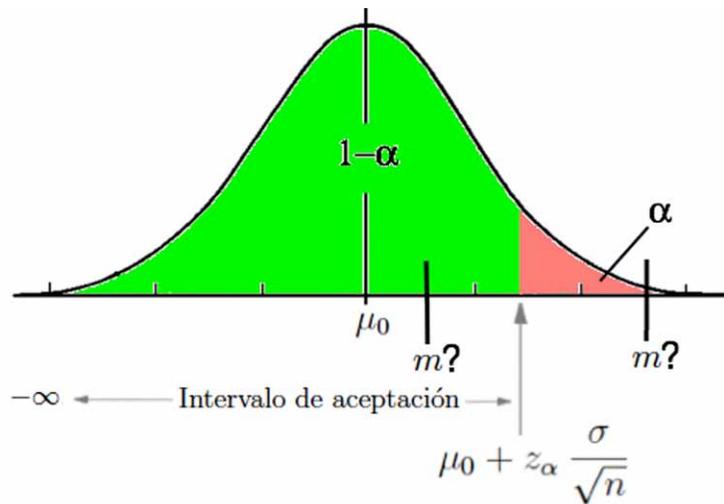
Así, la hipótesis nula expresa que la media de la población es menor o igual que el valor supuesto, mientras que la hipótesis alternativa afirma que es mayor. Aunque la media de la muestra con la que se trabaja es mayor que el valor supuesto, podría deberse al azar, y la probabilidad de que eso ocurra es lo que se calcula en el contraste. El intervalo de aceptación para la media muestral sería en este caso el de todos los valores menores que un cierto límite superior, es decir, desde menos infinito hasta ese límite, que viene dado por:

$$m_{lim} (sup) = \mu_0 + z_{\alpha} \frac{\sigma}{\sqrt{n}}$$

Si la media de la muestra es menor que el valor supuesto para la población ( $m < \mu_0$ ), las hipótesis nula y alternativa se formulan al contrario, es decir, la nula expresa que la media poblacional es mayor o igual que el valor supuesto. En este caso, el intervalo de aceptación es el opuesto al anterior: todos los valores mayores que un cierto límite inferior, que es igual al superior, pero reemplazando el signo más por el signo menos en su fórmula.

Este tipo de contraste se llama unilateral, porque el intervalo de aceptación se extiende desde un único límite hasta el infinito hacia la izquierda o hacia la derecha, y la única región de rechazo se extiende hasta el infinito opuesto. Esa única región de rechazo delimita el área total que corresponde al nivel de significación  $\alpha$ , razón por la cual el valor límite del intervalo depende de  $z_\alpha$  y no de  $z_{\alpha/2}$  como en el contraste bilateral.

El intervalo de aceptación y las áreas bajo la curva normal resultantes pueden verse en la siguiente gráfica<sup>9</sup>, análoga a la anterior, pero para un contraste unilateral con hipótesis nula  $\mu \leq \mu_0$ :



Fuente: elaboración propia.

**Contraste de hipótesis para la diferencia de medias**

También se puede llevar a cabo un contraste de hipótesis para la diferencia entre las medias de dos poblaciones distintas, estableciendo en la hipótesis nula que ambas medias son iguales y en la hipótesis alternativa, que son distintas (contraste bilateral):

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_1 : \mu_1 - \mu_2 \neq 0$$

Siguiendo un procedimiento análogo al del contraste para una única media, los límites del intervalo de aceptación de la hipótesis nula se obtienen a partir del estadístico de inferencia (diferencia de medias muestrales tipificada) y de la curva normal estándar, y resultan:

$$(m_1 - m_2)_{lim} = (\mu_1 - \mu_2) \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

<sup>9</sup> Enlace relacionado: <http://view.genial.ly/5d69775dd81fe03b6969b618/interactive-content-contraste-unilateral>

donde, según la hipótesis nula,  $\mu_1 - \mu_2$  es 0. Si la diferencia de medias muestrales,  $m_1 - m_2$ , se encuentra dentro de este intervalo, se acepta la hipótesis nula, es decir, las dos poblaciones tienen la misma media.

Al igual que en el caso de una sola media, también son posibles aquí los contrastes unilaterales. Y el contraste también puede ser más general fijando un valor distinto de 0 para la diferencia de medias en las hipótesis nula y alternativa.

El contraste para la diferencia de medias es muy importante en la investigación cuantitativa porque permite comparar los valores de la variable dependiente en dos grupos distintos, cada uno de los cuales está asociado a un valor (una categoría) diferente de la variable independiente. Se puede estudiar así si existe algún efecto estadísticamente significativo de la variable independiente sobre la dependiente. Ya se introdujeron en capítulos anteriores los diseños de investigación en los que se forman varios grupos según los valores de la variable independiente, que pueden usarse para estudiar, por ejemplo, la eficacia de un nuevo tratamiento médico, una nueva metodología educativa, una nueva política económica, etc.

### EJEMPLO

Dos grupos de 40 alumnos cada uno, elegidos al azar, han seguido metodologías distintas para el aprendizaje de inglés (variable independiente) durante un curso académico, al final del cual se han medido las competencias adquiridas mediante la realización de un test (variable dependiente). El grupo 1 ha obtenido en el test una puntuación media de 8,24 con varianza 4,49 y el grupo 2 ha obtenido una puntuación media de 7,32 con varianza 4,21. Para un nivel de confianza del 99 %, ¿puede decirse que las dos metodologías aplicadas producen diferencias significativas en los resultados de la competencia en inglés?

La resolución comienza estableciendo el contraste de hipótesis, que en este caso es para la igualdad de medias de dos poblaciones, o diferencia de medias entre ambas:

$H_0: \mu_1 = \mu_2$ , que equivale a  $H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 \neq \mu_2$ , que equivale a  $H_1: \mu_1 - \mu_2 \neq 0$

A continuación, se obtiene el intervalo en el que cabría esperar que se encontrase la diferencia de medias muestrales suponiendo cierta la hipótesis nula, con una confianza  $(1-\alpha) = 0,99$  (significación  $\alpha = 0,01$ ). Se trata de un contraste bilateral cuyo valor  $z_{\alpha/2} = z_{0,005}$  es el que deja a su derecha un área bajo la curva normal estándar de 0,005, y por tanto a su izquierda un área de 0,995. Buscando esta área en la tabla de áreas bajo la curva normal estándar, se encuentra que corresponde al valor  $z_{0,005} = 2,58$ . Los límites del intervalo de aceptación vienen dados entonces por  $0 \pm 2,58 \times \sqrt{(4,49/40 + 4,21/40)} = \pm 1,20$ . Por otro lado, el valor empírico de la diferencia de medias en la muestra es  $m_1 - m_2 = 8,24 - 7,32 = 0,92$ . Este valor se encuentra dentro del intervalo que va desde  $-1,20$  hasta  $1,20$ , es decir, está en el intervalo de aceptación de la hipótesis nula. La conclusión del contraste es por tanto que las dos metodologías de enseñanza de inglés no introducen diferencias estadísticamente significativas en cuanto a los resultados académicos, con un nivel de confianza del 99 %. Si alguno de los dos métodos era novedoso y se pretendía comparar con el tradicional, la investigación ha concluido que no introduce ningún cambio significativo. En los dos grupos de la muestra en estudio sí existía una clara diferencia de medias en la variable dependiente, pero el contraste estadístico ha establecido que se debe a efectos del azar, no a una influencia real de la variable independiente (todo ello dentro de la probabilidad fijada del 99 %).

### Contraste de hipótesis para la proporción

En este caso las hipótesis para el contraste (bilateral) son las siguientes:

$$H_0 : \pi = \pi_0$$

$$H_1 : \pi \neq \pi_0$$

Los límites del intervalo de aceptación de la hipótesis nula se obtienen a partir del estadístico de inferencia (proporción muestral tipificada) y de la curva normal estándar (para muestras grandes,  $n > 30$ ), y resultan:

$$plim = \pi_0 \pm z_{\alpha/2} \sqrt{\frac{\pi_0(1 - \pi_0)}{n}}$$

Si la proporción en la muestra,  $p$ , se encuentra dentro de este intervalo, se acepta la hipótesis nula, es decir, la población de la que se ha extraído la muestra tiene la proporción supuesta  $\pi_0$ .

### EJEMPLO

Se quiere comprobar la hipótesis de que el nivel de desempleo de una región es igual al del promedio de las regiones de su entorno, que se sitúa en un 12,6 %. Para ello se entrevista a un grupo de 800 personas escogidas al azar, de las que 120 resultan encontrarse en situación de desempleo. Para un nivel de confianza del 99 %, ¿se puede aceptar la hipótesis de partida? La hipótesis de partida (nula), junto a su hipótesis alternativa, pueden expresarse como:

$$H_0 : \pi = 0,126$$

$$H_1 : \pi \neq 0,126$$

Se trata de un contraste bilateral para la proporción poblacional. Dado que la muestra es grande, se puede suponer una distribución normal para la proporción muestral tipificada, con  $z_{\alpha/2} = z_{0,005} = 2,58$ , de donde se obtienen los siguientes límites del intervalo de aceptación:  $0,126 \pm 2,58 \times \sqrt{(0,126 \times (1-0,126) / 800)} = 0,126 \pm 0,030$ .

Por otro lado, el valor empírico de la proporción en la muestra es  $p = 120/800 = 0,15$ . Este valor se encuentra dentro del intervalo que va desde 0,096 hasta 0,156, es decir, está en el intervalo de aceptación de la hipótesis nula. Se concluye por tanto que no existen evidencias estadísticas que permitan descartar que el porcentaje de población desempleada en la región es igual que el promedio en las regiones de su entorno, con un nivel de confianza del 99 %.

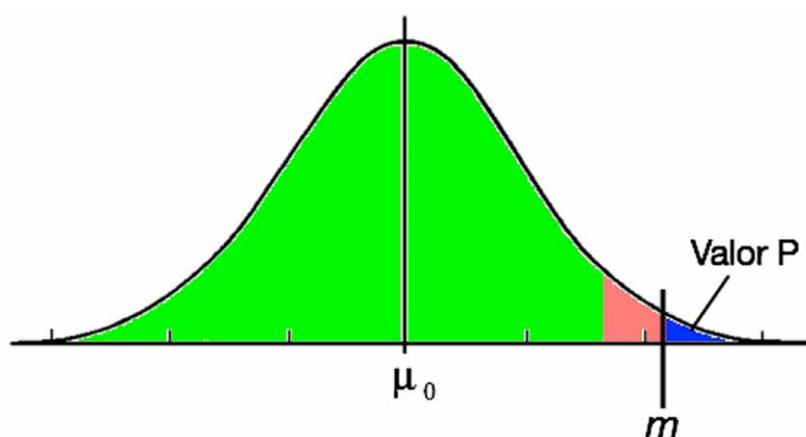


Licencia: dominio público. Fuente: Pixabay.

## AMPLIACIÓN. El valor $P$

El valor  $P$  en un contraste de hipótesis es la probabilidad de extraer al azar muestras tan diferentes de la población supuesta en la hipótesis nula como lo es la muestra que se maneja en la investigación. Al hablar de muestras «tan diferentes», se incluyen las que difieren lo mismo que la muestra en estudio y las que lo hacen aún más.

Gráficamente, el valor  $P$  corresponde al área bajo la curva de distribución muestral que queda a la derecha del valor empírico del estadístico en la muestra, es decir, para valores iguales o mayores. La hipótesis nula se rechaza si el valor  $P$  obtenido para el valor empírico del estadístico en la muestra es menor que el nivel de significación que se había fijado en el contraste. Así, en un contraste unilateral con nivel de confianza 0,95, que corresponde a un nivel de significación de 0,05, la hipótesis nula se rechaza si el valor  $P$  es menor de 0,05. Esta condición es totalmente equivalente a que el valor empírico se encuentre fuera del intervalo de aceptación, que era la forma en que se había expresado antes.



Fuente: elaboración propia.

Un valor  $P$  muy bajo significa que es muy poco probable que una muestra como la que se maneja haya salido de una población como la descrita por la hipótesis nula. En consecuencia, se rechaza la hipótesis nula, es decir, la población no tiene las propiedades que describe la hipótesis.

En realidad, las investigaciones no deben enfocarse tanto en la decisión de rechazo o no rechazo de la hipótesis nula. Al fin y al cabo, esa decisión depende del valor del nivel de confianza, y este se decide libremente. Lo que deben hacer es cuantificar cómo de probable es obtener una muestra al azar tan diferente a la población hipotética como lo es la muestra con la que se trabaja, medida que viene dada por el valor  $P$ . En muchas investigaciones, de hecho, se sobreentiende un nivel de significación de 0,05, y se rechaza la hipótesis nula si el valor  $P$  obtenido es menor, pero ni siquiera se dice cuánto vale. Lo más adecuado es dar el valor  $P$  para que cualquier investigador pueda valorar el grado en que los datos empíricos (muestrales) apoyan o no la hipótesis sobre la población.

### EJEMPLO

La hipótesis nula de una investigación supone que la estatura media de las mujeres españolas es 1,63 metros. Se extrae una muestra al azar de 10 mujeres, cuya estatura media resulta ser 1,84 metros, valor que se aleja considerablemente de la media muestral supuesta en la hipótesis nula. ¿Cómo de fácil, o de probable, sería encontrar en esa población muestras aleatorias de 10 personas tan alejadas o más de la media poblacional como la que se ha extraído? Para averiguarlo, se podrían sacar muchas muestras aleatorias distintas, cada una con 10 personas, y ver qué proporción de ellas presentan una media de estatura tan alejada o

más de la media poblacional como en nuestra muestra. Esta proporción es una probabilidad, y es la que se denomina valor  $P$  del contraste. En la práctica, se calcula el área bajo la curva normal hacia la derecha del valor dado por la media de la muestra,  $m$ . El área en esa cola de la curva normal (en azul en la figura anterior) será muy pequeña, porque la media de nuestra muestra está muy alejada de la media poblacional que supone la hipótesis nula. Eso quiere decir que es muy poco probable que nuestra muestra haya salido de una población con media 1,63 metros. Como la muestra ha sido extraída al azar de la población, lo que se rechaza es que la población tenga esa media, es decir, se rechaza la hipótesis nula.

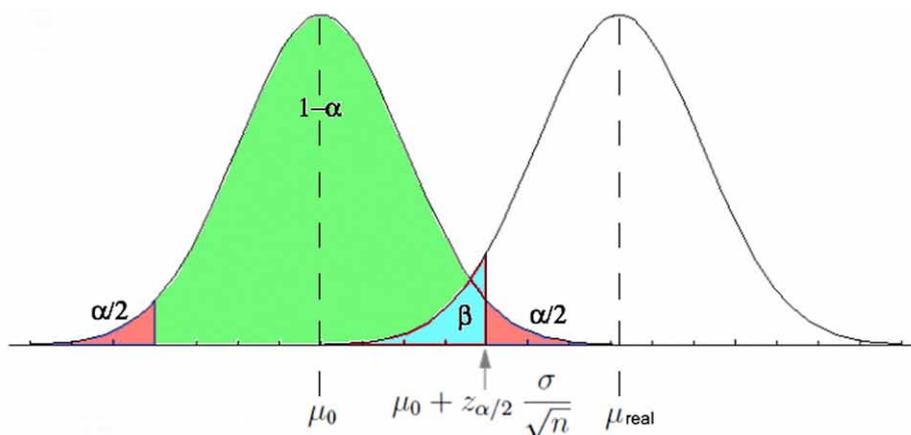
### 15.3. Errores en el contraste

El rechazo o no rechazo de una hipótesis solo se puede establecer con cierta probabilidad, como ocurre con todos los procedimientos en estadística inferencial. Por tanto, pueden producirse errores en el contraste, que consisten en rechazar hipótesis que son ciertas o en no rechazar hipótesis que son falsas.

Como la hipótesis nula planteada inicialmente puede ser cierta o falsa y el análisis estadístico puede rechazarla o no rechazarla, son posibles las cuatro situaciones siguientes:

1. La hipótesis nula es verdadera, pero el análisis estadístico la rechaza: se está cometiendo un error de tipo I, que ocurre con una probabilidad  $\alpha$  (nivel de significación).
2. La hipótesis nula es falsa, pero el análisis estadístico no la rechaza: se está cometiendo un error de tipo II, que ocurre con una probabilidad  $\beta$ .
3. La hipótesis nula es falsa y el análisis estadístico la rechaza: se trata de una decisión correcta que ocurre con una probabilidad dada por la potencia del contraste,  $(1-\beta)$ .
4. La hipótesis nula es verdadera y el análisis estadístico no la rechaza: se trata de una decisión correcta que ocurre con una probabilidad dada por el nivel de confianza del contraste,  $(1-\alpha)$ .

Ya definimos antes las probabilidades  $\alpha$  y  $(1-\alpha)$  y sus correspondientes áreas bajo la curva de la distribución muestral, usualmente la normal estándar. La probabilidad  $\beta$ , por su parte, es un área bajo la curva normal centrada en la media real de la población (que es una curva distinta a la centrada en la media supuesta en la hipótesis nula). Esta área se extiende desde el infinito negativo hasta el límite superior del intervalo de aceptación definido en la curva supuesta por la hipótesis nula (para  $\mu_{\text{real}} > \mu_0$ ), como se muestra en la siguiente figura<sup>10</sup>:



Fuente: elaboración propia.

<sup>10</sup> Enlace relacionado: <http://view.genial.ly/5d6983c0961a560fa9ca4dc2/interactive-content-errores-contraste>

Si la media real no se conoce, tampoco se conoce la posición de la distribución normal asociada, y por tanto la probabilidad  $\beta$  (área azul en la figura) no se puede calcular. Esta probabilidad de error es mayor cuanto más próxima se encuentra la media real de la población a la media supuesta en la hipótesis nula. Lo ideal en una investigación es que las probabilidades de los dos tipos de errores,  $\alpha$  y  $\beta$ , sean lo más pequeñas posible, pero para un tamaño de muestra fijo,  $\beta$  aumenta cuando  $\alpha$  disminuye (como se puede ver en la figura si se desplaza horizontalmente el segmento que delimita ambas áreas). Lo que se puede hacer es fijar inicialmente la probabilidad  $\alpha$  en el valor que se considere adecuado y aumentar el tamaño de la muestra en estudio,  $n$ ; así se reduce la desviación típica de la distribución muestral (que es  $\sigma/\sqrt{n}$ ), y por tanto la anchura de la curva normal correspondiente, con la consiguiente disminución del área  $\beta$ .

A los errores de tipo I también se les llama falsos positivos, y a los errores de tipo II, falsos negativos, especialmente en pruebas clínicas donde la hipótesis nula significa estado normal, es decir, que la persona está sana. Así, en este caso un error de tipo I significa que la persona está sana, pero la medida que se le toma (un análisis de sangre, por ejemplo) da un resultado que indica que está enferma, de donde viene el nombre de falso positivo. Un error de tipo II significa que la persona está enferma, pero la medida que se le toma da un resultado que no indica enfermedad, de donde viene el nombre de falso negativo. En este contexto, la población se podría asociar a todas las posibles medidas que se podrían tomar sobre un mismo sujeto, mientras que la muestra correspondería a la medida que realmente se toma; del resultado de esta última se tiene que inferir información sobre la población, es decir, sobre todas las posibles medidas, cuyo promedio sería un indicador adecuado de la presencia real o no de enfermedad en el sujeto.

Los falsos positivos y los falsos negativos no se deben a errores del experimentador o de los instrumentos, sino que dependen del propio diseño estadístico de las pruebas. Los falsos positivos aparecen porque la prueba se diseña para ser muy sensible a ligeras desviaciones respecto a los valores de una persona sana, aun a riesgo de que esas pequeñas variaciones se deban a otras causas y no a la presencia de enfermedad. En muchos casos son más peligrosos los falsos negativos que los falsos positivos, por lo que es preferible diseñar pruebas con alta sensibilidad (alta potencia,  $1-\beta$ ), es decir, que reduzcan la probabilidad de falso negativo aun a costa de aumentar la de falso positivo. Pero en otras ocasiones es necesario que la prueba tenga alta especificidad (alto nivel de confianza,  $1-\alpha$ ), es decir, que identifique con claridad una determinada circunstancia y no la confunda con otra, lo que implica reducir la probabilidad de falso positivo aun a costa de aumentar la de falso negativo.

### HISTORIAS DE LA CIENCIA. Falsos positivos y falsas intuiciones

El test de Mantoux consiste en una inyección intradérmica de antígenos que permite detectar la presencia en el organismo de la bacteria causante de la tuberculosis. Esta prueba tiene un porcentaje de falsos positivos de 7,5 %, que pueden deberse a la presencia de un cierto nivel de antígenos por haber recibido la vacuna antituberculosa o haber sufrido una infección de otro tipo de bacterias. El porcentaje de falsos negativos de la prueba es del 0,69 %. Para conseguir este porcentaje relativamente bajo de falsos negativos, que son más peligrosos, la prueba tiene que ser muy sensible, lo que aumenta el porcentaje de falsos positivos.

Supongamos que el 10 % de las personas que se someten a la prueba tienen realmente la enfermedad. En realidad, la prevalencia de la tuberculosis, es decir, el número de personas que padecen la enfermedad en un momento dado respecto a la población total, es muy baja (0,06 % en Europa), pero en las personas que se someten a las pruebas, muchas de las cuales presentan síntomas, la proporción es mayor.

Supongamos que se hacen la prueba diez mil personas. De ellas, 1000 (el 10%) estarán enfermas y 9000 sanas. De entre las enfermas, 69 dan (falso) negativo en el test (el 0,69 % de todas las sometidas a la prueba), mientras que las otras 931 dan positivo (real). De entre las sanas, 750 dan (falso) positivo (el 7,5 % de todas las sometidas a la prueba), mientras que las otras 8250 dan negativo (real). En total, entre falsos y correctos, se dan 8319 negativos y 1681 positivos.

Con estos datos se puede calcular que, si la prueba da negativo, aún existe una probabilidad de  $69/8319$ , o del 0,83 %, de estar enfermo (número de falsos negativos entre número total de negativos). Por otro lado, una persona que da positivo en la prueba tiene una probabilidad de  $931/1681$ , o del 55,4 %, de estar realmente enferma (número de positivos reales entre número total de positivos).

Se define la tasa de falsos positivos como el número de personas sanas que dan positivo dividido entre el número total de personas sanas que se hacen la prueba, en este caso  $750/9000$  o 8,33 %. Análogamente, la tasa de falsos negativos es el número de personas enfermas que dan negativo dividido entre el número total de personas enfermas que se hacen la prueba, en este caso  $69/1000$  o 6,9 %.

Es un error bastante común pensar que, como la tasa de falsos positivos es 8,33 %, la probabilidad de estar realmente enfermo cuando la prueba da positivo (verdadero positivo) es su complementaria,  $100 \% - 8,33 \% = 91,67 \%$ , en lugar del 55,4% obtenido antes. Pero ambos cálculos tienen significados distintos:

- 55,4% es la probabilidad de estar enfermo habiendo dado positivo en la prueba.
- 91,67 % es la probabilidad de dar negativo en la prueba estando sano (lo complementario al falso positivo), que se denomina especificidad de la prueba.

Tampoco debe confundirse con la probabilidad de dar positivo en la prueba estando enfermo, que se denomina sensibilidad de la prueba y en este caso es  $931/1000$  o 93,1 %.

La probabilidad de estar enfermo habiendo dado positivo en la prueba, 55,4 %, es relativamente baja porque influye mucho la incidencia de base del hecho que se estudia. En este caso, la incidencia de base de la tuberculosis es baja (10 % de los que se hacen la prueba) y el porcentaje de falsos positivos es alto (7,5 % de los que se hacen la prueba), de modo que, habiendo dado positivo en la prueba, es casi tan probable que se trate de uno de los falsos positivos como que se trate de uno de los verdaderos.



Test de Mantoux. Licencia: dominio público. Fuente: CDC-PHIL, Wikimedia Commons.

### AMPLIACIÓN. Contraste con otros parámetros

En este capítulo se han tratado en detalle tres parámetros poblacionales representativos: la media, la diferencia de medias de dos poblaciones y la proporción, cada uno de los cuales tiene asociado un estadístico de inferencia (tipificado) que sigue una distribución muestral de tipo normal estándar o  $t$  de Student.

Existen otros parámetros poblacionales además de los anteriores, algunos de los cuales están asociados a estadísticos de inferencia que siguen distribuciones muestrales distintas a la normal o a la  $t$  de Student, como son la distribución  $\chi^2$  (chi cuadrado) de Pearson o la distribución  $F$  de Fisher-Snedecor. Estas distribuciones son asimétricas respecto al valor de la media (es decir, diferentes a ambos lados), en contraste con la normal y la  $t$  de Student, que son simétricas.

En las tablas siguientes se recogen los parámetros más importantes tanto de una sola población como para la comparación entre dos poblaciones, junto con los estadísticos correspondientes que actúan como estimadores puntuales, los estadísticos de inferencia y la distribución muestral que sigue cada uno de ellos. Con todo ello se pueden calcular intervalos de confianza o contrastar hipótesis para esos parámetros.

## UNA POBLACIÓN

Parámetro poblacional	Estadístico estimador puntual	Estadístico de inferencia	Distribución muestral
Proporción $\pi$	Proporción $p$	$\frac{p-\pi}{\sqrt{p(1-p)/n}}$	Normal est. $N(0,1)$
Media $\mu$ ( $\sigma^2$ conocida)	Media $m$	$\frac{m-\mu}{\sigma/\sqrt{n}}$	Normal est. $N(0,1)$
Media $\mu$ ( $\sigma^2$ desconocida, $n > 30$ )	Media $m$	$\frac{m-\mu}{s/\sqrt{n}}$	Normal est. $N(0,1)$
Media $\mu$ ( $\sigma^2$ desconocida, $n \leq 30$ )	Media $m$	$\frac{m-\mu}{s/\sqrt{n-1}}$	$t$ de Student $t_{n-1}$
Varianza $\sigma^2$	Varianza insesgada $\frac{n}{n-1}s^2$	$\frac{ns^2}{\sigma^2}$	$\chi^2$ de Pearson $\chi_{n-1}^2$

## DOS POBLACIONES

Parámetro poblacional	Estadístico estimador puntual	Estadístico de inferencia	Distribución muestral
Diferencia de proporciones $\pi_1 - \pi_2$  ( $n_1 + n_2 > 30$ )	Diferencia de proporciones $p_1 - p_2$	$\frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$	Normal est. $N(0,1)$
Diferencia de medias $\mu_1 - \mu_2$  ( $\sigma_1^2 \neq \sigma_2^2$ conocidas)	Diferencia de medias $m_1 - m_2$	$\frac{(m_1 - m_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	Normal est. $N(0,1)$
Diferencia de medias $\mu_1 - \mu_2$  ( $\sigma_1^2 \neq \sigma_2^2$ desconocidas, $n_1 + n_2 > 30$ )	Diferencia de medias $m_1 - m_2$	$\frac{(m_1 - m_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	Normal est. $N(0,1)$
Diferencia de medias $\mu_1 - \mu_2$  ( $\sigma_1^2 = \sigma_2^2$ desconocidas, $n_1 + n_2 \leq 30$ )	Diferencia de medias $m_1 - m_2$	$\frac{(m_1 - m_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$t$ de Student $t_{n_1 + n_2 - 2}$
Cociente de varianzas $\sigma_1^2 / \sigma_2^2$	Cociente de varianzas insesgadas $\frac{\frac{n_1}{n_2} \frac{(n_2 - 1)}{(n_1 - 1)} \frac{s_1^2}{s_2^2}}$	$\frac{n_1}{n_2} \frac{(n_2 - 1)}{(n_1 - 1)} \frac{s_1^2}{s_2^2}$	$F$ de Fisher -Snedecor $F_{(n_1 - 1), (n_2 - 1)}$
Coefficiente de correlación $\rho_{12}$	Coefficiente de correlación $r_{12}$	$\frac{\sqrt{n-2} r_{12}}{\sqrt{1-r_{12}^2}}$	$t$ de Student $t_{n-2}$

## AMPLIACIÓN. Análisis de varianza

Un aspecto importante de investigaciones cuantitativas algo más avanzadas es el análisis de varianza, o ANOVA (del inglés *analysis of variance*). Es propio de las investigaciones donde se forman varios grupos dentro de la muestra según los valores de la variable independiente, y se mide en cada uno de ellos los valores resultantes de la variable dependiente. El análisis consiste en identificar y cuantificar distintos tipos de varianza en esos grupos de valores.

Se llama varianza sistemática a la variabilidad o dispersión de los valores de la variable dependiente causada por variables conocidas, como la variable independiente o cualquier variable perturbadora que se esté teniendo en cuenta en la investigación. Puede ser de dos tipos: entre grupos y entre filas.

La varianza entre grupos proviene de la dispersión en las medias de la variable dependiente de cada uno de los grupos formados en la muestra según los valores de la variable independiente. En diseños de investigación experimentales, donde el valor de la variable independiente en cada grupo es manipulado por el investigador, se denomina varianza experimental. Esta varianza es muy importante porque ofrece indicios del efecto de la variable independiente sobre la dependiente. Si no tuviese ningún efecto, los grupos tendrían medias muy parecidas en la variable dependiente. En cambio, si la variable independiente tuviese un efecto importante, cada grupo tendría una media de la variable dependiente muy diferente. Este análisis ya se ha llevado a cabo en un apartado anterior para el caso de dos grupos, a través del contraste de hipótesis para la diferencia de medias, pero los procedimientos de ANOVA permiten hacerlo para más de dos grupos.

### EJEMPLO

En una investigación se establece como variable independiente el método didáctico aplicado a un conjunto de estudiantes y como variable dependiente su resultado académico. Se forman al azar cuatro grupos de estudiantes, en cada uno de los cuales se aplica un método didáctico diferente. Se miden los resultados académicos de los alumnos con un examen y se obtienen las medias de las puntuaciones en cada grupo, que son: 5,8; 6,4; 7,5; 8,3. La varianza sistemática entre grupos (experimental en este caso) corresponde entonces a la varianza entre todas esas medias, que resulta 0,935, y es un indicador del efecto que ha tenido el empleo de diferentes métodos didácticos en los resultados académicos.

Pueden existir en el estudio variables perturbadoras que es necesario controlar para que la relación entre la independiente y la dependiente no resulte alterada. Entre las técnicas de control de estas variables, que vimos en un capítulo anterior, se encuentra el emparejamiento o el bloqueo, cuya aplicación da lugar a parejas o a bloques de elementos según los valores de la variable de control. Se denomina varianza entre filas a la que proviene de la dispersión entre las medias de la variable dependiente de las diversas parejas o bloques. El cálculo es totalmente análogo al de la varianza entre grupos, solo que ahora esos grupos se denominan filas y se forman según los valores de la variable de control, no de la independiente.

		Grupo con distinto valor de la variable independiente		
		Grupo 1	Grupo 2	Grupo...
Grupos formados por emparejamiento o bloques con distinto valor de una variable perturbadora	Fila 1			
	Fila 2			
	Fila...			

Otro componente de la varianza, distinta de la sistemática, es la varianza intragrupo, que se debe a la dispersión que sigue existiendo en los valores de la variable dependiente dentro de los grupos formados según la variable independiente. Esa dispersión puede deberse a otras variables perturbadoras que no han sido controladas (quizá ni siquiera han sido identificadas), o bien puede ser el resultado de influencias

aleatorias sobre los elementos o en el propio proceso de toma de datos. Se llama varianza intragrupo porque se calcula dentro de los grupos formados en la investigación, no comparando unos grupos o filas con otros. Para obtener esta varianza se calcula en primer lugar la varianza de los datos dentro de cada grupo, y a continuación se calcula la media aritmética de las varianzas de todos los grupos.

Es importante señalar que parte de la varianza intragrupo puede pasar a formar parte de la sistemática si se identifican y controlan nuevas variables perturbadoras. Cuantas más variables de este tipo se consideren, más se reducirá la varianza intragrupos, y la que quede se acercará cada vez más a la causada específicamente por influencias aleatorias, tomando entonces el nombre de varianza aleatoria o varianza de error. Nunca se podrá tener la certeza de haber identificado todas las variables perturbadoras, pero al menos puede hacerse con las más importantes, dejando que el resto contribuya a la varianza de error.

En resumen, la varianza de los valores de la variable dependiente tiene un componente sistemático (que puede ser experimental), formado por varianza entre grupos y varianza entre filas, y un componente intragrupos (que puede ser principalmente aleatorio).

Además de esta identificación de componentes de la varianza, ANOVA permite llevar a cabo un contraste de hipótesis mediante estadística inferencial para comprobar si las medias de los diversos grupos de la muestra presentan diferencias estadísticamente significativas extrapolables a la población en general. El objetivo final es establecer si, a partir de los datos de la muestra y para un nivel de confianza dado, existe en la población una influencia de la variable independiente o de las variables perturbadoras controladas sobre la variable dependiente. El procedimiento resumido se recoge a continuación:

Para contrastar las hipótesis sobre la influencia de ciertas variables sobre la variable dependiente, es decir, si las medias de los diversos grupos formados son iguales o son diferentes en la población, se comprueba si la  $F$  empírica de la muestra, calculada según el procedimiento de la tabla de abajo, se encuentra dentro del intervalo de aceptación, que se obtiene a partir de la distribución  $F$  de Fisher-Snedecor.

El límite superior del intervalo de aceptación para estudiar la influencia de la variable independiente que da origen a los grupos viene dado por  $F(\alpha; gl_{\text{grupos}}; gl_{\text{error}})$ , mientras que para estudiar la influencia de la variable perturbadora controlada que da lugar a las filas viene dado por  $F(\alpha; gl_{\text{filas}}; gl_{\text{error}})$ . Estos valores, que dependen del nivel de significación del contraste ( $\alpha$ ) y de los grados de libertad ( $gl$ ) que se calculan como indica la tabla de abajo, se pueden buscar en una tabla de áreas bajo la curva de la distribución  $F$ .

Fuente de variación	Suma cuadrática SC	Grados de libertad gl	Media cuadrática MC	$F$ empírica
<b>Sistemática entre grupos</b>	Sumar cada grupo, elevarlo al cuadrado, dividir entre el número de elementos en el grupo; sumar esos valores y restar $K$	Nº grupos menos 1	$\frac{SC_{\text{grupos}}}{gl_{\text{grupos}}}$	$\frac{MC_{\text{grupos}}}{MC_{\text{error}}}$
<b>Sistemática entre filas</b>	Ordenar la variable dependiente según la variable de control. Sumar cada fila, elevarlo al cuadrado, dividir entre el número de grupos; sumar esos valores y restar $K$	Nº filas menos 1	$\frac{SC_{\text{filas}}}{gl_{\text{filas}}}$	$\frac{MC_{\text{filas}}}{MC_{\text{error}}}$
<b>Intragrupos (error)</b>	$SC_{\text{total}} - SC_{\text{grupos}} - SC_{\text{filas}}$	$gl_{\text{grupos}} \times gl_{\text{filas}}$	$\frac{SC_{\text{error}}}{gl_{\text{error}}}$	
<b>Total</b>	Sumar el cuadrado de todos los elementos y restar $K$	Nº total de elementos menos 1		

\* La cantidad  $K$  se obtiene sumando todos los elementos, elevando al cuadrado y dividiendo entre el número total de elementos.

Esquema del capítulo<sup>11</sup>



Fuente: elaboración propia con Genially.

<sup>11</sup> Enlace relacionado: <http://view.genial.ly/5d191532dcd5930f446c664e/horizontal-infographic-diagrams-estadistica-inferencial>

# Capítulo 16.

## Medidas numéricas

Las investigaciones cuantitativas manejan valores numéricos desde la fase de recogida de datos hasta la fase de elaboración del informe final, pasando por la fase de análisis, en la que se aplican los procedimientos de la estadística descriptiva o de la estadística inferencial. Resulta esencial expresar correctamente esos datos numéricos de manera que aporten toda la información correcta y relevante y eviten la espuria.

Los datos numéricos suelen obtenerse con instrumentos de medida que, entendidos en sentido amplio, incluyen desde un termómetro en una investigación de ciencias naturales hasta un test de personalidad en un estudio de ciencias sociales. No obstante, el contenido de esta sección tiene mayor aplicación en ciencias naturales, y especialmente en el trabajo de laboratorio. Cuando los valores numéricos se obtienen directamente de un instrumento de medida, el resultado se denomina medida directa, mientras que al combinar varias de estas se obtiene como resultado una medida indirecta.

La información que aporta una medida no consiste únicamente en su valor,  $x$ , sino también en su incertidumbre asociada,  $E$ , y en las unidades en las que están expresados ambos números, si las tienen:

$$\text{Medida} = (\text{valor} \pm \text{incertidumbre}) \text{ unidades} = (x \pm E) u$$

En ocasiones se emplea la incertidumbre relativa, que es el cociente entre la incertidumbre de la medida y su propio valor:  $\varepsilon = E/x$ . Se puede dar en porcentaje multiplicando su valor por 100.

### EJEMPLO

Una expresión correcta de una medida de longitud es la siguiente:  $(52,0 \pm 0,1)$  cm. Esta medida viene dada en unidades de centímetro, con símbolo cm, y el valor de 52,0 cm va asociado a una incertidumbre de 0,1 cm, es decir, la longitud real puede ser 0,1 cm

(un milímetro) mayor o menor que 52,0 cm, razón por la cual se sitúa entre ambos el símbolo conjunto de más y menos ( $\pm$ ). Dicho de otro modo, el valor real de la medida estará, según nuestras estimaciones, entre 51,9 cm y 52,1 cm. La incertidumbre relativa de esta medida es  $\varepsilon = 0,1/52,0 = 0,0019$ , o un 0,19 %.

La incertidumbre es muy importante porque los valores de la medida nunca son totalmente precisos. Por un lado, los instrumentos empleados tienen asociada una escala con un cierto tamaño o rango mínimo, por debajo del cual no se puede afinar la medida de manera rigurosa. Por otro lado, en el proceso de medida se pueden cometer errores, ya sean de apreciación de la persona que toma los datos o por la influencia del entorno. Por ejemplo, para medir la longitud se puede usar como instrumento una regla cuyas unidades más pequeñas marcadas sean los milímetros. En consecuencia, las medidas que se tomen con ella no podrán darse en general con mayor precisión que el milímetro. Además, para medir un objeto hay que alinear sus extremos con dos marcas de la regla, una inicial y otra final, y existe el riesgo de que el experimentador se confunda al mirarlo, o de que un pequeño movimiento interfiera con la lectura.

En las ciencias sociales, las incertidumbres en las medidas cuantitativas son inevitables, debido a la dificultad de medir, e incluso de definir, las variables relacionadas con las actividades humanas, y a la enorme cantidad de factores que intervienen en ellas. En ciencias naturales, los avances tecnológicos en la instrumentación permiten reducir cada vez más las incertidumbres de las medidas, pero sean más grandes o más pequeñas, siempre van a existir. De hecho, en el nivel más fundamental que conocemos de la naturaleza, descrito por la física cuántica, no se pueden evitar los efectos del azar, de la interacción de los instrumentos con los propios objetos que se quieren medir o de la imprecisión causada en ciertas variables cuando se miden otras distintas.

Aunque a veces se usan indistintamente los términos incertidumbre o error, tiene más sentido el primero de ellos. Estrictamente, el error es la diferencia entre el valor de nuestra medida y el valor real de la variable que estamos midiendo. Como este último no lo conocemos nunca directamente, sino a través de las medidas, nunca se podrá determinar con certeza esa diferencia. El término incertidumbre, en cambio, hace referencia a la duda existente en cuanto al valor real de una variable, y expresa la diferencia máxima, tanto por encima como por debajo, que se estima que podría existir entre el valor de nuestra medida y el valor real.

Por supuesto, existen medidas que son exactas, como las que provienen de un conteo preciso, por ejemplo cien personas o diez átomos. También carecen de incertidumbre los valores que provienen de definiciones y no de medidas, como son las constantes matemáticas ( $\pi$ ,  $e$ ).

En cuanto a las unidades, no describiremos aquí todas las que existen, porque es un tema demasiado amplio y muy específico de cada ciencia, aunque sí se deben mencionar algunos aspectos generales. Siempre hay que indicar las unidades en las que están expresadas las medidas, excepto en aquellas que no las tienen (cantidades adimensionales). En ciencias naturales existen diversos sistemas de unidades, y debe elegirse uno que resulte conveniente para el tema de la investigación en cuestión y mantenerlo a lo largo de todo el informe o artículo de investigación. Si no existen buenas razones para escoger otro sistema, lo más adecuado es usar el Sistema Internacional (SI). La definición nominal de las unidades se acuerda en organismos y reuniones internacionales, como la Conferencia General de Pesos y Medidas, y su definición experimental y estandarización se lleva a cabo en laboratorios de todo el mundo, como el Instituto Nacional de Estándares y Tecnología (NIST) de Estados Unidos<sup>12</sup>.

---

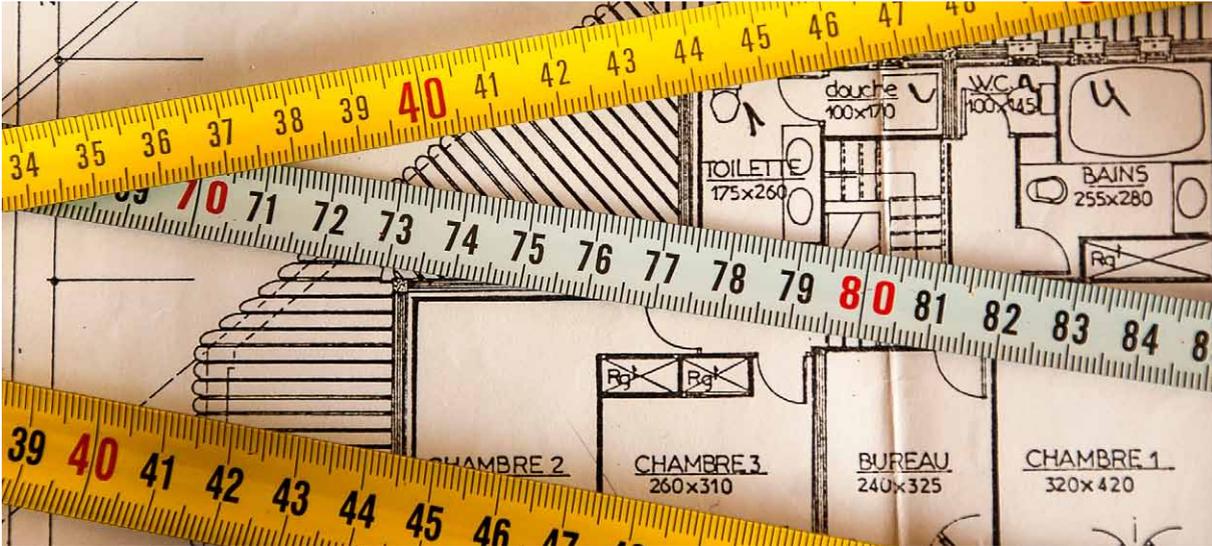
<sup>12</sup> Enlaces relacionados:

Sitio web del NIST y algunas de sus publicaciones sobre el Sistema Internacional de Unidades:

<http://www.nist.gov>

<http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.330-2019.pdf>

<http://physics.nist.gov/cuu/pdf/sp811.pdf>



Licencia: dominio público. Fuente: Pixabay.

### 16.1. Incertidumbre en medidas directas

En las medidas directas se distinguen dos fuentes de incertidumbre, la aleatoria,  $E_a$ , y la sistemática,  $E_s$ . La incertidumbre total de las medidas directas se obtiene combinando ambas a través de su suma cuadrática, es decir, elevando al cuadrado cada una de ellas, sumándolas y haciendo la raíz cuadrada del resultado:

$$E = \sqrt{E_a^2 + E_s^2}$$

#### **Incertidumbre aleatoria**

Está asociada a errores aleatorios, que en cada medida actúan con intensidad y dirección diferentes, incrementando o disminuyendo el valor real de la variable. Se deben a factores ambientales muy difíciles de controlar incluso en condiciones de laboratorio, como pequeñas variaciones de temperatura, vibraciones de los objetos que se miden o de los instrumentos de medida, interferencias de campos electromagnéticos, y muchas otras. También pueden tener su origen en la lectura de los instrumentos de medida o en la manipulación de los elementos en estudio por parte del experimentador.

Una manera de estimar la incertidumbre aleatoria es efectuar la misma medida varias veces y con las mismas condiciones, al menos con aquellas que son controlables por el experimentador. El hecho de que no se obtenga el mismo resultado en todas las medidas se debe precisamente a esos factores incontrolables que actúan de forma aleatoria. Como valor más fiable se toma la media aritmética de todas las medidas, en la que se compensan los posibles errores por exceso y por defecto de cada una de ellas. Cuantas más medidas se efectúen, más se reducirá esta incertidumbre, como se verá cuantitativamente a continuación.

Las medidas que se toman de una misma variable en condiciones similares constituyen la muestra, que se puede considerar un subconjunto de la hipotética población formada por las infinitas medidas de ese tipo que podrían efectuarse. Se calcula la media aritmética de las medidas realizadas, es decir, de los valores de la muestra, pero en realidad interesa la media de la población de la que han sido

extraídos, porque la media del conjunto infinito de medidas se supone igual al valor real de la variable. Se puede estimar un rango de valores para la media poblacional a partir de la media muestral haciendo uso de las técnicas de la estadística inferencial descritas en capítulos anteriores.

Como se deduce en estadística inferencial, las medias de las muestras extraídas de una población con media  $\mu$  y varianza  $\sigma^2$  siguen una distribución normal cuya media es la de la población y cuya varianza es la de la población dividida entre el tamaño de la muestra,  $\sigma^2/n$ . Cuando la varianza poblacional no se conoce, que es lo habitual, se supone igual a la varianza de la muestra,  $s^2$ , y la distribución muestral que sigue la media pasa a ser la  $t$  de Student. El intervalo de confianza para la media poblacional  $\mu$  viene definido entonces por los límites:

$$\mu_{lim} = m \pm t_{\alpha/2, n-1} \frac{s}{\sqrt{n-1}}$$

La distancia entre la media aritmética y los límites de este intervalo es lo que se define como la incertidumbre aleatoria (o estadística) de la medida:

$$E_a = t_{\alpha/2, n-1} \frac{s}{\sqrt{n-1}}$$

En una muestra aleatoria suficientemente grande, que en este caso significa un número suficientemente grande de medidas realizadas ( $n > 30$ ), la distribución de la media tiende a la normal, y se puede sustituir el factor  $t_{\alpha/2, n-1}$  por  $z_{\alpha/2}$ . En cualquier caso, siempre ha de indicarse el nivel de confianza ( $1-\alpha$ ) empleado para el cálculo de la incertidumbre aleatoria, por ejemplo 95 % o 99 %.

### **Incertidumbre sistemática**

La incertidumbre sistemática incluye todas las fuentes de error no aleatorias, es decir, que se supone que afectan de la misma manera a todas las medidas. Parte de esta incertidumbre proviene de la diferencia más pequeña de una variable que es capaz de distinguir un aparato de medida, que es lo que se conoce como su sensibilidad (a veces se usa también el término precisión), y suele coincidir con la división más pequeña de la escala del instrumento. Esta incertidumbre no disminuye con el número de medidas tomadas.

### **EJEMPLOS**

Las medidas de un termómetro con una escala que va de grado en grado irán asociadas a una incertidumbre sistemática de un grado ( $E_s = 1 \text{ }^\circ\text{C}$ ). Las medidas con una regla que marca los milímetros irán asociadas a una incertidumbre sistemática de un milímetro ( $E_s = 1 \text{ mm}$ ). Las lecturas de tiempo de un cronómetro digital que marca hasta las centésimas de segundo irán asociadas a una incertidumbre sistemática de una centésima de segundo ( $E_s = 0,01 \text{ s}$ ). En un instrumento sin escala propiamente dicha de lectura de medidas también se puede establecer una incertidumbre sistemática. Por ejemplo, en una balanza de platillos donde la masa de un objeto se obtiene equilibrándolo con el de pesas de masas conocidas, la incertidumbre sistemática será la de la pesa más ligera utilizada, que podría ser un gramo ( $E_s = 1 \text{ g}$ ).

A veces es necesario ajustar las incertidumbres sistemáticas asociadas a la sensibilidad del aparato de medida para tener en cuenta la habilidad del experimentador. Probablemente no habrá problema en leer las marcas de grado en un termómetro o las de milímetro en una regla, pero nuestra capacidad de reacción probablemente no será suficiente como para detener un cronómetro justo en la centésima de segundo en la que vemos terminar un proceso. En este caso lo más sensato sería usar una incertidumbre de una décima de segundo como mínimo, o incluso un segundo, independientemente de la sensibilidad del cronómetro. Este problema no surgiría, o se reduciría mucho, si la toma de medidas estuviese controlada electrónicamente, como ocurre en gran parte de la investigación científica actual.

Otra fuente habitual de incertidumbre sistemática es el denominado error de cero de los instrumentos de medida, relacionado con el origen de su escala. Existe un error de cero cuando el aparato arroja un valor de medida distinto de cero incluso cuando la variable en cuestión presenta un valor nulo. Es lo que ocurre, por ejemplo, si una balanza de precisión marca una masa antes de haber colocado ningún cuerpo sobre ella, o si un voltímetro o amperímetro marcan voltajes o intensidades eléctricas antes de que circule corriente por el circuito. Sin embargo, esta incertidumbre es fácilmente identificable y se puede eliminar reajustando el instrumento o restando a cada medida ese error.

La incertidumbre sistemática que se asocia a cada medida tiene que basarse siempre en el instrumento y en el método experimental empleado, y en el informe de la investigación ha de especificarse su valor y explicar su origen.



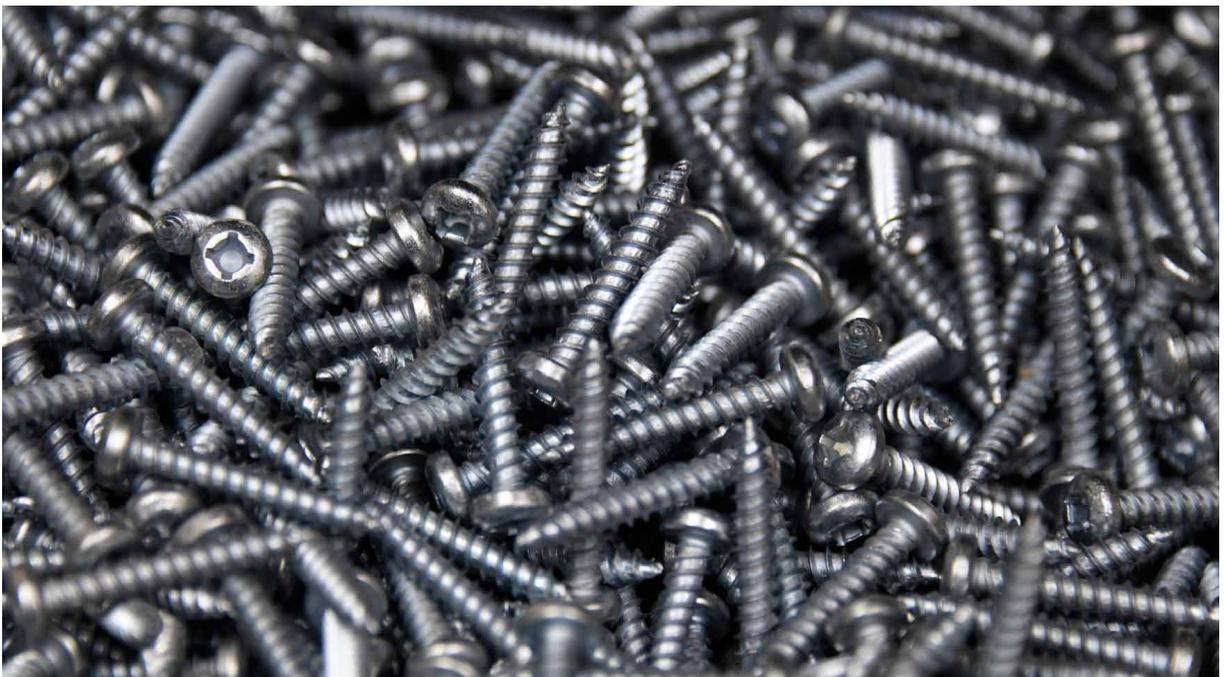
Licencia: dominio público. Fuente: Pixabay.

### ▶ AMPLIACIÓN. Matices sobre incertidumbres

Existe un matiz de interpretación en las incertidumbres aleatoria y sistemática aquí definidas y sus análogos en el desarrollo lógico que vimos en el capítulo de estadística inferencial. Allí se hablaba de determinar un parámetro poblacional, como la media, pero la población no era hipotéticamente infinita (aunque fuese muy grande) y el valor de la media no tenía una interpretación teórica o ideal como valor «real» de una variable, más allá de ser la media que se obtendría si se midiesen todos los elementos de la población.

Por ejemplo, diversas medidas de la longitud de un mismo tornillo pueden diferir de su longitud real debido a influencias aleatorias en el proceso de medida. Por otro lado, las longitudes de varios tornillos distintos pueden diferir de la longitud media de los tornillos de la caja de la que se extraen porque realmente cada tornillo tiene una longitud ligeramente diferente; esto se deberá probablemente también a influencias aleatorias, pero en el proceso de fabricación, no en el de medida. En el primer caso se habla de incertidumbre aleatoria de la medida, en el segundo se podría hablar de incertidumbre de muestreo, también de carácter aleatorio, porque depende de qué elementos salgan escogidos por azar para la muestra. En ambos casos, las incertidumbres disminuyen al aumentar el número de medidas o el tamaño de la muestra.

Este matiz también se puede trasladar a la incertidumbre sistemática. En el caso de las medidas se habla de desviaciones permanentes del método o instrumento de medida, mientras que en la estadística inferencial del capítulo anterior se refería a sesgos de selección de la muestra. Así, diversas medidas de la longitud de un mismo tornillo pueden diferir de su longitud real (todas ellas de la misma manera) debido a que el instrumento empleado tenga un error sistemático. Por otro lado, las longitudes de varios tornillos distintos pueden diferir de la longitud media de los tornillos de la caja de la que se extraen (todas ellas de la misma manera), porque se tiende a escoger, por ejemplo, los tornillos que sobresalen más porque son más largos, lo que da lugar a un sesgo de selección de la muestra. En ninguno de los dos casos disminuyen las incertidumbres al aumentar el número de medidas o el tamaño de la muestra.



Licencia: dominio público. Fuente: Pixabay.

## 16.2. Exactitud y precisión

Después de introducir las incertidumbres de tipo aleatorio y sistemático, es buen momento para aclarar la diferencia entre dos propiedades, la exactitud y la precisión, que se aplican al conjunto de medidas obtenidas con un determinado método experimental o instrumento.

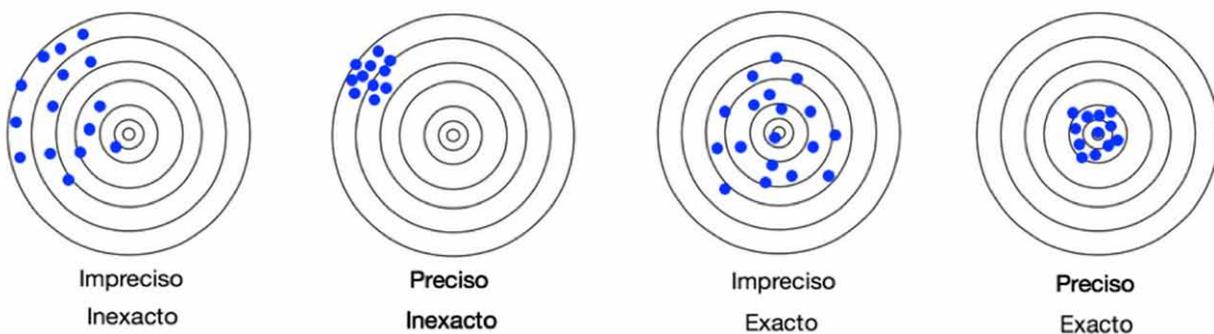
La exactitud es el grado de concordancia entre el resultado de las medidas y el valor verdadero de la magnitud que se pretende medir. No puede determinarse cuando el valor verdadero es desconocido, pero sí tiene sentido cuando este se conoce o cuando se comparan las medidas con la que

proporciona un método del que se sabe de antemano que tiene un alto grado de exactitud, y se emplea como referencia para la calibración. La exactitud está relacionada con la incertidumbre sistemática, que introduce desviaciones siempre en el mismo sentido que alejan el valor medido del verdadero (particularmente el error de cero).

La precisión, por su parte, es el grado de concordancia entre diferentes medidas realizadas en las mismas condiciones, al nivel en que pueden ser controladas por el método experimental en cuestión; si el conjunto de medidas presenta mucha dispersión, el método es impreciso, mientras que si están muy concentradas en torno a un valor, el método es preciso. La precisión está relacionada con la incertidumbre aleatoria, ya que el hecho de que se obtengan diferentes medidas en las mismas condiciones se debe a las influencias aleatorias.

La exactitud y la precisión son independientes entre sí. Por ejemplo, un método puede ser impreciso, dando medidas bastante distintas entre sí, pero cuyo promedio se acerca al valor verdadero y por tanto es exacto; y, al contrario, un método puede dar medidas muy similares entre sí pero estar todas ellas muy alejadas del valor verdadero. Lo habitual, sin embargo, es que la exactitud lleve asociada también buena precisión, pero no necesariamente al revés (debido a errores sistemáticos).

Una manera sencilla y visual de distinguir ambas propiedades es con el ejemplo de las dianas, cuyo centro representa la medida verdadera y donde los puntos de impacto de las flechas representan los resultados de diferentes medidas:



Fuente: elaboración propia.

### ▶ AMPLIACIÓN. Exactitud y precisión en ciencias sociales

Los conceptos de exactitud y precisión tal y como han sido descritos están adaptados preferentemente a las ciencias naturales, especialmente a las ciencias más cuantitativas como la física o la química. En ciencias sociales se pueden definir conceptos relacionados con ellos, pero con algunos matices diferentes.

Un concepto relacionado con la exactitud es la validez, que en este contexto se refiere a la concordancia entre lo que realmente se quiere medir (el constructo teórico) y las variables o indicadores que se definen para medirlo. Esta definición de validez está enfocada en la operacionalización, es decir, en la traducción de conceptos teóricos a variables medibles.

Por otro lado, un concepto relacionado con la precisión es la fiabilidad, que se refiere a la capacidad para obtener resultados estables y consistentes en distintas medidas realizadas sobre la misma magnitud.

Los conceptos de validez y fiabilidad son comunes en ciencias sociales, y especialmente en la teoría clásica de los test de psicometría, que es la rama de la psicología que se ocupa de la medición de variables psicológicas. En esta teoría, tanto la validez como la fiabilidad se definen a través de correlaciones (variaciones comunes de dos o más variables):

- La validez se mide a través de la correlación entre dos conjuntos de puntuaciones obtenidas con dos test diferentes diseñados para medir la misma variable (es decir, usando dos métodos distintos), o entre las puntuaciones de un test y las medidas de una cierta conducta (criterio).
- La fiabilidad se mide a través de la correlación entre dos conjuntos de puntuaciones obtenidas tras dos aplicaciones del mismo test en momentos distintos, o en dos mitades del mismo test, o en dos test similares (test paralelos).



Licencia: dominio público. Fuente: Pixabay.

### 16.3. Incertidumbre en medidas indirectas

Las medidas indirectas resultan de efectuar operaciones matemáticas entre medidas directas, como por ejemplo su suma o su producto. Si las medidas directas tienen incertidumbres asociadas, a la medida indirecta resultante le corresponderá también una cierta incertidumbre, que se obtiene a partir de las primeras a través de un cálculo de propagación de incertidumbres.

La medida indirecta es una función matemática de varias medidas directas, es decir, de varias variables, y su incertidumbre asociada se obtiene a través de derivadas parciales sobre esa función. La derivada de una función expresa cómo cambia su valor cuando cambia el valor de alguna de sus variables en una cantidad pequeña, que es lo que se conoce como diferencial. Cuando la función depende de una sola variable, se trata de la derivada habitual (derivada total). Cuando depende de varias variables, se puede calcular una derivada distinta respecto a cada una de ellas (derivada parcial), manteniendo constante el resto de variables.

Una pequeña variación de una función puede expresarse como la suma de las derivadas parciales respecto a cada una de sus variables multiplicada por una pequeña variación de esa variable. Tomando como ejemplo sencillo una función de dos variables,  $f(x,y)$ , su variación se expresaría matemáticamente de la siguiente manera:

$$df(x, y) = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

Como se puede observar, los símbolos de derivación parcial son distintos que en las derivadas totales, y para las pequeñas variaciones de la función y de las variables se ha empleado el símbolo de diferencial ( $df$ ,  $dx$ ,  $dy$ ).

La clave del método de propagación de incertidumbres consiste en que la incertidumbre en la medida directa de cada una de las variables se puede considerar una pequeña variación de esta; es decir, se puede identificar, por ejemplo, la incertidumbre de la variable  $x$ ,  $E(x)$ , con un diferencial de esa variable,  $dx$ . De hecho, la incertidumbre expresa un aumento o una disminución, generalmente pequeños, sobre el valor de la medida directa que se considera más fiable. La pequeña variación que se obtiene en la función que proporciona la medida indirecta,  $df$ , se puede identificar entonces como la incertidumbre de la medida indirecta,  $E(f)$ :

$$E(f) = \left| \frac{\partial f}{\partial x} \right| E(x) + \left| \frac{\partial f}{\partial y} \right| E(y)$$

Cada término se ha considerado en valor absoluto para que los signos de las derivadas parciales no compensen unas contribuciones con otras, porque no se sabe cómo se combinarán los errores en la realidad. Las incertidumbres de las medidas directas  $E(x)$  y  $E(y)$  son las totales, que contienen incertidumbre aleatoria, obtenida a partir de su desviación típica  $s$ . Teniendo esto último en cuenta, la combinación de las incertidumbres de medidas directas puede afinarse sabiendo que la varianza de una suma es igual a la suma de las varianzas. Recordando que la varianza es el cuadrado de la desviación típica, esto implica que habría que sumar los cuadrados de cada término (no los términos directamente), y el resultado sería el cuadrado de la incertidumbre indirecta, de modo que esta se obtendría según la expresión:

$$E(f) = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 [E(x)]^2 + \left(\frac{\partial f}{\partial y}\right)^2 [E(y)]^2}$$

Esta incertidumbre cuadrática es siempre menor que la obtenida sin introducir los cuadrados. El cálculo cuadrático puede ser más conveniente cuando la contribución aleatoria a las incertidumbres directas sea más relevante, pero en general ambos son válidos.

**EJEMPLO**

En una investigación sobre electricidad se cuenta con un circuito en el que se establece una diferencia de potencial  $V$  proporcionada por una pila, que se mide varias veces con un voltímetro. El circuito es atravesado por una corriente eléctrica de intensidad  $I$ , medida varias veces con un amperímetro. Cada una de estas dos medidas directas va asociada a una incertidumbre total,  $E(V)$  y  $E(I)$  respectivamente, que es combinación de incertidumbre sistemática y aleatoria.

La resistencia al paso de corriente ofrecida por los elementos del circuito se calcula, según la ley de Ohm, con la fórmula  $R = V / I$ . Se trata por tanto de una medida indirecta, cuya incertidumbre asociada (usando la expresión no cuadrática) viene dada por:

$$E(R) = \left| \frac{\partial R}{\partial V} \right| E(V) + \left| \frac{\partial R}{\partial I} \right| E(I) = \frac{1}{I} E(V) + \frac{V}{I^2} E(I) = R [\epsilon(V) + \epsilon(I)]$$

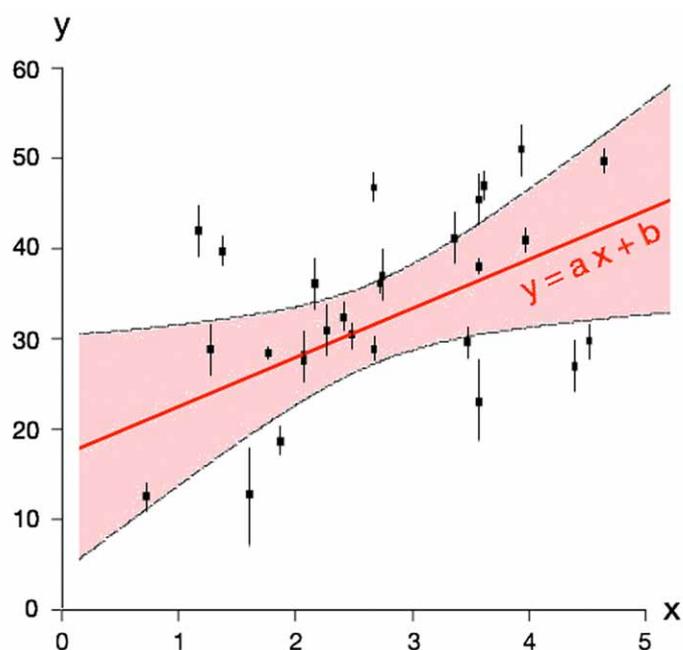
En el último paso la incertidumbre se ha escrito de manera más compacta expresándola en función de la propia medida indirecta  $R$  y de las incertidumbres relativas de las medidas directas,  $\epsilon(V) = E(V) / V$  y  $\epsilon(I) = E(I) / I$ .

Como ejemplo adicional se puede dar el de la medida indirecta de la potencia eléctrica del mismo circuito, que se obtiene como  $P = V \times I$ . Su incertidumbre (no cuadrática) viene dada por:

$$E(P) = \left| \frac{\partial P}{\partial V} \right| E(V) + \left| \frac{\partial P}{\partial I} \right| E(I) = I E(V) + V E(I) = P [\epsilon(V) + \epsilon(I)]$$

También llevan asociadas incertidumbres indirectas otros cálculos realizados sobre medidas directas, como los parámetros de una recta de regresión: pendiente,  $a$ , y ordenada en el origen,  $b$ . De hecho, al tener en cuenta la incertidumbre de cada uno de los datos de partida,  $x_i$  e  $y_i$ , el propio cálculo

de  $a$  y  $b$  es diferente del que se realizó en el capítulo de estadística descriptiva, porque debe asignarse mayor peso a los datos con menor incertidumbre. Calcular a mano los valores y las incertidumbres de los parámetros de la recta resulta tedioso, pero se pueden obtener fácilmente con *software* estadístico. En una representación gráfica, el efecto de estas incertidumbres se traduce en que la recta de regresión se extiende por una región delimitada por dos curvas, una situada por encima y otra por debajo de la recta original; se trata de curvas, y no de rectas, debido al efecto combinado de la incertidumbre en la ordenada en el origen (que traslada la recta en el eje  $y$ ) y de la incertidumbre en la pendiente (que hace bascular la recta).



Fuente: elaboración propia.

#### 16.4. Cifras significativas

En la expresión numérica de una medida se consideran cifras significativas aquellas que recogen información relevante y son fiables. En la incertidumbre se suelen tomar una o dos cifras significativas. En el valor de la medida, el número de cifras significativas depende de la incertidumbre asociada, como explicaremos en el apartado siguiente sobre presentación de medidas.

Para expresar correctamente medidas numéricas resulta a menudo necesario identificar cuáles son sus cifras significativas, reducir su número a través de un redondeo o expresar la medida en notación científica:

##### **Identificación de las cifras significativas**

Cuentan como cifras significativas las siguientes:

- Todas las cifras distintas de cero.
- Los ceros situados entre cifras distintas de cero.
- Los ceros situados a la derecha del valor, si lleva coma decimal.

Los ceros a la izquierda no cuentan como cifras significativas porque solo sirven para situar la coma decimal en su posición correcta. Tampoco cuentan los ceros a la derecha si no son decimales, porque solo sirven para colocar las cifras a su izquierda en su posición correcta.

**EJEMPLO**

Una cierta medida toma el valor 0,00000**520078900**, que contiene nueve cifras significativas, señaladas en negrita. Cuentan como significativas las cinco cifras distintas de 0, las dos cifras 0 situadas entre las cifras 2 y 7, y las dos cifras 0 situadas más a la derecha. No cuentan como cifras significativas las seis cifras 0 situadas a la izquierda, que sirven para colocar correctamente la coma decimal, de manera que el valor expresa algo más de cinco millonésimas.

La medida 0,00000**5200789** tiene el mismo valor que la del ejemplo anterior, en el sentido de que no es ni mayor ni menor, pero su precisión es diferente porque tiene siete cifras significativas en lugar de nueve (no tiene las dos cifras 0 a la derecha).

**Redondeo**

La correcta presentación de una medida exige a menudo eliminar cifras no significativas situadas a la derecha, redondeando el valor de la última cifra no eliminada. Cuando la primera de las cifras que hay que eliminar (empezando por la izquierda) es mayor o igual que 5, la cifra anterior se aumenta en una unidad, y si es menor que 5, la cifra anterior se queda como estaba.

**EJEMPLO**

Aplicando una serie de criterios que se describirán posteriormente, se determina que las medidas tomadas en una investigación tienen que presentarse con cifras significativas hasta la novena cifra decimal. Así, en la medida 0,00000520078900 la primera cifra que hay que eliminar es el 7 situado en la décima posición decimal; la última de las cifras significativas con las que quedará el número, que inicialmente es el 0 situado a la izquierda del 7, aumenta en una unidad por redondeo, dando como resultado correctamente presentado 0,000005201. Otros ejemplos: si se establece que son necesarias cifras significativas hasta la primera cifra decimal, la medida 26,3419 tiene que presentarse como 26,3 y la medida 26,3502 tiene que presentarse como 26,4.

**Expresión en notación científica**

La notación científica es muy útil porque facilita la comprensión del valor de las medidas y la comparación entre varias de ellas, especialmente cuando tienen valores muy grandes o muy pequeños. Además, ayuda a la identificación de las cifras significativas.

Para expresar una medida en notación científica se coloca la primera cifra significativa empezando por la izquierda en la posición de las unidades, seguida por la coma decimal y por el resto de cifras significativas. A continuación, el número obtenido se multiplica por la potencia de 10 necesaria para recuperar el número original.

**EJEMPLO**

La medida 0,00000520078900 se expresa en notación científica como  $5,20078900 \times 10^{-6}$ . La primera cifra significativa empezando por la izquierda, el 5, se ha trasladado a la posición de las unidades, y tras la coma decimal se han colocado el resto de cifras significativas. A continuación, se ha añadido un factor 10 elevado a la sexta potencia negativa, que es lo mismo que dividir entre un millón y que equivale a trasladar la cifra situada en la posición de las unidades hasta la sexta posición tras la coma decimal, la de las millonésimas, que era donde se encontraba en el número inicial.



Licencia: dominio público. Fuente: Pixabay.

Si se quisiera comparar la medida  $0,00000520078900$  con la medida  $0,000000520085610$ , habría que contar la posición tras la coma decimal en la que se sitúa la primera cifra significativa en cada una de ellas. Expresadas en notación científica resultan  $5,20078900 \times 10^{-6}$  y  $5,20085610 \times 10^{-7}$ , respectivamente, donde el exponente del factor 10 muestra con claridad que la segunda medida es menor que la primera casi en un factor 10, algo que resulta más difícil de apreciar en la notación convencional.

## 16.5. Presentación de medidas

Las incertidumbres de medidas directas, especialmente cuando contienen una contribución aleatoria que proviene de un cálculo estadístico, pueden obtenerse con una gran cantidad de cifras. Lo mismo ocurre con las incertidumbres de medidas indirectas, que también involucran operaciones matemáticas. Al tratarse de incertidumbres, y no de valores de la medida, es suficiente con conocerlas de manera aproximada, sin mucha precisión, por lo que muchos de los dígitos a la derecha pueden ignorarse. Así, por ejemplo, si la incertidumbre obtenida es  $9,175678\dots$ , será suficiente con saber que es de nueve unidades (9), o de nueve unidades y dos décimas (9,2), aplicando redondeo en ambos casos, pero no más.

En consecuencia, la incertidumbre que acompaña a un valor se redondea para dejar solo dos cifras significativas, a menos que ya de inicio solo contenga una, en cuyo caso se mantiene únicamente esa. Esta regla es una convención, y podría usarse otra diferente (por ejemplo, dejar la incertidumbre con una sola cifra significativa siempre). Lo importante es que se use siempre el mismo criterio a lo largo del documento que refleja el desarrollo o las conclusiones de una investigación.

En cuanto al valor de la medida, no tiene sentido darlo con más precisión que lo que indica su incertidumbre asociada. Por tanto, el valor de la medida se redondea a tantas posiciones decimales como las que tiene la incertidumbre. Por ejemplo, si el valor de la medida es  $619,52464$  y su incertidumbre es la del ejemplo anterior, que ya se había redondeado a 9,2, no tiene sentido mantener todas esas cifras decimales en el valor, porque no son fiables; como la incertidumbre tiene una posición decimal, el valor también debe redondearse hasta esa posición, quedando  $619,5$ . Siempre conviene que la medida contenga coma decimal, expresándola en notación científica si hace falta, para evitar posibles confusiones en cuanto a si los ceros a la derecha son o no cifras significativas.

Es muy importante presentar los resultados siempre de esta manera, con el número apropiado de cifras significativas, bien redondeados, mejor en notación científica (con la misma potencia de diez para el valor y para la incertidumbre), y por supuesto indicando las unidades si las hay (que también han de ser las mismas en el valor y en la incertidumbre).

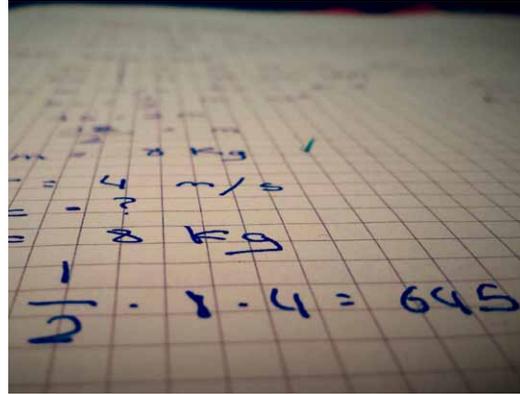
La reducción de cifras significativas a través del redondeo debe efectuarse solo en los resultados presentados, pero no en los que se van obteniendo en cada cálculo intermedio, porque podrían acumularse errores en el proceso. En los resultados intermedios conviene mantener dos cifras significativas más que las que quedarán en el valor final correctamente presentado.

### EJEMPLO

Se obtiene la medida  $0,00000520078900 \pm 0,0000000015$ . Su incertidumbre tiene dos cifras significativas (el 1 y el 5) situadas en la novena y en la décima posición decimal. El valor debe quedar por tanto con diez posiciones decimales tras su redondeo, resultando  $0,0000052008$ . En notación científica, tanto el valor como la incertidumbre correctamente presentados son:  $(5,2008 \pm 0,0015) \times 10^{-6}$ . Se observa que la misma potencia de diez multiplica tanto al valor como a su incertidumbre, encerrados en el mismo paréntesis.

## EJEMPLO

La medida  $5549392 \pm 2259$  no está correctamente expresada. La incertidumbre ha de redondearse hasta dejar dos únicas cifras significativas: 2300. Los dos últimos ceros no son en este caso cifras significativas (lo serían si hubiese coma decimal), pero deben mantenerse porque marcan las posiciones correctas de las cifras significativas (evidentemente, no tendría sentido redondear 2259 a 23). A continuación, el valor también ha de redondearse hasta las centenas, que es la posición de la cifra significativa más a la derecha en la incertidumbre, resultando 5549400 (de nuevo, los dos últimos ceros no son significativos). La expresión correcta de la medida resulta por tanto:  $5549400 \pm 2300$ . En notación científica se escribe como:  $(5,5494 \pm 0,0023) \times 10^6$ . Si se hubiera usado desde el principio esta notación, habría sido más fácil determinar las posiciones decimales que habría que mantener, cuatro en este caso, evitando posibles confusiones acerca del carácter significativo o no de los ceros a la derecha.



Licencia: dominio público. Fuente: Pixabay.

## 16.6. Representación gráfica

Las investigaciones cuantitativas, tanto del ámbito de las ciencias sociales como de las naturales, pueden reunir una gran cantidad de datos numéricos. A la hora de comunicar los resultados es necesario plantearse cuántos y cuáles de esos datos es razonable aportar en la memoria de la investigación o artículo científico, y cómo hacerlo.

Un conjunto de datos numéricos puede presentarse en forma de tabla, que permite clasificarlos, por ejemplo, según las condiciones en que han sido obtenidos o según los valores que tomaban las variables de control. Si el conjunto es demasiado amplio, quizá sea más conveniente representarlos gráficamente, que requiere menos espacio y además evidencia las regularidades o tendencias en los valores. Si la cantidad de datos brutos fuera tal que incluso su representación gráfica resultase abrumadora, puede optarse por mostrar algunos subconjuntos representativos, aunque el análisis estadístico y las conclusiones finales se basen en todo el conjunto recogido.

Dada la importancia que tienen las representaciones gráficas en los procedimientos científicos, se describirán aquí algunos de los tipos más importantes y algunas claves para su correcta elaboración.

### Gráfica de puntos o de dispersión

La gráfica de puntos o de dispersión se utiliza para representar pares de datos, correspondientes normalmente a valores de una variable independiente y de una variable dependiente. Consta de dos ejes perpendiculares entre sí que representan los valores de ambas variables, habitualmente los de la independiente en el eje horizontal y los de la dependiente en el eje vertical. Cada uno de los valores del par se sitúa en su eje correspondiente y desde esos puntos se trazan rectas perpendiculares a ese eje; en

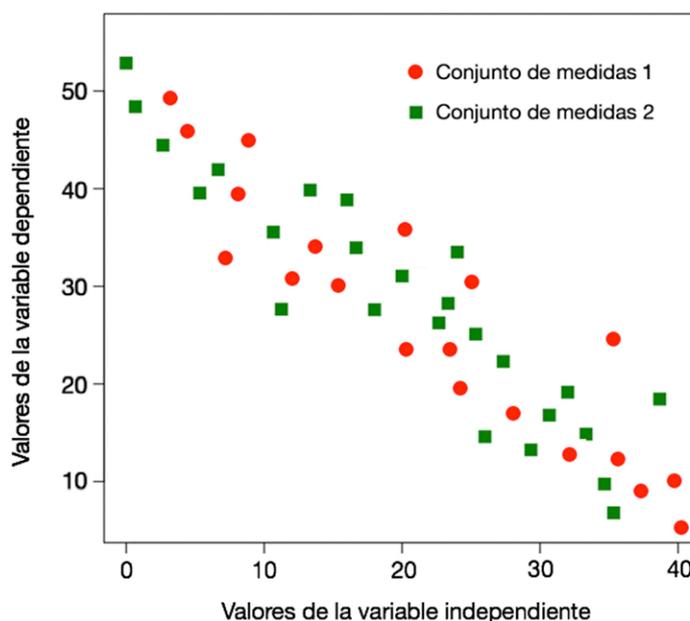


Licencia: dominio público. Fuente: Pixabay.

la intersección entre ambas se sitúa un punto, que representa la pareja de datos. Se trata de una representación cartesiana, como la que se usó cuando se introdujo la correlación entre dos series de datos.

Distintos grupos de datos pueden representarse en una misma gráfica usando símbolos o colores diferentes para los puntos de cada conjunto. Los símbolos deben ser suficientemente grandes como para distinguirlos claramente, y deben aparecer en una leyenda adjunta a la gráfica donde se especifique a qué grupo de datos corresponden.

En los ejes se indican los valores cada cierto intervalo, no muy corto para no saturarlos de números, pero tampoco muy amplio, de modo que se pueda estimar visualmente el valor que corresponde a los puntos de la gráfica de la manera más precisa posible. En cada eje ha de indicarse además la variable que representa y, muy importante, las unidades en las que están expresados sus valores, si tienen. Esto último es esencial, porque sin las unidades no se puede entender cuantitativamente la información de la gráfica. Todo este texto (valores en los ejes, nombre de la variable representada en cada uno de ellos, sus unidades) debe tener un tamaño adecuado para su fácil lectura.



Fuente: elaboración propia.

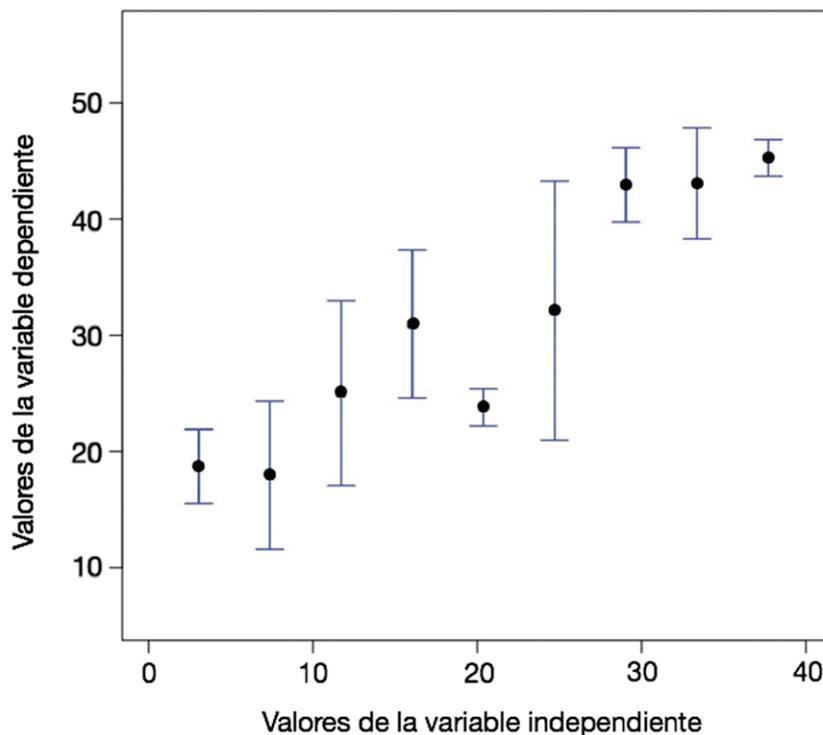
Los intervalos de valores que se señalan en los dos ejes pueden ser diferentes entre sí, y de hecho es lo habitual porque se refieren a variables distintas. Además, alguno de ellos o ambos pueden remplazar la escala lineal (la usual) por una escala logarítmica, en la que cada intervalo de la misma amplitud implica que el valor de la variable se multiplica por un mismo factor. Por ejemplo, en escala lineal los intervalos iguales en que se divide un eje pueden corresponder a un aumento de 10 unidades en el valor de la variable, mientras que en escala logarítmica pueden expresar que el valor de la variable se multiplica por 10; en el primer caso los intervalos podrían ir señalados en el eje con marcas de 0, 10, 20, 30, etc., mientras que en el segundo caso podrían ir con 1, 10, 100, 1000, etc. (el cero no puede aparecer en escala logarítmica, ni tampoco valores negativos). Con la escala logarítmica se consigue que el eje abarque un rango de valores muchísimo más amplio en el mismo espacio, lo que puede ser necesario en ocasiones; la desventaja es que la distribución de los puntos en la gráfica es más difícil de interpretar intuitivamente.

Los ejes no tienen por qué comenzar en el valor cero de su respectiva variable. Si los valores representados son muy altos, quedaría mucho espacio vacío en la gráfica, que no aporta nada. Para evitarlo, los ejes pueden cruzarse en el valor más conveniente de cada uno de ellos. A veces será necesario realizar esta modificación en los programas de trazado de gráficas, en algunos de los cuales los ejes se cruzan por defecto en sus respectivos valores cero.

Una vez representadas todas las parejas de valores, los puntos contiguos pueden unirse entre sí por segmentos rectos, dando lugar a un gráfico de líneas. Esta unión de los puntos se introduce para «guiar el ojo», pero más bien consigue confundir al cerebro, porque la línea resultante no contiene datos reales ni surge de un ajuste estadístico, por lo que su contenido científico es espurio. En términos generales, es mejor no trazar esos segmentos, y quitarlos en caso de que el programa estadístico empleado los introduzca por defecto.

Lo que sí puede tener sentido es incluir en la gráfica una curva de ajuste de los puntos, que se obtiene por procedimientos estadísticos, por ejemplo minimizando la separación al cuadrado de cada punto a esa curva (ajuste por mínimos cuadrados). En un capítulo anterior se trató el caso de la recta de regresión, pero lo mismo puede hacerse con otras curvas (por ejemplo, polinomios de cualquier grado o curvas exponenciales). Estas curvas de ajuste son suaves, porque no están formadas por segmentos concatenados que unen los puntos; si el ajuste estadístico es bueno, la curva sigue aproximadamente la distribución de los puntos, sin alejarse demasiado de ellos, aunque no tiene por qué pasar necesariamente por encima de todos (quizá de ninguno). Si se incluyen curvas de ajuste, es conveniente escribir su ecuación y los parámetros estadísticos del ajuste en la propia gráfica; en el caso de una recta de regresión habría que incluir la pendiente y la ordenada en el origen, así como el coeficiente de correlación o su cuadrado, el coeficiente de determinación.

Si los valores de las variables tienen incertidumbres asociadas es necesario reflejarlas en la gráfica mediante barras de error. Son segmentos que parten de cada uno de los puntos representados, y su longitud viene dada por la incertidumbre de los valores en la escala indicada por los ejes. Las barras de error se extienden hacia la derecha y hacia la izquierda de cada punto según la incertidumbre de la variable representada en el eje horizontal, y hacia arriba y hacia abajo según la incertidumbre de la variable representada en el eje vertical. Las barras de error pueden tener distinta longitud en cada punto de la gráfica, porque las medidas se pueden haber tomado en diferentes condiciones, dando lugar a distintas incertidumbres.



Fuente: elaboración propia.

### Gráfica de barras

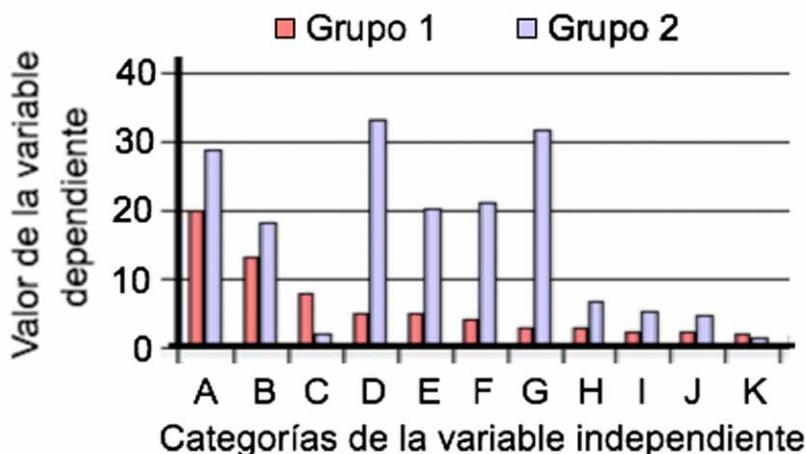
La gráfica de barras también se emplea para representar parejas de datos de dos variables relacionadas. Una de ellas, normalmente la independiente, ha de ser cualitativa o discreta, cuyos valores son categorías, o también puede ser una variable continua en la que se han establecido intervalos de valores.

Esta gráfica consta de dos ejes perpendiculares entre sí, en uno de los cuales (por ejemplo, en el horizontal) se sitúan las categorías de la variable independiente y en el otro (el vertical), los valores de la variable dependiente correspondientes a cada categoría. Cada categoría se asocia entonces a una barra rectangular vertical cuya altura representa el valor de la variable dependiente. Si las categorías de la variable independiente se sitúan en el eje vertical, se obtiene una gráfica de barras horizontales.

Cuando las categorías han sido construidas a partir de intervalos de valores con diferentes amplitudes, las barras pueden tener anchuras distintas, proporcionales a la amplitud de los intervalos. Si existen varios grupos diferentes de datos se pueden representar todos juntos trazando varias barras juntas dentro de cada categoría, o bien añadiendo unas barras a otras, por ejemplo apilándolas en vertical, si la suma de sus valores tiene algún significado. Las barras asociadas a diferentes grupos de datos pueden distinguirse con colores, indicando la correspondencia en una leyenda adjunta a la gráfica.

Como ocurre en el gráfico de puntos, todos los textos (valores numéricos, nombres de las categorías y de las variables en los ejes, la leyenda, etc.) deben tener un tamaño y una disposición adecuados para que la gráfica se entienda con claridad.

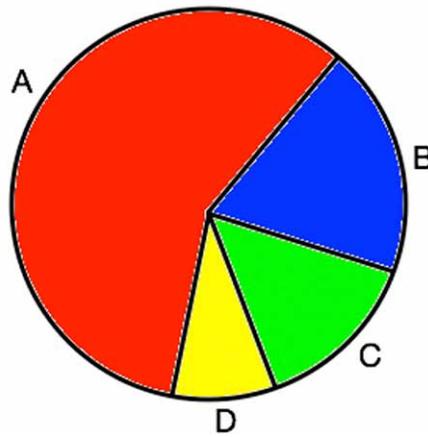
Las incertidumbres asociadas a los valores de la variable continua pueden incluirse como segmentos; si las barras son verticales, estos segmentos se extienden hacia arriba y hacia abajo a partir del extremo de la barra (hacia la izquierda y hacia la derecha en el caso de una barra horizontal).



Fuente: elaboración propia.

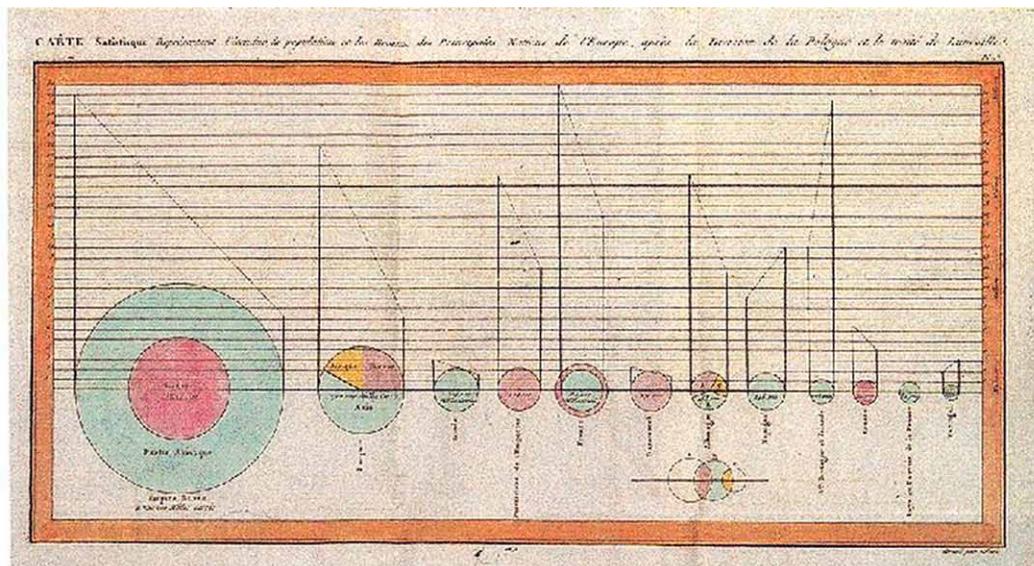
### Gráfica de sectores circulares

La gráfica de sectores circulares (llamada comúnmente «de tarta») se emplea, de nuevo, para representar parejas de datos de dos variables relacionadas, una de las cuales se mide en categorías. En este caso, la otra variable debe corresponder a la proporción o porcentaje respecto a un cierto valor total. El valor total es representado por un círculo, y cada una de las categorías se representa con un sector circular, similar a una porción de tarta, cuyo ángulo central o área es proporcional al valor asociado a la otra variable. En cada sector debe indicarse el nombre de la categoría a la que corresponde, o bien pueden llevar colores diferentes que se asocian a las categorías en una leyenda adjunta.



Fuente: elaboración propia.

Se pueden representar varias gráficas de tarta distintas para diferentes grupos de medidas. Si los valores totales sumando todas las categorías son diferentes en cada grupo, sus tartas correspondientes pueden dibujarse de tamaños distintos, proporcionales a ese valor total, lo que permite comparar los grupos entre sí; varios círculos de diferentes tamaños pueden dibujarse superpuestos. Las porciones pueden separarse ligeramente entre sí para mayor claridad (especialmente las que resultan muy finas y son difíciles de apreciar).



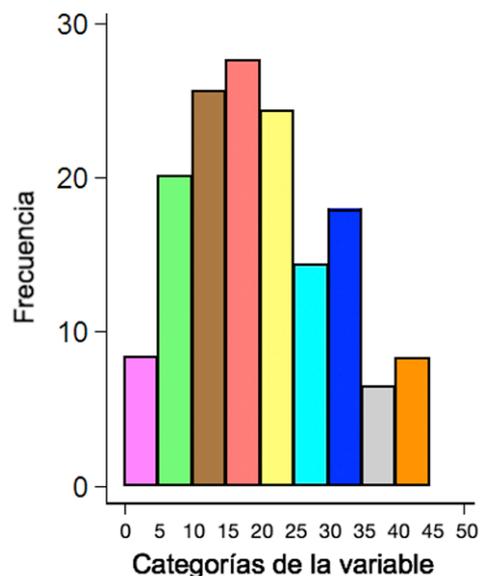
Primer uso conocido de gráficas de sectores circulares, en Statistical Breviary (1801) del ingeniero y economista escocés William Playfair (1759-1823). Licencia: dominio público. Fuente: Wikimedia Commons.

## Histograma

En las variables discretas (o continuas discretizadas mediante intervalos) puede registrarse el número de veces que aparece cada una de las categorías en los datos del estudio, que es su frecuencia. Su representación gráfica toma el nombre de histograma, y es similar a la gráfica de barras descrita anteriormente, pero en este caso la altura de cada barra representa la frecuencia de cada categoría. La frecuencia puede ser absoluta, si se refiere al número total de veces que ocurren elementos de cada categoría, o relativa, que es el

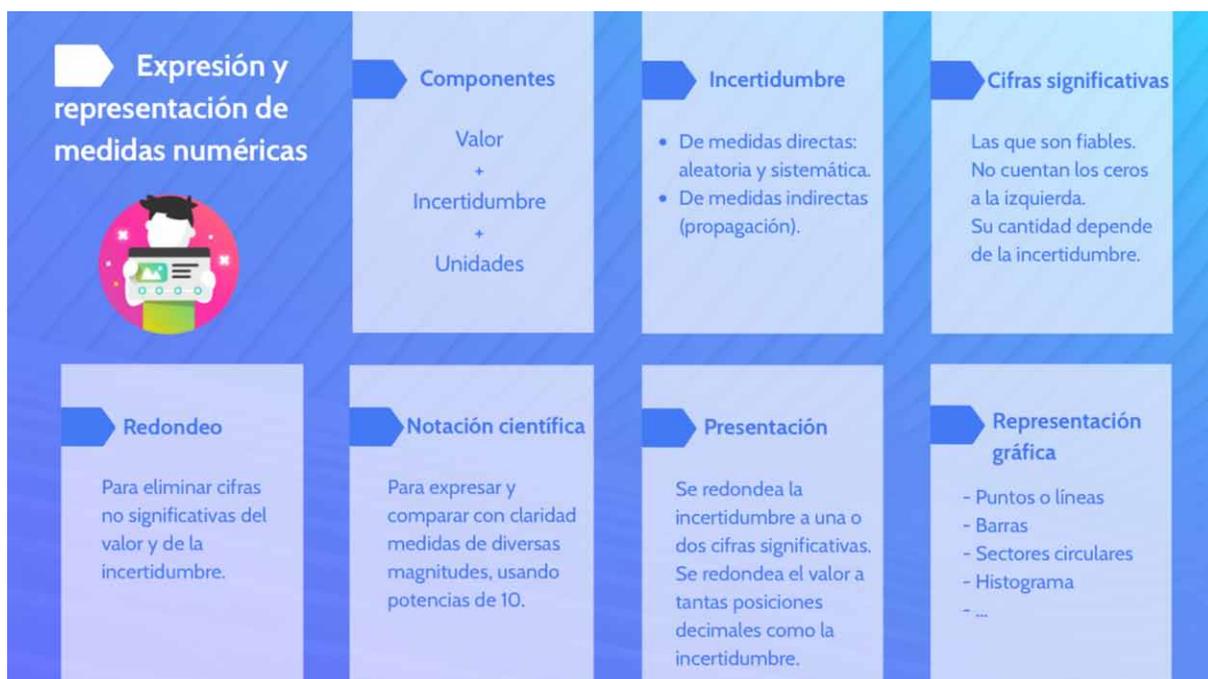
número de ocurrencias entre el número total de elementos; la suma de estas últimas es uno (o cien si están expresadas en porcentaje), y se pueden interpretar como las probabilidades de que ocurra cada categoría en ese conjunto de elementos. Es decir, el histograma con frecuencias relativas es análogo a una función de probabilidad, concepto que se introdujo al hablar de estadística inferencial.

Si las categorías han sido construidas como intervalos de valores de diferente longitud, los rectángulos tendrán diferentes tamaños en su base y la frecuencia será proporcional al área de la barra (no a su altura). Cuando los intervalos construidos son cada vez más cortos, la variable categórica se parece cada vez más a una continua, y el histograma con frecuencias relativas se convierte en una función de densidad de probabilidad; en ella, la probabilidad o frecuencia relativa en un cierto intervalo de la variable corresponde al área de la gráfica delimitada por ese intervalo, como se vio en capítulos anteriores sobre estadística inferencial.



Fuente: elaboración propia.

### Esquema del capítulo<sup>13</sup>



Fuente: elaboración propia con Genially.

<sup>13</sup> Enlace relacionado:

<http://view.genial.ly/5c7ac82008403f026132d84f/interactive-content-presentacion-de-medidas-numericas>

## **PARTE IV**



# Capítulo 17.

## La investigación científica en el aula

Los principios generales del método científico y las técnicas y procedimientos específicos vistos en los capítulos anteriores pueden aplicarse en el aula de secundaria para llevar a cabo una investigación científica adaptada al nivel de conocimientos de los estudiantes. El desarrollo de un trabajo de investigación puede introducirse como actividad dentro de una asignatura del currículo, como una materia independiente o incluso formar parte de una rama específica de Bachillerato. En todos los casos, el trabajo de investigación contribuye a adquirir y a profundizar las competencias clave de secundaria y a alcanzar algunos de los objetivos de las asignaturas de esta etapa.

Ciertas metodologías innovadoras son especialmente adecuadas, con las adaptaciones necesarias, para enseñar en el aula los procedimientos de la ciencia, pero lo más importante es transmitir una motivación por la actividad científica. Para contribuir a ella, en los próximos capítulos hablaremos de los certámenes científicos dirigidos a jóvenes investigadores, donde los estudiantes pueden presentar los trabajos que han llevado a cabo.



Licencia: dominio público. Fuente: Pixabay.

### 17.1. Objetivos y competencias clave

La introducción de la investigación científica en la enseñanza secundaria contribuye a una serie de objetivos de aprendizaje y a la adquisición de algunas de las competencias clave propias de este periodo.

Entre los objetivos perseguidos figuran los siguientes:

- Utilizar de manera autónoma conocimientos básicos de cada ciencia y las estrategias empleadas en su desarrollo, aplicando en particular el método científico en sus diversas variantes.
- Ser capaz de aplicar los conocimientos teóricos de cada ciencia a situaciones prácticas de interés, incluyendo aplicaciones tecnológicas.
- Desarrollar habilidades de diseño y realización de investigaciones científicas simplificadas, incluyendo protocolos de observación, experimentación, medida o recogida de datos, análisis, etc., con el fin de contrastar hipótesis científicas según los procedimientos del método científico general y específico de cada ciencia.
- Interpretar y expresar adecuadamente información científica transmitida oralmente o por escrito, a través de expresiones matemáticas, algoritmos y códigos informáticos, diagramas, gráficas, tablas, etc.
- Valorar las relaciones de las diversas ciencias con la tecnología, la sociedad, la economía, la salud y el medio ambiente.
- Entender los avances de las ciencias como aportaciones culturales, no solo porque el conocimiento científico forma parte de la cultura en sentido general, sino por su capacidad para influir en otras formas de expresión cultural y artística.
- Extrapolar los valores presentes en la investigación científica a otros ámbitos de las relaciones humanas: actitud abierta y flexible, fomento de la creatividad y la libertad, recurso a la evidencia y a la lógica, crítica rigurosa pero constructiva, etc.

Algunos de estos objetivos están relacionados con los recogidos en las legislaciones educativas vigentes, pero aquí han sido adaptados a la enseñanza de la investigación científica, es decir, a la creación de conocimiento científico, y no exclusivamente a su transmisión, que es lo que reflejan en mayor medida los currículos oficiales.

En cuanto a las capacidades o aptitudes que pueden adquirirse o afianzarse trabajando la investigación científica en la enseñanza secundaria, y que se relacionan con las competencias clave recogidas en las legislaciones estatal y autonómicas y en las recomendaciones del Consejo de la Unión Europea, se pueden nombrar las siguientes:<sup>14</sup>

- Competencia en comunicación lingüística.

Toda investigación parte del trabajo realizado previamente por otros científicos, cuyas publicaciones, muchas veces en un idioma distinto al nuestro, es necesario leer y comprender en profundidad. Además, la investigación propia ha de ser reflejada en una memoria o en un artículo, quizá también en un idioma distinto al nuestro, para darla a conocer al resto de la comunidad científica. En el contexto del aula de secundaria, los jóvenes autores de una investigación pueden exponer oralmente su trabajo a sus compañeros, o elaborar un póster que será expuesto en el centro. En todos estos actos comunicativos debe hacerse especial énfasis en la adaptación al público receptor, así como en la claridad y el rigor de la exposición.

- Competencia matemática y en ciencias y tecnología.

Las investigaciones en ciencias de la naturaleza y en matemáticas contribuyen de manera directa a la adquisición de esta competencia clave. Pero, más en general, toda investigación cuantitativa, incluyendo las de ciencias sociales, requiere un tratamiento matemático de los datos que afianza la competencia matemática, y además lo hace de una manera muy práctica, ya que el objetivo último es la expresión correcta de los datos recogidos (usando estadística descriptiva, gráficas, etc.) y el contraste de las hipótesis de investigación (usando estadística inferencial).

---

<sup>14</sup> Enlaces relacionados:

Orden ECD/65/2015, de 21 de enero, por la que se describen las relaciones entre las competencias, los contenidos y los criterios de evaluación de la Educación Primaria, la Educación Secundaria Obligatoria y el Bachillerato: <http://www.boe.es/buscar/doc.php?id=BOE-A-2015-738>

Recomendación del Consejo de la Unión Europea de 22 de mayo de 2018 relativa a las competencias clave para el aprendizaje permanente: [http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32018H0604\(01\)&from=EN](http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32018H0604(01)&from=EN)

Dentro de esta competencia, la Orden ECD/65/2015, de 21 de enero de 2015, incluye la siguiente descripción sobre la formación en investigación científica y en comunicación de la ciencia:

[...] la adquisición de las competencias en ciencia y tecnología requiere, de manera esencial, la formación y práctica en los siguientes dominios:

- Investigación científica: como recurso y procedimiento para conseguir los conocimientos científicos y tecnológicos logrados a lo largo de la historia. El acercamiento a los métodos propios de la actividad científica –propuesta de preguntas, búsqueda de soluciones, indagación de caminos posibles para la resolución de problemas, contrastación de pareceres, diseño de pruebas y experimentos, aprovechamiento de recursos inmediatos para la elaboración de material con fines experimentales y su adecuada utilización– no solo permite el aprendizaje de destrezas en ciencias y tecnologías, sino que también contribuye a la adquisición de actitudes y valores para la formación personal: atención, disciplina, rigor, paciencia, limpieza, serenidad, atrevimiento, riesgo y responsabilidad, etcétera.
- Comunicación de la ciencia: para transmitir adecuadamente los conocimientos, hallazgos y procesos. El uso correcto del lenguaje científico es una exigencia crucial de esta competencia: expresión numérica, manejo de unidades, indicación de operaciones, toma de datos, elaboración de tablas y gráficos, interpretación de los mismos, secuenciación de la información, deducción de leyes y su formalización matemática. También es esencial en esta dimensión competencial la unificación del lenguaje científico como medio para procurar el entendimiento, así como el compromiso de aplicarlo y respetarlo en las comunicaciones científicas.

– Competencia digital.

La recomendación del Consejo de la Unión Europea de 22 de mayo de 2018 relativa a las competencias clave para el aprendizaje permanente se refiere a la competencia digital en los siguientes términos:

La competencia digital implica el uso seguro, crítico y responsable de las tecnologías digitales para el aprendizaje, en el trabajo y para la participación en la sociedad, así como la interacción con estas. Incluye la alfabetización en información y datos, la comunicación y la colaboración, la alfabetización mediática, la creación de contenidos digitales (incluida la programación), la seguridad (incluido el bienestar digital y las competencias relacionadas con la ciberseguridad), asuntos relacionados con la propiedad intelectual, la resolución de problemas y el pensamiento crítico.

La mayor parte de la información científica necesaria para conocer el estado de la cuestión de un tema o para llevar a cabo una investigación secundaria se encuentra en formato digital y puede encontrarse en internet. Así pues, la investigación científica potencia las habilidades de búsqueda, filtrado y análisis de información en entornos digitales. Estos procedimientos forman parte de una actividad más amplia denominada curación de contenidos, en este caso de contenidos digitales de carácter científico.

Además, toda investigación hace uso actualmente de herramientas digitales para procesar los datos recogidos (bases de datos, hojas de cálculo, *software* de cálculo estadístico o simbólico, de representación gráfica, etc.) y para comunicar la investigación (procesadores de texto o sistemas de composición de texto como LaTeX, aplicaciones de creación de transparencias para presentaciones, infografías, ilustraciones, vídeos, etc.).

En definitiva, una investigación científica, por sencilla que sea, hace un uso intensivo de herramientas digitales diversas, por lo que trabaja de manera muy completa la competencia digital: utilización, acceso, filtrado, evaluación, creación, programación y difusión de contenidos digitales.

– Competencia para aprender a aprender.

Por la propia naturaleza de una investigación científica, es casi imposible saber *a priori* qué conocimientos concretos serán necesarios para llevarla a cabo. Un científico es una persona que aprende continuamente, por necesidad. Y por esa razón es alguien que ha desarrollado una especial capacidad para aprender a aprender, es decir, para identificar los conocimientos que necesita adquirir, para encontrar las fuentes más adecuadas y para asimilarlos de manera autónoma.

Un estudiante de secundaria que comience una investigación se dará cuenta enseguida que los contenidos de sus asignaturas son una base necesaria, pero seguramente no suficiente para culminar con éxito su tarea. Tendrá que aprender contenidos específicos por su cuenta, aunque sea guiado por sus profesores.

- Competencia del sentido de la iniciativa y del espíritu emprendedor.

La investigación científica exige siempre iniciativa, empezando por la elección del tema según las motivaciones e intereses propios. Requiere además una continua toma de decisiones para ajustarse a las dificultades que surgen en el proceso. No hay que olvidar que una investigación, aunque parta de unos objetivos muy concretos, nunca tiene un final cerrado y previamente conocido.

- Competencias sociales y cívicas.

La competencia social se ve fortalecida por el conocimiento nuevo generado en las investigaciones en ciencias sociales, relativo a la estructura y la dinámica de la sociedad, que permite entender sus fenómenos y sus problemas, además de buscar soluciones. La competencia cívica es impulsada desde las investigaciones en derecho, politología, historia, etc., que producen nuevo conocimiento y una crítica renovada de conceptos fundamentales como derechos y deberes, justicia, democracia, igualdad, etc.

Además, no deben olvidarse las profundas implicaciones sociales que pueden tener muchas investigaciones en ciencias naturales. Los avances en salud, que se traducen en mayor esperanza y calidad de vida, y las nuevas tecnologías, que trastocan el mercado de trabajo y los hábitos de consumo, producen cambios cada vez más rápidos en la organización social.

Así pues, la investigación científica en cualquier área contribuye a desarrollar las competencias sociales y cívicas en el aula, siempre con la orientación de un profesor que impulse la reflexión acerca del impacto que pueden tener esas investigaciones.

- Competencia en conciencia y expresión cultural.

Los resultados de la actividad científica influyen de manera esencial en la cultura de cada sociedad en cada periodo de tiempo. Llevar a cabo una investigación invita a (y a veces exige) reflexionar sobre los efectos que podría tener en el modo de vida individual y en la organización social y económica, lo que se traduce en una influencia en las expresiones culturales tanto artísticas (relacionadas con las bellas artes) como más cotidianas. A su vez, estas manifestaciones culturales pueden servir de inspiración para la creación de teorías científicas.



Licencia: dominio público. Fuente: Pixabay.

## 17.2. Metodologías de aprendizaje de la investigación científica

Algunas de las metodologías de aprendizaje más innovadoras, o que habitualmente se consideran como tales, presentan características muy apropiadas para la enseñanza del método científico y del proceso de investigación. No se trata aquí de describir metodologías para la didáctica de las ciencias, que son muchas y muy variadas, sino las que mejor pueden adaptarse a la enseñanza de la investigación científica, como las siguientes:

- Aprendizaje basado en proyectos.

Es una metodología con ciertas similitudes con la investigación científica, pero también con importantes diferencias.

El objetivo del aprendizaje basado en proyectos es que el estudiante adquiera un conocimiento ya establecido y con el que el profesor está familiarizado. La característica de esta metodología es que el profesor no transmite ese conocimiento en forma de clase magistral, sino que guía a los estudiantes para que lo encuentren acudiendo a diversas fuentes y lo combinen, sinteticen, asimilen

y valoren críticamente. Por el contrario, en una investigación científica se crea conocimiento nuevo aplicando el método científico<sup>15</sup>.

El proyecto de aprendizaje en esta metodología es elaborado por el profesor, partiendo de un problema complejo formulado habitualmente en forma de preguntas. El profesor se encarga de facilitar las herramientas y recursos para que los alumnos encuentren la solución correcta, que el profesor puede valorar porque ya la conoce, al menos en gran parte. En una investigación científica, en cambio, son los propios alumnos los que elaboran el proyecto de investigación, partiendo también de un problema para el que se construye una solución *a priori*, en forma de hipótesis de trabajo. El profesor puede ayudar aportando sus conocimientos y facilitando recursos, pero su guía no va encaminada a que los alumnos alcancen una respuesta correcta ya establecida, porque nadie sabe cuál es, ni si existe. En su lugar, las indicaciones del profesor deben ir encaminadas a que los alumnos apliquen correctamente el método científico en todos sus pasos, tanto en su versión general como en la específica para el área de la investigación.

Por supuesto, el aprendizaje basado en proyectos no se centra únicamente en la adquisición de conocimientos, sino también en la de una serie de habilidades y actitudes que son muy similares a las que se necesitan y se potencian en una investigación científica. Además, ambas actividades fomentan la colaboración entre los participantes y los convierten en protagonistas del proceso. Así pues, a pesar de las diferencias existentes, sería posible reformular y adaptar un aprendizaje basado en proyectos para transformarlo en una investigación científica en el aula, siempre que profesores y alumnos cuenten con la base necesaria para realizar ese salto.



Licencia: dominio público. Fuente: Pixabay.

– Metodologías basadas en la investigación.

Estas metodologías giran en torno a la resolución de problemas abiertos, partiendo de la formulación de hipótesis y procediendo a su contraste. Han recibido considerable atención en los últimos años por su capacidad de aumentar el interés de los alumnos por el aprendizaje de las ciencias, ya que imitan el proceso por el que históricamente se han adquirido los conocimientos científicos que se pretenden transmitir.

Entre estas metodologías se encuentran la educación científica basada en la indagación (*inquiry-based science education*, IBSE) o el aprendizaje basado en problemas (*problem-based learning*, PBL). Esta última tiene aspectos en común con el aprendizaje basado en proyectos, pero también se diferencia en algunos matices. Todas ellas comparten un papel muy activo por parte del alumno en el proceso de aprendizaje. Otra de estas metodologías es la de resolución de problemas como investigación (MRPI), que comienza con el análisis cualitativo de la situación problemática y su expresión en términos operativos, para proceder a con-



Licencia: dominio público. Fuente: Pixabay.

<sup>15</sup> Enlaces relacionados:

Para conocer de forma clara y concisa las claves del aprendizaje basado en proyectos se puede consultar:

<http://cedec.intef.es/7-elementos-esenciales-del-abp/>

<http://cedec.intef.es/7-practicas-docentes-abp/>

Para profundizar se puede seguir este curso ofrecido por el INTEF:

<http://formacion.intef.es/course/view.php?id=58>

tinuación a la formulación de hipótesis. El siguiente paso es el diseño de una estrategia de resolución y la identificación de los materiales y métodos necesarios. Ese diseño se lleva a la práctica para obtener unos resultados que serán finalmente analizados e interpretados en el marco de las hipótesis formuladas al inicio. Como se puede observar, se trata de un proceso muy inspirado en el método científico. Aunque su objetivo primario es el aprendizaje significativo de conceptos científicos ya establecidos, es evidente que también sirve para introducir a los alumnos en los procedimientos de la investigación propiamente dicha<sup>16</sup>.

– Gamificación.

Es una metodología que emplea elementos, mecánicas y diseños propios de los juegos para potenciar el proceso de aprendizaje y la resolución de problemas. El nombre proviene del inglés *game* (juego), aunque en castellano se prefiere usar el término ludificación.

Entre las ventajas que aportan las estructuras de juego en el aprendizaje destaca el aumento de la motivación extrínseca, al convertir en lúdica una actividad que requiere esfuerzo, lo que a su vez facilita una mayor retención en la memoria. A la motivación también contribuyen la asignación de puntuaciones, el paso de niveles o categorías, y otras recompensas habituales en el contexto de los juegos, que además fomentan al mismo tiempo la competitividad y la colaboración<sup>17</sup>.

Sin subestimar los beneficios de la gamificación en la didáctica de las ciencias, aquí interesa destacar el modo en que cierto tipo de juegos reproducen total o parcialmente el método científico, y que pueden servir por tanto para asimilarlo de una forma más amena. Se trata de aquellos juegos que requieren el uso de la inducción o de la deducción lógica, de la exploración de un entorno a través de «medidas», de la elaboración y contraste de hipótesis, del diseño de experimentos, etc., y que pueden contener además, como ocurre en la realidad, una cierta dosis de azar. En el Apéndice I se describen algunos juegos de este tipo.



Licencia: dominio público. Fuente: Pixabay.

De entre las metodologías anteriores, las dos primeras surgen de un enfoque cognitivista de la psicología, ya que exploran los mecanismos mentales involucrados en el aprendizaje, que coinciden en parte con los del método científico. En otras palabras, nuestro aprendizaje comparte ciertos procesos con el método científico, aunque este último presenta otros condicionantes. La gamificación, por su parte, surge de un enfoque más cercano al conductismo, ya que las reglas, premios o penalizaciones de los juegos están destinados a reforzar o debilitar conductas; no obstante, los juegos también pueden intentar combinar el enfoque conductista con el cognitivista si involucran mecanismos mentales de aprendizaje, como la deducción, la inducción, la causalidad, etc.

Algunas metodologías o recursos para la enseñanza de las ciencias (predominantemente las naturales) pueden encontrarse en los proyectos ENCIENDE (Enseñanza de las ciencias en la didáctica escolar), ACIERTAS (Aprendizaje de las ciencias por indagación en redes transversales colaborativas, de la Confederación de Sociedades Científicas de España, COSCE) o Engage (proyecto

<sup>16</sup> Enlaces relacionados:

Martínez Aznar, M. M. (2010) «La MRPI: una metodología investigativa para el desarrollo de las competencias científicas escolares en la educación primaria». En: Martínez Aznar, M. M., Garriga Mercader, L. y Langreo Valverde, S. (2010), *Educación científica «Abora»: el informe Rocard*, Madrid: Ministerio de Educación. Disponible en: <http://sede.educacion.gob.es/publiventa/PdfServlet?pdf=VP15136.pdf&area=E>

<sup>17</sup> Enlaces relacionados:

Para profundizar en la metodología de gamificación se puede seguir este curso del Intef: <http://formacion.intef.es/course/view.php?id=463>

de la Unión Europea destinado a capacitar a las generaciones futuras para participar en la ciencia y cambiar la forma en que se enseña)<sup>18</sup>.

### 17.3. La investigación como proyecto del centro

Una investigación científica en secundaria puede convertirse en un proyecto de todo el centro educativo. Además del profesor que ejerce como coordinador o tutor del trabajo, que probablemente pertenecerá al área de conocimiento del tema de la investigación, el resto del profesorado del centro puede contribuir aportando puntos de vista, sugerencias o directamente ayudando en problemas específicos. Por ejemplo:

- El profesorado del departamento de lengua puede sugerir mejoras en la redacción de los documentos escritos que recogen la investigación, como una memoria o un póster, o en la expresión de la exposición oral del trabajo.
- El profesorado de inglés puede revisar las secciones de la memoria de investigación que estén redactadas en esa lengua, que suele consistir únicamente en un breve resumen del trabajo, y puede ayudar también a los estudiantes a preparar breves conversaciones sobre su trabajo en inglés.
- El profesorado de matemáticas puede enseñar las técnicas estadísticas más adecuadas para el análisis de los datos recogidos.
- El profesorado de TIC puede ayudar con las herramientas informáticas necesarias para el desarrollo de la investigación y para la presentación de los resultados.
- Para investigaciones en ciencias naturales y matemáticas, el profesorado de física, química, biología, geología, tecnología, etc. puede contribuir a relacionar el tema de la investigación con otras áreas o con posibles aplicaciones tecnológicas, lo que sirve para contextualizar y motivar la investigación.
- Lo mismo puede darse en investigaciones de ciencias sociales y humanidades, donde el profesorado de historia, arte, lenguas o filosofía, así como los profesionales de orientación del centro, pueden contribuir a relacionar el tema de la investigación con otros campos, a contextualizar el estudio o a valorar sus implicaciones éticas. Esto último es muy importante cuando los sujetos de experimentación son personas o animales, o cuando los resultados de la investigación pueden tener efectos negativos para la sociedad.

Además del profesorado y de los autores, otros muchos estudiantes del centro pueden participar en la investigación de muy diversas maneras, tanto de forma activa, contribuyendo a las tareas de investigación, como de forma pasiva, siendo ellos mismos sujetos de investigación en trabajos de ciencias sociales y humanas.

La implicación por parte del centro educativo en el desarrollo de un trabajo de investigación puede institucionalizarse de diversas maneras, algunas de las cuales requieren el permiso de las administraciones educativas. Por ejemplo, puede contemplarse una modalidad de Bachillerato de Investigación, en el que los alumnos tienen que realizar un trabajo científico después de recibir la formación necesaria y específica para ello. En algunas comunidades autónomas esta modalidad forma parte del currículo oficial, como en la Región de Murcia (Bachillerato de Investigación), Castilla y León (Bachillerato de Investigación/Excelencia), o Cataluña (en forma de trabajo de investigación, *treball de recerca*, obligatorio para todo el alumnado). En otros países también es habitual una modalidad de Bachillerato o una materia

<sup>18</sup> Enlaces relacionados:

Proyecto ENCIENDE: <http://enciende.cosce.org>

Proyecto ACIERTAS: <http://aciertas.org>

Proyecto Engage: <http://www.engagingscience.eu/es/>

específica sobre investigación. Por ejemplo, en el Telebachillerato Comunitario de México, para el que se ha desarrollado un estupendo texto de libre distribución para alumnos de un nivel aproximadamente equivalente al de secundaria (15-19 años)<sup>19</sup>.



Licencia: dominio público. Fuente: Pixabay.

#### 17.4. Motivación

De la misma manera que la motivación y el aprendizaje están íntimamente ligados, también lo están la motivación y la investigación científica. La aplicación rigurosa del método científico para generar conocimiento requiere de un trabajo arduo de muchas etapas que difícilmente puede concluirse con éxito si el investigador no posee un gusto notable por la actividad que desarrolla.

Desde el aula puede trabajarse una motivación extrínseca hacia la investigación, acudiendo por ejemplo a la ciencia recreativa, basada en experimentos visualmente atractivos o problemas teóricos curiosos o con soluciones inesperadas, a la ciencia ficción de calidad o a la divulgación científica. También se pueden poner de relieve las implicaciones beneficiosas que tienen la ciencia y los avances tecnológicos en la salud y el bienestar, la sociedad o el medio ambiente. O incluso, desde una perspectiva más individualista, se puede mencionar el éxito profesional y prestigio social que puede alcanzarse si se cuenta con una carrera investigadora sólida.

Llevar a cabo un trabajo de investigación puede requerir además una importante motivación intrínseca, mediante la cual el investigador disfruta simplemente con la belleza de los hechos observados y analizados, explorando y comprendiendo su entorno natural o social, saciando su curiosidad científica y dando vía libre a su creatividad. Esta motivación intrínseca es más difícil de transmitir desde el aula, pero es muy importante porque es la que actúa en el largo plazo y en los momentos más duros de una investigación. Más que enseñarse, la motivación intrínseca se contagia, y ahí es fundamental el papel del profesorado, y en particular de los tutores y tutoras entusiastas que se animen a coordinar trabajos de investigación en el aula.

Como se mencionó antes, la divulgación científica es una buena herramienta para motivar el estudio de las ciencias y fomentar el interés por la investigación, ya que simplifica y hace accesible la ciencia sin recurrir en exceso al lenguaje matemático y siguiendo un estilo ágil, ameno y atractivo. Todo lo anterior ha de ser compatible, por supuesto, con la veracidad y el rigor: puede darse información científica limitada y simplificada, pero no incorrecta. El medio tradicional de la divulgación científica ha sido el ensayo escrito, pero cada vez hay más interés por parte de los medios de comunicación, y en los últimos años ha tenido un crecimiento exponencial en internet, especialmente en forma de vídeos o canales completos dedicados a ella.

De la misma manera que la didáctica de las ciencias ha de tener en cuenta la maduración intelectual y los conocimientos previos de los estudiantes, la divulgación también puede presentar distintos

---

<sup>19</sup> Enlace relacionado: <https://libros.conaliteg.gob.mx/TB11A.htm>

niveles de dificultad. Así, existe la divulgación para un público muy general (con conocimientos científicos preuniversitarios, seguramente ya bastante olvidados), para niños, para adolescentes, para estudiantes universitarios de ciencias e incluso para científicos profesionales.

Parte de la mejor divulgación es realizada por los propios investigadores que han contribuido notablemente al avance en sus respectivos campos, muchos de ellos galardonados con el Premio Nobel. Algunos de los nombres más conocidos son Stephen Hawking en física, Carl Sagan en astronomía, Marvin Harris en antropología, Richard Dawkins en genética y biología, Steven Pinker en psicología, etc. En España contamos también con investigadores que han contribuido a la divulgación, comenzando por Santiago Ramón y Cajal (*Reglas y consejos sobre investigación científica*, cuyo contenido es especialmente interesante para este curso)<sup>20</sup> y llegando hasta nuestros días con, por ejemplo, el paleoantropólogo Juan Luis Arsuaga, codirector del equipo de investigación de Atapuerca.

Otra actividad que puede motivar a los estudiantes para que se introduzcan en el mundo de la investigación científica es la participación en certámenes para jóvenes investigadores. El hecho de que los trabajos sean evaluados por científicos profesionales y de que puedan recibir premios, algunos de ellos importantes, puede dar un impulso considerable al interés de los jóvenes. En particular, dar a conocer y preparar la participación en el Certamen Jóvenes Investigadores de la Secretaría General de Universidades es uno de los objetivos principales de este libro, y a ello se dedicarán los siguientes capítulos.

## HISTORIAS DE LA CIENCIA. Isaac Asimov

Entre los escritores de divulgación científica, Isaac Asimov ocupa un lugar destacado por la amplísima variedad de temas científicos que trató en sus obras y por su capacidad para transmitir de forma accesible los conocimientos más avanzados en todas esas áreas.

Asimov nació en Rusia en 1920 pero a los tres años se trasladó a Nueva York con sus padres. Estudió bioquímica en la Universidad de Columbia y, tras doctorarse en química, trabajó en la Universidad de Boston como profesor asociado, pero pronto comenzó a ganar más dinero como escritor y se dedicó plenamente a ello. Es autor de centenares de obras, especialmente de ciencia ficción (como la saga *Fundación* o *Yō, robot*) y de divulgación científica. Entre estas últimas se encuentra una serie sobre divulgación histórica (desde la civilización egipcia hasta el nacimiento de los estados modernos) y numerosos ensayos sobre diversas ciencias naturales: química, bioquímica, biología, física, geología, astronomía, etc.; como ejemplo, y como invitación a su lectura, se pueden nombrar los siguientes títulos (de sus ediciones en castellano): *Momentos estelares de la ciencia*, *Breve historia de la química*, *Grandes ideas de la ciencia*, *El universo*, *100 preguntas básicas sobre la ciencia*, *El electrón es zurdo y otros ensayos científicos*, *Fotosíntesis*, *Nueva guía de la ciencia*, *Alpha Centauri, la estrella más próxima...*



Isaac Asimov. Licencia: dominio público.  
Fuente: Wikimedia Commons.

En definitiva, en lo que a divulgación científica se refiere, Asimov fue un auténtico hombre del Renacimiento, pero en el siglo xx. Aunque algunos contenidos de sus ensayos están algo desactualizados, siguen siendo un excelente recurso para introducir a los jóvenes y a los mayores en el mundo de la ciencia en general o en disciplinas concretas.

<sup>20</sup> Enlace relacionado: Reglas y consejos sobre investigación científica, de Santiago Ramón y Cajal: [http://cvc.cervantes.es/ciencia/cajal/cajal\\_reglas/](http://cvc.cervantes.es/ciencia/cajal/cajal_reglas/)

# Capítulo 18.

## El Certamen Jóvenes Investigadores

La Secretaría General de Universidades del Ministerio de Universidades convoca anualmente desde hace más de 30 años el Certamen Jóvenes Investigadores, que se organiza conjuntamente con el Instituto de la Juventud (Injuve)<sup>21</sup>. Además, colaboran otras instituciones educativas y de investigación como el Consejo Superior de Investigaciones Científicas (CSIC), universidades o sociedades científicas.

El Certamen tiene como objetivo promover el interés por la investigación entre los jóvenes mediante la concesión de premios a investigaciones, tanto básicas como aplicadas, relacionadas con cualquiera de las áreas del currículo de Enseñanza Secundaria Obligatoria, Bachillerato y Formación Profesional. Además, pretende fomentar la cultura científica y el pensamiento crítico en la juventud y en la sociedad en general.

Numerosos participantes de las pasadas ediciones del Certamen han desarrollado posteriormente notables carreras científicas, que consisten típicamente en la obtención de un título universitario, seguido de la realización de una tesis doctoral y de una investigación postdoctoral, esta última a menudo en centros de investigación o universidades fuera de España. Otros antiguos participantes trabajan en el sector privado, donde pueden desarrollar también una labor investigadora en los departamentos de I+D+i de las empresas. Y algunos han continuado su pasión por adquirir y transmitir conocimientos dedicándose a la docencia preuniversitaria, donde tienen la oportunidad de llevar a cabo la misión más importante de todas: iniciar la formación y fomentar la vocación de los investigadores del futuro.



---

<sup>21</sup> Enlace relacionado: Maleta didáctica elaborada por el Injuve sobre el Certamen Jóvenes Investigadores en su 30º aniversario: <http://www.injuve.es/talentojuven/noticia/maleta-didactica-30-aniversario-jovenes-investigadores>

En este capítulo se describen los aspectos más relevantes de las bases del Certamen Jóvenes Investigadores, que se recogen en la convocatoria publicada anualmente en el Boletín Oficial del Estado: requisitos de los jóvenes participantes, documentación que debe presentarse, criterios de valoración de los trabajos y premios que se conceden. Se concluye con una serie de reflexiones sobre los temas de investigación más adecuados para el Certamen.

## 18.1. Requisitos

Los trabajos presentados al Certamen pueden ser llevados a cabo individualmente o por grupos de hasta tres estudiantes, que deben estar matriculados en centros docentes españoles de Enseñanza Secundaria Obligatoria, Bachillerato o Formación Profesional. Sus edades deben estar comprendidas entre los 15 y los 20 años, aunque este último límite puede ser ampliado para participantes con discapacidad.

Es imprescindible consultar la convocatoria más reciente del Certamen para asegurarse de las fechas en que deben cumplirse los requisitos anteriores. Por ejemplo, en la convocatoria del año 2020 se establecía que los estudiantes tenían que haber cumplido 15 años de edad antes del 31 de diciembre de 2019, y no haber cumplido los 21 a fecha 30 de septiembre de 2020. En cuanto al requisito de matriculación en centros docentes, se refería al curso académico 2019/2020.

Los jóvenes participantes, ya sea uno solo o un grupo, deben ser coordinados y supervisados por un tutor, que será un profesor del centro docente en el que estén matriculados o un responsable de la asociación juvenil a la que pertenezcan. Este tutor será el encargado de cumplimentar y firmar la solicitud de participación en el Certamen.

Aunque solo tres estudiantes como máximo pueden ser considerados autores responsables de la investigación, y solamente un profesor puede ejercer de coordinador oficial, las contribuciones de más profesores y alumnos son posibles y deseables, debiendo reconocerse en el apartado correspondiente de la memoria de investigación.



Licencia: dominio público. Fuente: Pixabay.

## 18.2. Documentación

La solicitud de participación en el Certamen, que debe ser firmada y enviada por el profesor coordinador, tiene que ir acompañada de los siguientes documentos:

- Memoria de la investigación realizada, que consta de portada, índice, breve resumen (hasta 250 palabras en castellano y a continuación traducido al inglés), introducción, antecedentes, hipótesis y objetivos, materiales y métodos, resultados, conclusiones, agradecimientos, bibliografía y webgrafía y anexo de figuras. La extensión máxima de la memoria es de 10 páginas de texto y otras 10 páginas adicionales para el anexo de figuras. Los apartados mencionados son esencialmente los mismos que conforman los artículos publicados en revistas científicas especializadas; la explicación del contenido de cada uno de ellos se recoge en un capítulo posterior.



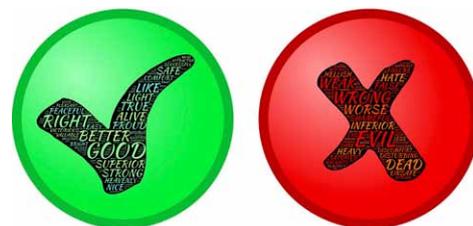
Licencia: dominio público. Fuente: Pixabay.

- Currículum del profesor coordinador, siguiendo el modelo especificado en la convocatoria.
- Fotocopia del Documento Nacional de Identidad de los jóvenes investigadores autores del trabajo.
- Certificado de matriculación de los jóvenes investigadores en un centro docente español de Enseñanza Secundaria Obligatoria, Bachillerato o Formación Profesional.

### 18.3. Criterios de valoración

Las investigaciones presentadas al Certamen se valoran atendiendo a los siguientes criterios:

- Originalidad.  
Las investigaciones deben ser novedosas, particularmente en sus objetivos y metodología, y no mera reproducción de otras realizadas con anterioridad por otras personas. También se valorará que la investigación presentada no haya sido publicada previamente por sus autores.
- Calidad del marco teórico.  
En la investigación debe quedar patente que los autores conocen las teorías y modelos científicos vigentes en los que se puede integrar el tema de la investigación llevada a cabo, y que son capaces de interpretar los resultados obtenidos en el contexto de una o varias de esas teorías.
- Correcto planteamiento de hipótesis científicas y objetivos de investigación.  
Se valora la correcta expresión y precisión científica de las hipótesis de trabajo planteadas, así como la claridad, adecuación y viabilidad de los objetivos propuestos para la investigación.
- Adecuación de la metodología empleada.  
Se considera en este criterio la correcta elección y justificación de la metodología necesaria para contrastar las hipótesis y alcanzar los objetivos planteados. Se tiene en cuenta el conocimiento por parte de los autores del método científico general y del particular en el ámbito del tema de investigación.
- Claridad en la exposición.  
Los aspectos más importantes de la investigación deben quedar reflejados de manera clara y precisa en la memoria escrita presentada. Se valora positivamente la organización de la memoria en los apartados descritos en la convocatoria, claramente identificados, paginados y recogidos en un índice inicial. Deben incorporarse en un anexo todas aquellas figuras que contribuyan a aclarar lo descrito en el texto.
- Correcta extracción y análisis de conclusiones.  
Las conclusiones extraídas deben reflejar el resultado del contraste de las hipótesis planteadas con los datos recogidos en la investigación, cumpliendo con los objetivos establecidos. Deben estar expresadas en la memoria con claridad y precisión, describiendo el proceso lógico mediante el que se ha llegado hasta ellas. Se tendrá en cuenta también el análisis e interpretación que se hace de las conclusiones en cuanto a alcance, implicaciones, aplicaciones, etc.
- Autonomía de diseño y realización de la investigación.  
La investigación debe haber sido diseñada y desarrollada por los jóvenes investigadores miembros del equipo, incluyendo la redacción de la memoria final presentada. No obstante, pueden haber recibido ayuda puntual del profesor coordinador y de otros adultos de su centro o asociación con conocimientos específicos de ciertos aspectos de la investigación o de la redacción de la memoria. Además, se valora positivamente el esfuerzo realizado por los jóvenes investigadores para obtener ayuda o apoyo puntual de profesionales del mundo académico e investigador.



Licencia: dominio público. Fuente: Pixabay.

– Utilización de recursos propios.

Para llevar a cabo el trabajo, los jóvenes investigadores deben haber empleado los recursos a su alcance, incluyendo los del centro educativo o asociación a la que pertenecen. Además, se valora positivamente que los jóvenes investigadores hayan realizado gestiones para usar de manera puntual recursos materiales y fuentes de información de universidades y centros de investigación.

#### 18.4. Premios

Para incentivar la participación de los jóvenes en el Certamen, la Secretaría General de Universidades y los organismos colaboradores destinan más de cien mil euros a la dotación de premios en metálico, además de conceder otros premios que financian estancias de investigación. Se seleccionan hasta 40 trabajos, cuyos autores, acompañados por su tutor, deben acudir al Congreso del Certamen Jóvenes Investigadores para presentar sus investigaciones.

Algunos ejemplos de premios concedidos en los Congresos más recientes son los siguientes:

- Un premio especial del Certamen a la investigación que reúna en mayor grado los méritos que se valoran, dotado con 5000 euros.
- Hasta seis primeros premios dotados con 4000 euros cada uno. Cinco de ellos irán destinados a la mejor investigación presentada en cada una de las cinco áreas consideradas en el Certamen, siempre que se hayan presentado trabajos en ellas: Artes y Humanidades, Ciencias Sociales y Jurídicas, Ciencias Físicas, Químicas y Matemáticas, Ciencias de la Tierra y de la Vida y Tecnología e Ingeniería. El premio restante podrá ser asignado a cualquiera de las áreas.
- Hasta diez segundos premios dotados con 3000 euros cada uno, de los que al menos uno de ellos será asignado a cada una de las cinco áreas del Certamen mencionadas antes.
- Hasta diez terceros premios dotados con 2000 euros cada uno, donde, de nuevo, al menos uno de ellos será asignado a cada una de las cinco áreas.
- Hasta diez accésits dotados con 1000 euros cada uno, repartidos por áreas como en los premios anteriores.
- Un premio, dotado con 3000 euros, a la mejor memoria de entre las investigaciones desarrolladas durante las estancias en el Consejo Superior de Investigaciones Científicas que se ofrecen como premio por esta institución, como se comentará más adelante.
- Hasta tres premios a tres trabajos para representar a España en el Certamen Europeo de Jóvenes Investigadores (EUCYS)<sup>22</sup> que se celebra cada año en un país distinto. La Comisión Europea se hace cargo de los gastos de desplazamiento desde un aeropuerto internacional español hasta el del lugar de celebración del Certamen Europeo y de los gastos de alojamiento y manutención para seis autores como máximo, de entre 14 y 20 años de edad y que no hayan completado su primer año de estudios universitarios en la fecha de celebración del Certamen Europeo.

Los premios en metálico se reparten equitativamente entre los autores de la investigación. Además, los profesores tutores de los siete trabajos que han recibido el premio especial o primeros premios reciben 2000 euros cada uno. Todos estos premios están exentos de IRPF.

Además de los premios propios del Certamen, las entidades colaboradoras ofrecen premios en metálico y en forma de estancias de investigación, estas últimas compatibles con los premios en metálico. Destacan los ocho premios otorgados por el Consejo Superior de Investigaciones Científicas (CSIC), que consisten en estancias para los autores de los trabajos de hasta dos semanas de duración en

<sup>22</sup> Enlace relacionado:

EUCYS: [http://ec.europa.eu/info/research-and-innovation/funding/funding-opportunities/eucys\\_en](http://ec.europa.eu/info/research-and-innovation/funding/funding-opportunities/eucys_en)

cualquiera de los centros de investigación del CSIC, para las que se cubren los gastos de desplazamiento y estancia hasta un máximo de 1000 euros por cada premio. También ofrecen estancias científicas la Sociedad Científica Informática de España y la Conferencia de Decanos y Directores de Ingeniería Informática de España, o las Reales Sociedades de Física y de Química; y ofrecen premios en metálico la Universidad de Málaga, dotado con 3000 euros, o la Fundación ONCE (al mejor trabajo de investigación dirigido a la mejora de la calidad de vida de las personas con discapacidad en ámbitos como la accesibilidad, la inclusión educativa y las ayudas técnicas).



Licencia: dominio público. Fuente: Pixabay.

Por último, el Certamen concede tres menciones de honor: al profesor coordinador de alguno de los equipos que haya llevado a cabo una labor especialmente destacable en cuanto al fomento de la investigación científica entre los jóvenes; al centro educativo o asociación juvenil que haya desarrollado una labor especialmente destacable en cuanto a la promoción de la cultura científica, tecnológica e investigadora entre los jóvenes; y a la persona o institución que haya destacado en el fomento de la cultura científica, tecnológica e investigadora entre los jóvenes.

## 18.5. Orientaciones sobre los temas de investigación

Las bases del Certamen Jóvenes Investigadores especifican que las investigaciones presentadas deben estar relacionadas con las áreas del currículo de la Enseñanza Secundaria Obligatoria, el Bachillerato o la Formación Profesional. En la práctica, este requisito apenas limita los temas investigados, lo que da lugar a una amplia variedad y riqueza.

Buscar un tema adecuado es la primera tarea de los jóvenes investigadores, ayudados por su tutor. Debe tenerse en cuenta los medios, conocimientos y capacidades de los autores, cuya edad estará comprendida entre los 15 y los 20 años. No obstante, se puede considerar, como ya se mencionó anteriormente, que los autores podrían obtener ayuda o apoyo puntual de profesores o investigadores, tanto en conocimientos como en recursos materiales o acceso a fuentes de información. También conviene mencionar que los conocimientos y competencias de los jóvenes no se tienen por qué limitar a los adquiridos en la enseñanza formal y reglada, ya que pueden alcanzar niveles mucho más altos en algunas áreas concretas si le dedican tiempo y esfuerzo; la curiosidad científica es sin duda una gran fuente de motivación para lograrlo.

Conviene elegir un tema original, interesante y atractivo; se puede recurrir a asuntos que estén de actualidad, o que puedan contribuir, aunque sea de forma modesta e indirecta, a solucionar problemas relevantes en ciencia básica o, más en general, para las personas y la sociedad. Puede resultar apropiado plantear investigaciones centradas en el entorno natural o social de los autores, donde es más improbable que se hayan realizado investigaciones similares en el pasado y por tanto resulten novedosas o incluso tengan aplicaciones beneficiosas en ese entorno. A veces el valor de una investigación puede residir principalmente en los materiales y métodos desarrollados, por ejemplo si resultan más sencillos, eficientes, baratos o accesibles que los ya existentes; podrían entonces ser usados en muchas otras investigaciones, no solo en la que presentan los autores.

Deben evitarse temas que sean ya muy conocidos en el ámbito docente o académico (por ejemplo, prácticas de laboratorio), porque el objetivo de toda investigación es crear conocimiento nuevo o herramientas novedosas para conseguirlo. También deben evitarse en este Certamen los trabajos de ciencia recreativa, porque la finalidad no es llamar la atención, ni siquiera hacer divulgación (al menos de forma prioritaria), sino, como ya se ha señalado, avanzar en el conocimiento. Ese avance

será probablemente modesto, pero ello no impide que sea científicamente riguroso y original en los resultados o en los métodos.

Para hacerse una idea del tipo de trabajos que son seleccionados para pasar a la segunda fase del Certamen Jóvenes Investigadores y son presentados en el Congreso de finalistas, se puede consultar por ejemplo el libro de actas del XXX Congreso<sup>23</sup>.

En el Apéndice III se describe una actividad que invita a reflexionar sobre una serie de aspectos importantes a la hora de plantear y desarrollar una investigación científica para el Certamen Jóvenes Investigadores: elección del tema de investigación y su tratamiento en el marco del método científico general, diseño de la investigación, análisis del control y la validez, tratamiento estadístico de datos numéricos, presentación de resultados, elaboración de la memoria escrita y de otros materiales para la presentación pública, y revisión y evaluación del trabajo una vez finalizado.

### AMPLIACIÓN. Otros certámenes para jóvenes investigadores

Además del Certamen Jóvenes Investigadores de la Secretaría General de Universidades, existen otros concursos de similares características convocados por organismos públicos, tanto nacionales como autonómicos, o por entidades privadas. En ocasiones, la temática se restringe a unas ciertas áreas científicas.

A continuación, se ofrece una lista, que pretende ser representativa pero no exhaustiva, de certámenes para investigadores jóvenes en España organizados por entidades públicas o privadas<sup>24</sup>:

- Feria Exporecerca Jove, organizada por Magma, Associació per Promoure la Recerca Jove (organización no gubernamental de Cataluña).
- Ciencia en Acción, concurso internacional dirigido a estudiantes, profesores, investigadores y divulgadores de la comunidad científica.
- Premios de investigación e innovación en ESO, Bachillerato y FP, organizado por la Consejería de Educación de Castilla y León.
- Congreso regional de investigadores junior, organizado por la Universidad de Murcia.
- Encuentro preuniversitario de jóvenes investigadores, organizado por la Universidad Complutense de Madrid.

<sup>23</sup> Enlace relacionado:

Libro de actas del XXX Congreso del Certamen Jóvenes Investigadores:

<http://www.injuve.es/conocenos/ediciones-injuve/actas-del-30-congreso-de-jovenes-investigadores>

<sup>24</sup> Enlaces relacionados:

- Exporecerca Jove: <http://www.magmarecerca.org/exporecerca/>

- Ciencia en Acción: <http://cienciaenaccion.org>

- Premios de investigación de Castilla y León:

<http://www.educacion.es/temas/calidad-evaluacion/excelencia-educativa/premios-investigacion-innovacion-bachillerato-formacion-pro>

- Congreso regional de investigadores, Universidad de Murcia:

<http://www.um.es/web/umusecundaria/contenido/actividades/ix-congreso-regional-de-investigadores-junior>

- Encuentro preuniversitario de jóvenes investigadores. Universidad Complutense de Madrid: <http://www.ucm.es/jovenesinvestigadores>

- SIMIP, Universidad Politécnica de Cartagena:

<http://www.upct.es/unidad-cultura-cientifica/es/0bachillerato-de-investigacion/proyectos-colaborativos-con-bachilleratos-de-investigacion>

- Premios a los mejores trabajos de investigación en Bachillerato, Universidad Pompeu Fabra: <http://www.upf.edu/es/web/graus/premios>

- Premio nacional de ciencia y tecnología para estudiantes de Bachillerato, Saint Louis University:

<http://www.slu.edu/madrid/academics/departments/science-engineering/science-technology-national-award.php>

- Jóvenes con investigadores, Sociedad Andaluza para la Divulgación de la Ciencia: <http://jovenesconinvestigadores.wordpress.com>

- Concurso para investigadores «Es de libro», CEDRO: <http://www.esdelibro.es/el-concurso/bases>

- Congreso internacional de jóvenes investigadores, Instituto de ciencia e investigación juvenil de Aragón: <http://www.icija.es>

- Certamen Tecnológico Efigy, Fundación Naturgy:

<https://www.fundacionnaturgy.org/certamen-tecnologico-efigy-de-fundacion-naturgy-2/>

- SIMIP, organizado por la Universidad Politécnica de Cartagena.
- Premios a los mejores trabajos de investigación en Bachillerato, organizado por la Universidad Pompeu Fabra.
- Premio nacional de ciencia y tecnología para estudiantes de Bachillerato, organizado por la Saint Louis University.
- Jóvenes con investigadores, organizado por la Sociedad Andaluza para la Divulgación de la Ciencia.
- Concurso para investigadores «Es de libro», organizado por el Centro Español de Derechos Reprográficos (CEDRO).
- Congreso internacional de jóvenes investigadores, organizado por el Instituto de ciencia e investigación juvenil de Aragón.
- Certamen Tecnológico Efigy, organizado por la Fundación Naturgy.

# Capítulo 19.

## La memoria de la investigación

En este capítulo se describen los principales contenidos de la memoria de una investigación ya concluida. En la actividad investigadora real esa memoria a menudo toma la forma de artículo científico, que puede ser publicado en una revista especializada o incluido en un libro técnico. Aquí el interés práctico se centra en la memoria que se presenta a un certamen para jóvenes investigadores, pero generalmente su estructura y contenido son muy similares a los de un artículo científico, y por tanto se describirán conjuntamente, haciendo notar las diferencias allí donde las hubiera.

Los apartados que se tratarán son: portada e índice, resumen (*abstract*), introducción, antecedentes (*state-of-the-art*), hipótesis y objetivos, materiales y métodos, resultados, conclusiones, agradecimientos, bibliografía y webgrafía, y anexo de figuras.



Licencia: dominio público. Fuente: Pixabay.

## 19.1. Portada e índice

El primer elemento de la memoria es la portada, en la que debe constar el título del trabajo de investigación, los nombres completos de todos los autores y del tutor, el centro docente o asociación juvenil al que pertenecen, el curso en el que están matriculados los autores y la fecha de presentación (año académico).

En un artículo científico no suele haber una portada como tal, pero una información muy similar figura en su encabezado, habitualmente con el siguiente orden: título de la investigación, nombres de todos los autores (debajo del título) y centro de investigación en el que realiza su labor cada uno de ellos (debajo de los nombres de los autores); para poder relacionar cada autor con su centro de investigación se suelen emplear superíndices numéricos o alfabéticos.

De entre toda la información que figura en el encabezamiento o portada, el título es la que requiere mayor reflexión. Evidentemente, ha de ser conciso y reflejar de manera concreta el tema de la investigación que se presenta. Se puede intentar redactar un título que atraiga la atención y despierte el interés del lector, especialmente en el ámbito de un certamen científico (con más cautela en un artículo de investigación real). Pero, para lograrlo, el título nunca debe dar a entender conclusiones o implicaciones de mayor alcance de lo que realmente contiene el trabajo realizado. En otras palabras, el título no puede ser engañoso, ni prometer más de lo que la investigación realizada ofrece. Si así se hiciera, se correría el riesgo de que la impresión global que se llevara el jurado sobre ese trabajo fuese la de decepción ante las expectativas no cumplidas. Como veremos más adelante, existen otros apartados de la memoria donde sí se pueden comentar las posibles implicaciones o aplicaciones futuras de la investigación, aunque no estén desarrolladas estrictamente en el trabajo realizado.

El segundo elemento de la memoria es un índice donde se recogen los apartados de los que consta y la página en la que comienza cada uno de ellos. Esos apartados son los mismos que estamos detallando aquí, aunque algunos pueden contener subapartados, que pueden aparecer en el índice diferenciándolos adecuadamente de los apartados principales (por ejemplo, dentro del apartado 3 pueden incluirse subapartados 3.a, 3.b, 3.c, etc. o bien 3.1, 3.2, 3.3, etc.; pueden incluirse también «subsubapartados», si ayudase a la organización de la memoria, como 3.1.1, 3.1.2, etc., pero sin llevar la subdivisión a un extremo que interrumpa la redacción fluida del texto).

Un índice de este tipo aparece al inicio de algunos artículos de investigación extensos, como los *reviews* o *reports* (artículos en los que se aporta una visión actualizada del conocimiento adquirido hasta la fecha en una determinada línea de investigación). En los artículos más breves, en cambio, no aparece un índice con paginación, pero suele incluirse un párrafo al inicio enumerando los apartados que conforman el artículo, para prevenir al lector sobre cómo está organizada la información.



Licencia: dominio público. Fuente: Pixabay.

## 19.2. Resumen

El primer apartado de contenido propiamente dicho es el resumen del trabajo de investigación, que ha de ser breve. En los artículos científicos se le denomina también con el término inglés *abstract*. En el Certamen Jóvenes Investigadores se establece un límite de 250 palabras para este resumen.

Al igual que ocurría con el título, el resumen debe reflejar fielmente la investigación realizada. Se trata de adelantar las ideas principales de algunos de los apartados que aparecen en la memoria que sigue al resumen. Es conveniente elaborar el resumen en último lugar, una vez que se tiene terminada la memoria, ya que al empezar por el resumen se corre el riesgo de incluir aspectos que luego no se traten en el resto de la memoria, o de omitir otros que sí aparecen y son muy relevantes.

No es habitual incluir en el resumen referencias a otros trabajos, a menos que la investigación realizada se apoye muy específicamente en otra anterior, o consista en confirmarla o refutarla.

El objetivo del *abstract* en los artículos científicos es permitir a los demás investigadores saber si el contenido del artículo es de su interés para el trabajo que están llevando a cabo. El título es a menudo demasiado breve como para transmitir detalles de la metodología empleada o las conclusiones extraídas, pero el *abstract* sí puede, y debe, aportar de manera concisa esa información. Tras la lectura del *abstract*, el investigador puede decidir si dedicar tiempo al artículo completo (algo que a veces requiere un gran esfuerzo), o bien descartarlo y buscar otros más útiles para su línea de investigación.

En el caso del Certamen Jóvenes Investigadores, o de cualquier otro, el jurado va a leer la memoria completa independientemente del contenido del resumen. Su utilidad aquí es otra: preparar al lector para lo que viene a continuación ayudándole a crear (junto con el índice) un esquema mental previo. Además, contribuye a la repetición de las ideas principales de la investigación, algunas de las cuales pueden aparecer en varias ocasiones (por ejemplo, en el resumen, en los resultados y en las conclusiones). Aunque en ciertos contextos la repetición puede resultar pesada, en un trabajo de investigación complejo puede ser de gran ayuda. No hay que olvidar que el lector puede no estar familiarizado estrictamente con el tema, que puede además ser muy abstracto y estar explicado con mucho detalle, por lo que repetir las ideas, de la manera más clara posible, facilita la asimilación.

En un artículo científico el resumen está redactado en un lenguaje igual de técnico que el resto del documento. Para la memoria del Certamen, quizá sea conveniente que el resumen sea algo más accesible, para que no solo el jurado, sino también los otros jóvenes participantes y el público en general, entiendan bien su contenido.

### 19.3. Introducción

Este apartado de la memoria permite acercarse de manera gradual al tema de la investigación, partiendo de un conocimiento general hasta llegar al área concreta de estudio. Puede servir para recopilar algunos datos y terminología generales del campo científico en el que se enmarca la investigación, que pueden ser bien conocidos, pero que conviene recordar. Por ejemplo, si se ha llevado a cabo una investigación sobre la geología de una zona, en la introducción se pueden recordar brevemente los procesos geológicos o los tipos de rocas, para refrescar la memoria de algo que en realidad la mayor parte de los lectores habrán estudiado en algún momento de su vida.



Fuente: elaboración propia con Genially.

Lo anterior es válido para la memoria del Certamen Jóvenes Investigadores, cuya introducción puede comenzar repasando conceptos básicos, como son los propios del nivel al que pertenecen los estudiantes, la Enseñanza Secundaria. En muchos artículos científicos, sin embargo, la propia introducción puede ser ya bastante técnica, y por tanto no muy accesible a investigadores ajenos al campo de la investigación.

La introducción, por tanto, ha de enfocarse desde el conocimiento más general hasta el más específico, o, si se prefiere, desde el más sencillo al más complejo, pero sin llegar al detalle concreto de la investigación, que se desarrollará en los apartados posteriores. Una de las claves de la introducción es su carácter más divulgativo y su capacidad de relacionar el área de conocimiento a la que pertenece la investigación con otras diferentes.

Otro de los contenidos fundamentales de este apartado introductorio es el de la motivación, es decir, las razones por las que la investigación realizada puede ser relevante. Estas razones pueden ser de tipo fundamental, para aclarar, confirmar o refutar teorías de ciencia básica, o de tipo práctico, para desarrollar tecnología, políticas, planes económicos, etc., en función del área de conocimiento de la investigación.

Esta parte de la introducción responde a la pregunta de por qué o para qué es interesante o relevante la investigación realizada. Debe reflejar las razones que han empujado a los investigadores a llevar a cabo su trabajo, pero también pueden aportarse implicaciones del trabajo realizado que se hayan encontrado *a posteriori*, conforme se profundizaba en el tema. Es decir, se puede comenzar la investigación por unas razones e intereses y acabar con otros, que pueden ser incluso más importantes que los originales. Las motivaciones de un estudio pueden expresarse muy a menudo en forma de problema, de pregunta, de inconsistencia o de inquietud, es decir, de todo aquello que despierta la curiosidad y nos mueve a buscar una respuesta o solución.

A la hora de enumerar las posibles aplicaciones o implicaciones futuras de la investigación realizada, es necesario ser imaginativo, sin que ello tenga ningún matiz peyorativo: la imaginación es un factor esencial en la actividad científica básica, y más aún en el desarrollo de aplicaciones.

#### 19.4. Antecedentes

En la sección de antecedentes, también llamada estado de la cuestión (en inglés, *state-of-the-art*, que se puede traducir como estado del arte), se recogen los últimos resultados científicamente establecidos sobre el tema de la investigación, tanto teóricos como experimentales. En el caso de los primeros, se deben incluir las interpretaciones teóricas propuestas más recientemente para el tema de la investigación, concretando el marco teórico general que se comenzó a exponer en el apartado de introducción.

Una de las principales razones para llevar a cabo esta labor, y reflejarla en la memoria, es asegurarse de que la investigación que se plantea no ha sido llevada a cabo con anterioridad. En el caso de que ya haya sido realizada, puede resultar interesante repetirla con otra metodología o en un contexto diferente, para ampliar el ámbito de aplicación de las conclusiones; en estas situaciones, será necesario comparar los resultados del nuevo estudio con los anteriores.

Es habitual que en este apartado se mencionen otros trabajos científicos previos, porque eso es precisamente lo que se está describiendo. Cada vez que se menciona un estudio es imprescindible citarlo, es decir, dar las indicaciones que permiten localizar el lugar en el que se encuentra publicado, que puede ser un libro, una revista, un sitio web, etc. El procedimiento habitual consiste en escribir un número o código entre paréntesis, corchetes o como superíndice cada vez que se menciona una investigación o sus resultados; cada número o código se asocia posteriormente a las indicaciones para encontrarlo, pero eso se hace en otra sección de la memoria, la de bibliografía o referencias, que se describirá posteriormente.

La ausencia de un apartado de antecedentes significaría que esa investigación es absolutamente pionera, lo que es enormemente improbable. Por tanto, este apartado es generalmente imprescindible.

ble, y además sirve como guía para que el lector, si lo considera necesario, acuda a las referencias para profundizar o comprender mejor el tema de la investigación.

El trabajo necesario para elaborar este apartado se engloba en lo que se denomina investigación secundaria, es decir, que recurre a investigaciones anteriores realizadas por otras personas, y consiste en este caso en la recopilación de fuentes documentales en forma de artículos de investigación, libros, etc. A este proceso se dedicará un capítulo posterior.



Licencia: dominio público. Fuente: Pixabay.

### 19.5. Hipótesis y objetivos

Como vimos en la parte I, el método científico parte de una o varias hipótesis generales de las que se pueden deducir consecuencias concretas susceptibles de ser comparadas con los datos obtenidos en la investigación, ya sean observaciones del entorno, resultados de experimentos, recopilación de información (por ejemplo, de fuentes históricas), etc. Si las consecuencias deducidas de la hipótesis no coinciden con los datos recogidos, la hipótesis queda falsada y ha de ser rechazada. A la hora de extraer las conclusiones de la investigación, que se reflejan en otro apartado de la memoria, será muy importante recordar lo que ya se indicó en la parte I: la falsación de una hipótesis, incluso en un solo caso (debidamente comprobado), resulta en su rechazo inmediato, pero su verificación nunca permite aceptarla de manera definitiva.

No conviene que las hipótesis de partida sean muy numerosas, porque resultaría una investigación demasiado amplia. A menudo, una sola hipótesis novedosa, profunda y relevante es suficiente. Contrastar varias hipótesis distintas exige recoger y analizar gran cantidad de datos, que puede ser una tarea inabarcable, especialmente en el contexto del Certamen Jóvenes Investigadores. Las hipótesis seleccionadas deben ser concretas y precisas, para facilitar su contraste con los datos recogidos.

Además de reflejar las hipótesis de investigación, o hipótesis de trabajo, en este apartado también han de enumerarse los objetivos de la investigación, es decir, los fines o metas que se pretenden alcanzar. El objetivo general de cualquier investigación, según el método científico, es contrastar hipótesis con los datos recogidos. De manera natural, por tanto, puede asociarse un objetivo a cada una de las hipótesis de trabajo.

Resulta adecuado enumerar también objetivos intermedios, a menudo de tipo instrumental. Por ejemplo, para lograr el objetivo principal de contrastar una de las hipótesis de partida, puede establecerse como objetivo intermedio construir un prototipo o un aparato de medida, elaborar una encuesta, establecer un protocolo, etc. Normalmente ninguno de esos procedimientos representa el objetivo final de la investigación, que debe ser el contraste de una hipótesis, pero pueden resultar imprescindibles para lograrlo. En cualquier caso, enriquecen mucho el trabajo realizado, porque producen una serie de instrumentos que pueden servir para otras investigaciones.



Licencia: dominio público. Fuente: Pixabay.

En conclusión, en este apartado se enumeran las hipótesis de trabajo, preferiblemente un conjunto reducido, y se asocian a objetivos. Además, se incluyen objetivos intermedios relacionados, por ejemplo, con la contextualización del tema de investigación, la elaboración de materiales y métodos o las aplicaciones de los resultados, siempre que tengan carácter novedoso.

## 19.6. Materiales y métodos

En este apartado debe describirse en detalle la metodología concreta empleada para recoger los datos, analizarlos y contrastar las hipótesis. La mayoría de estos procedimientos van asociados a materiales o instrumentos de muy diversos tipos: prototipos, aparatos de medida, cuestionarios o encuestas, bases de datos, programas informáticos de análisis estadístico y representación de datos, etc. Será necesario mencionar todos ellos y explicar su función y el modo en que se han empleado en la investigación. Siempre habrá que detenerse más en los más especializados, y que el lector pueda no conocer previamente, y sobre todo en aquellos que hayan sido creados o diseñados *ex profeso* para la investigación en marcha. La creación de materiales y métodos novedosos no siempre es el objetivo último de una investigación (aunque a veces pueda serlo), pero en todos los casos aportan un gran valor y es necesario ponerlo de relieve en la memoria.

En la parte I se introdujeron los principios generales del método científico, que deben quedar reflejados de manera clara a lo largo de toda la memoria, como se hace al enumerar las hipótesis de trabajo en el apartado anterior. Pero aquí ha de describirse el método científico específico para el área del estudio en cuestión, ya que las metodologías son muy diferentes para investigaciones en matemáticas, física, química, biología, geología, lingüística, filología, historia, arte, literatura, etc.

El grado de detalle en la descripción de materiales y métodos tiene que ser suficiente como para que una persona que lea la memoria sea capaz de reproducir la investigación y comprobar si se obtienen los mismos resultados. La reproducibilidad es una característica esencial de cualquier investigación científica, y es imposible alcanzarla si no se proporciona la información completa de cómo se ha llevado a cabo.

Opcionalmente, como parte de la metodología, se puede aportar un cronograma del desarrollo de la investigación, si se consideran relevantes los tiempos en que se lleva a cabo cada fase. Así puede ocurrir con la fase de toma de medidas de sistemas físicos, químicos o biológicos, en la aplicación de test a individuos, en la realización de encuestas, etc. Puede resultar práctico mostrar el cronograma en forma de tabla o empleando un diagrama de Gantt. Este último puede resultar útil también como documento interno para planificar el desarrollo de la investigación, incluyendo el posible reparto de tareas entre los miembros del equipo, pero en este caso quizá no sea necesario reflejarlo en la memoria.



Licencia: dominio público. Fuente: Pixabay.

## 19.7. Resultados

En este apartado de la memoria se consignan los datos recogidos en la investigación, ya sean cuantitativos o cualitativos.

Los datos cuantitativos, si no son una cantidad excesiva, pueden darse en forma de tabla, y si son muchos, quizás sea más conveniente representarlos gráficamente, de alguna de las maneras descritas en la parte III. Esto último es interesante en cualquier caso, porque transmite de una forma muy visual toda la información numérica recopilada.

Los resultados numéricos de las medidas, tanto directas como indirectas, deben expresarse correctamente, del modo en que se explicó en la parte III: valor e incertidumbre, ambos con la cantidad correcta de cifras significativas, y las unidades.

Junto a los datos brutos será necesario aportar los valores de los estadísticos calculados, como por ejemplo medidas de centralización, dispersión o correlación. Y finalmente, si se ha llevado a cabo un análisis de estadística inferencial, será necesario dar los intervalos de confianza para los parámetros poblacionales junto con su probabilidad (nivel de confianza), o el resultado del contraste de hipótesis estadística, también junto a su probabilidad de error.

En investigaciones con datos no cuantitativos, la información recogida suele ser muy extensa (por ejemplo, en el caso de entrevistas en profundidad), por lo que será necesario resumirla, seleccionar fragmentos representativos o elaborar un esquema o cuadro sinóptico de ideas principales.



Licencia: dominio público. Fuente: Pixabay.

## 19.8. Conclusiones

Las conclusiones más importantes de un trabajo de investigación son, evidentemente, las que se refieren a la validación o refutación de las hipótesis establecidas inicialmente. Aunque esta información ya se haya expresado en la sección de resultados, aquí se repetirá de forma concisa.

Este apartado puede iniciarse con una breve introducción al tema de la investigación y un resumen de la metodología empleada. La razón es que, en un vistazo rápido, cuando no interesan todos los detalles, se suele leer únicamente el resumen inicial o *abstract* y el apartado de conclusiones, y conviene que este último insista en los aspectos más relevantes de la investigación llevada a cabo. En el caso del Certamen Jóvenes Investigadores, evidentemente todos los trabajos van a ser leídos por completo, pero nunca está de más que las conclusiones vuelvan a resumir brevemente los aspectos más importantes de la investigación, para asentar la información o para refrescar la memoria.

Las conclusiones deben reflejar estrictamente lo que ha quedado científicamente demostrado en la investigación, y no tratar de ampliar su alcance o su relevancia injustificadamente.

## 19.9. Agradecimientos

En este apartado es necesario mencionar todas aquellas personas e instituciones que, sin formar parte del equipo de la investigación, hayan contribuido a su desarrollo de algún modo. Esa contribución puede haber sido directa, ayudando al equipo investigador en algún aspecto concreto, o indirecta, por ejemplo dando permiso para el uso de instalaciones o materiales de una institución.

Ya dijimos en otro capítulo que la realización de una investigación científica por parte de un grupo de estudiantes puede convertirse en un verdadero proyecto de centro, donde colaboren un gran número de profesores además del tutor, y quizá también con la participación secundaria de más estudiantes, aparte de los que forman el equipo. La contribución de todos ellos debe reflejarse en el apartado de agradecimientos.

En los artículos científicos se incluyen en este apartado, además de lo anterior, las instituciones o entidades que han financiado económicamente el trabajo de manera directa, a través, por ejemplo, de proyectos de investigación o de becas para investigadores. Además, cada vez más artículos incluyen

también los posibles conflictos de intereses, o una declaración de su ausencia. Un ejemplo de posible conflicto de intereses aparece cuando una investigación sobre la eficacia de un nuevo fármaco ha sido financiada, total o parcialmente, por la compañía farmacéutica que lo desarrolla, o si esta ha subvencionado de alguna manera cualquier tipo de actividad de los autores del trabajo.

En la memoria del Certamen Jóvenes Investigadores es poco probable que haya que mencionar fuentes de financiación del estudio y los posibles conflictos de intereses asociados. Pero, desde luego, si la investigación ha recibido cualquier tipo de apoyo económico, es necesario reflejarlo aquí.

## 19.10. Bibliografía y webgrafía

Este apartado recoge las fuentes de las que se ha tomado todo el contenido de la memoria de la investigación que no es original, es decir, que no es fruto directo del trabajo realizado por los autores. Esas fuentes se denominan en general referencias, y a veces también bibliografía, aunque este último término se refiere estrictamente a material escrito. En realidad, la variedad de formatos y soportes de las fuentes utilizadas puede ser muy amplia, incluyendo fotografías, vídeos, audios, etc., además del material escrito (libros, revistas, periódicos, leyes y reglamentos...). El acceso a estos otros tipos de fuentes es cada vez más sencillo y habitual gracias a internet, donde el material multimedia se encuentra alojado en páginas web. De ahí que este apartado se denomine conjuntamente «bibliografía y webgrafía».

Es necesario citar la fuente cuando se reproduce literalmente lo expresado por otros autores, pero también cuando se hace uso de una idea ajena, aunque esté escrita con otras palabras. La única excepción es aquella información que es ampliamente conocida y cuyo origen no puede atribuirse a personas o fuentes concretas. El objetivo de las citas suele ser doble: por un lado, permiten atribuir las ideas a sus autores legítimos, y por otro lado, facilitan al lector del texto la fuente original de esas ideas, de modo que pueda acudir a ella para obtener más detalles, interpretarla de distinta manera o extraer diferentes conclusiones. El trabajo propio también se ve reforzado al citar otros trabajos, porque aleja toda duda de plagio y porque sustenta las nuevas ideas en otras anteriores ya consolidadas científicamente.

Para la memoria del Certamen Jóvenes Investigadores se requiere emplear el estilo de citación de Harvard, también llamado de autor-fecha, que se describe en detalle en un capítulo posterior. Existen muchos estilos distintos de citación, algunos más típicos de las ciencias sociales y otros más habituales en ciencias naturales. De hecho, las editoriales pueden exigir el uso de un estilo propio en sus publicaciones, por ejemplo, en los artículos de investigación para sus revistas. Si los autores del trabajo pueden escoger qué estilo de citación usar, es necesario mantener el mismo durante todo el trabajo.



Licencia: dominio público. Fuente: Pixabay.

## 19.11. Anexo de figuras

En los artículos y libros de investigación, las figuras suelen ir intercaladas en el texto. En la memoria del Certamen Jóvenes Investigadores, en cambio, se reserva un apartado final, con una extensión máxima de 10 páginas, para incluir tablas, diagramas, gráficos, dibujos, fotografías o cualquier otro elemento

de tipo gráfico que complemente la información aportada en los apartados de texto previos. De esta manera, las 10 páginas de extensión máxima de la memoria, que no incluyen el anexo, pueden aprovecharse íntegramente para el texto.

El apartado de la memoria en el que las figuras pueden resultar más necesarias es el de resultados, pero también pueden ir asociadas a los apartados de introducción, antecedentes y materiales y métodos, en este último caso en forma de diagramas de flujo o cronogramas sobre la metodología, esquemas o fotografías de los dispositivos experimentales empleados, etc. Ninguna de las figuras debe estar presente como mera decoración, sino que deben tener una finalidad específica e ir asociadas a pasajes concretos del texto de la memoria.

Las figuras deben ir numeradas y acompañadas de un pie en el que se explique brevemente su contenido. En cualquier caso, toda figura debe estar explicada en detalle en el texto de la memoria, donde se hará referencia a cada una de ellas (por ejemplo, con la frase «ver figura X en el anexo» o «como se observa en la figura X del anexo», donde X es su numeración). En este anexo no debe aparecer ninguna figura que no esté explicada y referenciada en el texto, entre otras razones porque sería muy fácil saltársela y además estaría completamente descontextualizada.

En la memoria de una investigación es especialmente cierto aquello de que las imágenes valen más que mil palabras, pero para ello es necesario que sean suficientemente claras. A ello a menudo contribuye que tengan un tamaño, un contraste y una resolución adecuados. En las fotografías deben observarse nítidamente aquellos elementos relevantes. En las gráficas deben leerse claramente las etiquetas de los ejes, la numeración de estos y las unidades asociadas a esa numeración, así como el significado de cada serie de datos, que vendrán recogidos en una leyenda (o también en el pie de la figura).



Licencia: dominio público. Fuente: Pixabay.

### AMPLIACIÓN. Tipografía y edición de la memoria

Una gran parte de los textos científicos, como son memorias de investigación, artículos, libros, tesis doctorales, etc., se componen actualmente con el sistema LaTeX, creado por el matemático estadounidense Leslie Lamport en 1984, que contiene paquetes o preformatos basados a su vez en el sistema TeX, ideado por el también matemático estadounidense Donald Knuth. El sistema LaTeX ofrece una alta calidad tipográfica, de tipo profesional, y además es muy flexible en cuanto a la creación de expresiones matemáticas de muy diversos tipos.

El *software* necesario para componer documentos en LaTeX es libre, y puede descargarse desde latex-project.org para ser instalado en diversos sistemas operativos, o bien puede usarse en línea.

La mayoría de los artículos científicos que se pueden encontrar en internet (ya sean libres o de pago) están escritos con LaTeX. Aunque requiere un pequeño esfuerzo inicial, merece la pena que tanto profesores como alumnos se familiaricen con LaTeX, y que, llegado el momento, lo usen para producir sus documentos científicos y académicos.



Licencia: dominio público. Fuente: Wikimedia Commons.

# Capítulo 20.

## Fuentes documentales

Las fuentes documentales recogen datos o medidas tomados por otras personas, o pueden consistir en investigaciones completas ya publicadas en forma de artículos o libros. Puede resultar necesario acudir a ellas al realizar una nueva investigación en casos como los siguientes:

- Para obtener datos de una realidad que ya ha pasado (fuentes históricas).
- Para elaborar el estado de la cuestión sobre un tema y ajustar así el diseño y los objetivos de una nueva investigación sobre el mismo.
- Para ampliar las conclusiones de una investigación mediante comparaciones entre los nuevos datos recogidos (datos primarios) y los extraídos de fuentes documentales, por ejemplo, para comparar los resultados con otras regiones geográficas, otras sociedades, culturas o grupos de población, otras épocas, etc.
- Para seleccionar correctamente los elementos de una muestra a partir de fuentes documentales relativas a la población completa, por ejemplo, empleando los datos del censo para extraer la muestra del estudio.

El uso de fuentes documentales por parte de un investigador, en contraste con los datos que recoge él mismo, puede presentar una serie de problemas. En ocasiones puede no haber información suficiente sobre su calidad, en relación con posibles fuentes de error o con sesgos en las muestras. También puede ocurrir que las variables medidas en otro estudio no estén definidas exactamente como las del actual, lo que dificulta la combinación de los datos o incluso su simple comparación. Esto último es muy común en investigaciones realizadas en épocas muy distintas, pero cada vez es más raro en las llevadas a cabo en distintos países, por el creciente nivel de estandarización de la actividad científica. Cabe señalar que este último problema es más típico de las investigaciones en ciencias sociales, donde existe más arbitrariedad en la definición de conceptos y sus variables asociadas, y donde cobra más relevancia el contexto histórico y sociocultural.

Para identificar y, en la medida de lo posible, mitigar los problemas asociados al uso de fuentes documentales, es necesario someter estas a un estricto control para determinar su autenticidad, credibilidad (autoridad y prestigio de las personas o entidades que han elaborado los documentos), rigurosidad y representatividad (que puede deducirse de los errores estadísticos, si se incluyen). Es necesario conocer también los protocolos empleados para la recogida de datos, la definición de las variables empleadas, cuándo y dónde se recogieron, el tamaño de las muestras, el tipo de análisis realizado, etc. Si toda esta información no está expresada en la fuente y no puede deducirse de su contenido o averiguarse fácilmente por otros medios, es posible que no debamos utilizar los datos de esa fuente en la nueva investigación, o que tengamos que hacerlo con mucha cautela.

Las fuentes documentales pueden consistir en recopilaciones de datos (fuentes de datos brutos), como por ejemplo un censo poblacional, o bien en investigaciones completas (fuentes de datos elaborados), por ejemplo, un artículo que analiza las causas y consecuencias de la evolución de la pirámide poblacional de un país. En este último caso, los datos brutos originales pueden no estar contenidos en la propia fuente, y habrá que buscarlos en otra, si fueran necesarios.



Licencia: dominio público. Fuente: Pixabay.

## 20.1. Fuentes de datos brutos

En las ciencias naturales, los conjuntos de datos, que son generalmente medidas experimentales, se encuentran a menudo publicadas en artículos científicos o en recopilaciones como el *CRC Handbook of Chemistry and Physics* (editorial CRC Press), un grueso volumen que reúne una amplísima colección de datos físicos, químicos, bioquímicos e incluso algunos de biología, geología y astronomía. Si bien muchos libros y artículos son de pago, existen en internet bases de datos científicos y tecnológicos a las que se puede acceder libremente. Suelen estar alojadas en sitios web de centros u organismos de investigación, especialmente de carácter público, como por ejemplo los laboratorios e institutos nacionales estadounidenses. A modo de ejemplo se pueden citar las bases de datos del Instituto Nacional de Estándares y Tecnología (*National Institute of Standards and Technology*, NIST), del Centro Nacional de Datos Nucleares (*National Nuclear Data Center*, NNDC, perteneciente al Brookhaven National Laboratory) o del Centro de Control y Prevención de Enfermedades (*Center for Disease Control and Prevention*, CDC, cuyos datos sobre temas de salud se refieren normalmente a EE. UU.)<sup>25</sup>.

En el caso de las ciencias sociales, ciertos datos se encuentran publicados en artículos de investigación, como ocurre en el área de ciencias naturales. Pero también existen amplias y variadas bases de datos que elaboran, por ejemplo, los organismos públicos de estadística. El principal en España es el Instituto Nacional de Estadística (INE), organismo autónomo adscrito al Ministerio de Economía, que reúne datos de muy diversas áreas. Algunas comunidades autónomas cuentan también con organismos propios de estudios estadísticos, y un organismo similar al INE, pero a nivel europeo, es la Oficina Europea de Estadística (Eurostat)<sup>26</sup>.

<sup>25</sup> Enlaces relacionados:

- National Institute of Standards and Technology (NIST): <http://data.nist.gov/sdp/>
- National Nuclear Data Center (NNDC): <http://www.nndc.bnl.gov>
- Center for Disease Control and Prevention (CDC): <http://www.cdc.gov/DataStatistics/>

<sup>26</sup> Enlaces relacionados:

- Instituto Nacional de Estadística (INE): <http://www.ine.es>
- Oficina Europea de Estadística (Eurostat): <http://ec.europa.eu/eurostat/>

El servicio de datos estadísticos mundiales de la ONU (UNdata) ofrece información por países sobre población y migraciones, desarrollo, economía, educación, salud, género, agricultura, energía, medio ambiente, ciencia y tecnología, transporte, comunicaciones, comercio, turismo, etc. La ONU ofrece otras amplias bases de datos en el boletín mensual de estadísticas (MBS), y más específicas sobre objetivos de desarrollo sostenible y sobre comercio global. Otro organismo internacional que ofrece datos de carácter económico es el Banco Mundial<sup>27</sup>.

Our World in Data es un proyecto de la Universidad de Oxford (Reino Unido), financiado por fundaciones y por donaciones individuales, que contiene una base de datos a nivel mundial sobre población, salud, alimentación, energía, medio ambiente, tecnología, crecimiento y desigualdad, política, violencia, guerras, derechos, educación, medios de comunicación, cultura, etc. Los datos aparecen representados en diversos tipos de gráficos y en mapas. También resulta interesante la base de datos sobre los países del mundo que actualiza continuamente la Agencia Central de Inteligencia de los Estados Unidos (CIA), llamada The World Factbook, y que (curiosamente) es de acceso libre<sup>28</sup>.

## 20.2. Fuentes de datos elaborados

Las investigaciones ya completadas, que son las fuentes de datos elaborados, se encuentran generalmente publicadas en libros y artículos científicos. Los primeros, a menudo de carácter técnico, se pueden encontrar más fácilmente en bibliotecas de universidades y centros de investigación. También es cada vez más sencillo, aunque no necesariamente barato, adquirir esos libros a través de internet.

Algo similar ocurre con los artículos de revistas científicas. Aunque ya apenas circulan en formato impreso, sigue siendo necesario acudir a una biblioteca de universidad o de centro de investigación para, desde su red, acceder por internet a la versión digital. La razón es que esas instituciones pagan una cuota anual a las diversas editoriales para tener acceso completo a sus revistas. Si se carece de suscripción, todavía es posible acceder a los artículos individualmente, pero pagando por cada uno de ellos un precio que suele ser bastante alto. Si la investigación que se está llevando a cabo cuenta con un colaborador o contacto en alguna institución con suscripción, él puede obtener los artículos necesarios.

Cabe preguntarse por qué los investigadores envían sus artículos a publicar a revistas que luego cobran por la consulta, pudiendo poner sus trabajos en internet a libre disposición de cualquier persona interesada. En primer lugar, es necesario aclarar que los investigadores no cobran por la publicación de los artículos en esas revistas. A veces, incluso, ellos o las instituciones a las que pertenecen tienen que pagar por publicar. El principal motivo de publicar en revistas de pago es que los artículos enviados son sometidos a una revisión por pares, que significa que especialistas del área del trabajo presentado leen el artículo y sugieren o exigen mejoras antes de dar su visto bueno; normalmente son entre uno y tres revisores, que permanecen en el anonimato. El comité editorial de la revista es el encargado de organizar el proceso de revisión y de tomar la decisión final sobre la publicación del trabajo. Este mecanismo confiere mayor fiabilidad y garantía de calidad a los artículos publicados en esas revistas. Además, los artículos son editados y maquetados por profesionales de la editorial.

No obstante, es muy habitual que los propios investigadores suban a internet sus artículos en una versión muy parecida a la que ha sido publicada en una revista. A menudo, se diferencian únicamente en la maquetación y en pequeños detalles de edición. Un sitio web muy conocido que recopila

---

<sup>27</sup> Enlaces relacionados:

- Servicio de datos estadísticos mundiales de la ONU (UNdata): <http://data.un.org>  
- Boletín mensual de estadísticas (MBS) de la ONU: <http://unstats.un.org/unsd/mbs/>  
- Banco Mundial: <http://data.worldbank.org>

<sup>28</sup> Enlaces relacionados:

- Our World in Data (Oxford Univ.): <http://ourworldindata.org>  
- The World Factbook de la CIA: <https://www.cia.gov/library/publications/resources/the-world-factbook/>

estas versiones de libre acceso de los artículos es arXiv, que es gestionado por la biblioteca de la Universidad de Cornell (EE. UU.). Actualmente contiene cerca de dos millones de artículos (llamados *e-prints*) sobre física, matemáticas y estadística, computación, biología, ingeniería eléctrica y de sistemas y economía. Se pueden buscar artículos a partir del nombre de uno o varios de los autores, de palabras clave del título o del resumen (*abstract*), de diversos códigos de identificación del artículo, del área científica o de la fecha de publicación<sup>29</sup>.

Muchos autores publican también versiones de sus trabajos en el sitio web de la institución a la que pertenecen o en su web personal. Además, cada vez hay más revistas que publican en régimen de acceso libre (*open access*), es decir, que permiten la consulta gratuita de sus artículos.

Si se busca un artículo concreto y se conoce su referencia bibliográfica, puede introducirse en un buscador general de internet, como Google, y revisar la lista de enlaces que devuelve, porque en alguno de ellas puede encontrarse una versión de libre acceso. También puede acudir directamente a buscadores específicos de documentos científicos o académicos, como Google Scholar (Google Académico) o Crossref, en los que se puede buscar por referencia bibliográfica, por palabras del título o del resumen, por los nombres de los autores, etc.<sup>30</sup>

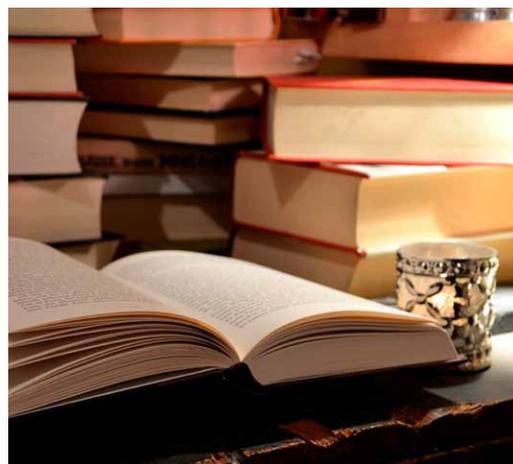
Los artículos publicados en versión digital vienen identificados por un código denominado DOI (*Digital Object Identifier*), que, si se conoce, puede servir para encontrarlo rápidamente en un buscador general o específico, o escribiendo directamente en el navegador la URL <http://dx.doi.org/código>, donde la palabra «código» ha de sustituirse por el DOI del artículo. Ejemplo de DOI para el artículo en el que se publicó el descubrimiento de la famosa partícula llamada bosón de Higgs en 2012: 10.1016/j.physletb.2012.08.021; por tanto, se accede a él directamente con la URL <http://dx.doi.org/10.1016/j.physletb.2012.08.021>.

En caso de no encontrar en internet el artículo buscado, siempre existe la opción de ponerse en contacto con alguno de los autores, preferentemente por correo electrónico (que a menudo se encuentra fácilmente introduciendo en un buscador el nombre y la institución a la que pertenece). Normalmente, los autores no tendrán inconveniente en enviar su artículo, y además, dependiendo del tiempo del que dispongan, puede que incluso puedan contestar preguntas sobre su trabajo. Quién sabe si, además, ese contacto puede ser el inicio de una colaboración científica en el futuro.

### 20.3. Cómo citar las fuentes

Citar una fuente significa aportar las indicaciones o referencias necesarias para localizar un documento, del tipo que sea, cuyo contenido ha sido empleado para desarrollar una investigación o para redactar la memoria final de la misma. Las citas pueden aparecer a lo largo del texto de la memoria o en forma de listado en un apartado específico que se suele denominar «Bibliografía» o «Referencias».

Existen diversos sistemas de citación, que organizan de manera distinta la información necesaria para localizar la fuente sin ambigüedad. Esa información incluye, entre otros datos, el nombre del autor o autores, el título del documento, el año de publicación, el libro o la revista en la que está incluido, etc. Aquí se describirá el sistema



Licencia: dominio público. Fuente: Pixabay.

<sup>29</sup> Enlaces relacionados:

arXiv: <http://arxiv.org>

<sup>30</sup> Enlaces relacionados:

Google Académico: <http://scholar.google.es>

Crossref: <http://search.crossref.org>

de citación de Harvard, que es el que se pide en las normas del Certamen Jóvenes Investigadores.

### **Citas en el texto**

Las citas a lo largo del texto deben incluirse en los siguientes casos:

- Cuando se está reproduciendo literalmente un fragmento de otra fuente. En este caso, además de citar la fuente, es necesario que el fragmento esté entrecomillado, o que vaya en párrafos aparte con mayor tabulación (menor longitud de los renglones que en el texto principal) o en cursiva.
- Cuando se está expresando una idea recogida por sus autores originales en otra fuente, aunque se haga con palabras distintas.

En el sistema de citación de Harvard las citas a lo largo del texto están abreviadas, ya que se volverán a incluir, pero con más datos, en el listado de referencias recogido en un apartado específico.

Las citas en este sistema se componen de los apellidos de los autores (hasta tres), del año de publicación del documento y, si se quiere precisar, del número de página (precedida de p.) o páginas (precedidas de pp.) en las que se encuentra la información específica usada. Si hay más de tres autores se usa solo el apellido del primero acompañado de *et al.*, que es la abreviación de la locución latina *et alii*, que significa «y otros». Estos componentes de la cita se separan por comas y se escriben entre paréntesis inmediatamente después de reproducir literalmente o de expresar una idea tomada de la fuente.

Ejemplos:

(Martínez, Carrasco y Tovar, 2014)  
(Carrasco Marín y Pérez Olmo, 2012, p. 5)  
(Horowitz *et al.*, 2016, pp. 126-128)

Si el apellido de los autores se menciona en el texto principal ya no se repite en el interior del paréntesis. Ejemplo: «Tal y como señalan Martínez, Carrasco y Tovar (2014)...». Si se citan varias obras de un mismo conjunto de autores publicadas en el mismo año, se distinguen entre sí añadiendo letras al año en orden alfabético. Ejemplo: (2012a, 2012b).

Algunos documentos tienen autoría corporativa o han sido elaborados por un conjunto de personas, por ejemplo empresas, organizaciones, comités, etc., y su nombre sustituye al de los apellidos de los autores. Si no se conoce el autor, se usa el término anónimo (o anon.) y si no se conoce la fecha de publicación se escribe n. d. (del inglés *no date*) o s. f. (sin fecha).

### **Citas listadas en la bibliografía**

Las fuentes empleadas deben citarse también en un apartado específico del documento en forma de lista ordenada alfabéticamente según el primer apellido del primer autor o del autor corporativo. Si se citan varias obras de un mismo primer autor con los mismos coautores y todas publicadas el mismo año, estas se ordenan por fecha; el año en cada una de ellas se acompaña de una letra minúscula distinta, en orden alfabético, que es como se indicaban también en las citas a lo largo del texto (por ejemplo, 2017a, 2017b, 2017c).

Las referencias en esta lista son completas, es decir, además de los autores y el año de publicación, incluyen todos los datos adicionales necesarios para localizar la fuente. La estructura y contenido de la referencia completa difiere ligeramente entre los diversos tipos de documento. Aquí se describirán los más usuales: libro completo, capítulo de un libro, artículo de revista, artículo de periódico, trabajo académico y página o sitio web de internet.

- Para citar un libro con el estilo de Harvard debe emplearse la siguiente estructura:

Apellido(s), inicial(es) de los autores. (Año de publicación) *Título*. Edición. Lugar de publicación: editorial.

Si hay varios autores, se separan mediante una coma, manteniendo en todos ellos la misma estructura: en primer lugar los apellidos, y tras una coma, las iniciales del nombre (cada una de ellas seguida de un punto). El último autor no se separa de los anteriores mediante una coma, sino mediante la conjunción «y». El año de publicación va a continuación, entre paréntesis. A continuación (sin punto intermedio), el título del libro se escribe en cursiva y acaba en punto. Si el libro que se ha consultado no es la primera edición, debe incluirse a continuación en forma abreviada (por ejemplo, 3ª ed.). La referencia termina con el lugar de publicación y, seguido de dos puntos, el nombre de la editorial.

Ejemplo:

Pérez López, J., García Martín, J. M. y González Díez, A. (2016) *Introducción a la biología*. 2ª ed. Barcelona: Editores Reunidos.

- Para citar un capítulo concreto dentro de un libro, la estructura es la siguiente:

Apellido(s), inicial(es) de los autores. (Año de publicación) «Título del capítulo». En: Apellido(s), inicial(es) de los editores del libro, seguido de ed. o eds. *Título del libro*. Edición. Lugar de publicación: editorial.

La diferencia con la cita de un libro completo es que aquí se empieza por los autores y el título del capítulo (entre comillas y sin cursiva), y a continuación se dice en qué libro se encuentra (comenzando por «En:»), indicando primero sus editores, es decir, las personas que coordinan a todos los autores de los capítulos, seguido de la abreviatura ed. (si es uno) o eds. (si hay varios).

Ejemplo:

García Barreiro, E., Benito Marín, L., González Díez, A. y Agudo Suárez, A. (2012) «Nuevas perspectivas en biología evolutiva». En: Asunción Rodríguez, P. y Anguita Romero, F. eds. *Los nuevos caminos de la ciencia biológica*. Barcelona: Libros Universitarios.

- Para citar artículos de revistas, la estructura es:

Apellido(s), inicial(es) de los autores. (Año de publicación) «Título del artículo», *Título de la revista*, volumen (número), p. o pp. página(s).

Las revistas que contienen artículos de investigación suelen organizarse en volúmenes y números, y por eso es necesario incluirlos en la cita. También se indica el intervalo de páginas que abarca el artículo, precedido de la abreviatura pp. (o p. si solo es una página).

Ejemplo:

Benito Marín, L. y García Barreiro, E. (2011) «El papel de la P-proteína en la producción de melatonina», *Avances en biología molecular*, 78 (2), pp. 127-138.

- Para citar un artículo de periódico impreso:

Apellido(s), inicial(es) de los autores. (Año de publicación) «Título del artículo», *Nombre del periódico*, fecha, p. o pp. página(s).

Ejemplo:

Fernández, M. (2019) «El nuevo plan del Gobierno para regular el autoconsumo eléctrico», *Diario Sur*, 2 de enero, p. 10.

- Para citar un trabajo académico:

Apellido(s), inicial(es) de los autores. (Año de presentación del trabajo académico) *Título del traba-*

jo. Tipo de trabajo (tesis, de fin de estudios...). Universidad en la que se presentó.

Un trabajo académico puede ser una tesis doctoral, un trabajo de fin de grado, de fin de licenciatura, de fin de máster, etc. Esta manera de citarlos es válida para aquellos que no han sido publicados. Si lo están, hay que seguir el formato de cita del medio en que se han publicado (usualmente como libro).

Ejemplo:

Cabo Antunes, J. L. (2018) *Impacto económico de los procesos de desamortización en España*. Tesis doctoral. Universidad Complutense de Madrid.

– Para citar una página o sitio web se procede del siguiente modo:

Apellido(s), inicial(es) de los autores. (Año de publicación) *Título de la página o sitio web* [Online]. Disponible en: URL (Consultado el «fecha»).

La URL es la dirección completa de internet en la que se puede encontrar la página web consultada, y a continuación se indica entre paréntesis la fecha de consulta (día, mes y año). En general, para citar una fuente (de cualquier tipo) que se ha consultado en internet habrá que incluir en la cita la dirección URL, en la forma «Disponible en: URL» y a continuación entre paréntesis la fecha de consulta, en la forma «(Consultado el «fecha»)».

Ejemplo:

Rodríguez, J. A. (2016) *La web de ecología* [Online]. Disponible en: <http://www.lawebdeecologia.com> (Consultado el 4 de enero de 2017).

Como se mencionó al inicio, existen diversos tipos de formato de citación o referencia de fuentes, además del de Harvard descrito con detalle aquí, pero es esencial mantener el mismo formato en todo el documento. Un formato típico de muchos artículos de investigación publicados en revistas especializadas consiste en asociar cada cita en el interior del texto con un número, que se escribe entre corchetes. Los números van ordenados, es decir, la primera cita es [1], la segunda cita es [2] y así sucesivamente. Si una misma fuente es citada varias veces, se usa en todas ellas el mismo número con el que se identificó la primera vez. Al final del artículo se incluye una sección (*References*) donde aparece la lista completa según el orden de cita (es decir, con los números ordenados), y donde cada uno de ellos se asocia a la referencia completa de la fuente. Como se puede observar, en este sistema no se da ninguna información sobre las fuentes a lo largo del texto, y hay que acudir a la sección de referencias cada vez que se quiera conocer una. En esta última sección solo se listan las referencias que han ido apareciendo en el texto y que tienen un número asociado.

En otro tipo de trabajos, como en la memoria que se ha de presentar al Certamen Jóvenes Investigadores, conviene que la lista final de referencias contenga no solo las fuentes citadas a lo largo del texto, sino también todo aquel documento que se haya consultado y haya servido en general para llevar a cabo el trabajo, aunque no esté citado específicamente en ningún punto de la memoria. Tanto las referencias específicas citadas en el texto como las generales se listan juntas en el apartado «Bibliografía y webgrafía» de la memoria, en orden alfabético según el apellido del primer autor.

## 20.4. La investigación secundaria

La investigación secundaria, que requiere el uso de fuentes documentales, consiste en utilizar datos o medidas recogidos por otros investigadores, en contraste con la primaria, que conlleva la recogida de sus propios datos.

Existen diversas razones para llevar a cabo una investigación exclusivamente secundaria. Por un lado, puede deberse a cuestiones prácticas: tiempo y recursos insuficientes para recoger datos propios,

o imposibilidad de hacerlo por referirse a una realidad pasada (caso de investigaciones históricas). Por otro lado, puede tratarse de investigaciones específicamente diseñadas para recopilar datos de una o más investigaciones previas, con dos objetivos principales:

- 1) Para someterlos a un análisis conjunto o metaanálisis, en el caso de emplear varias investigaciones. El objetivo es ampliar, profundizar y aumentar la validez de las conclusiones obtenidas previamente, ya que ahora se cuenta con un mayor número de datos que en cada una de las investigaciones anteriores por separado. Por ejemplo, en la literatura pueden encontrarse diversos estudios acerca del efecto del consumo de diferentes tipos de alimentos sobre el riesgo de desarrollar una cierta enfermedad. Un metaanálisis puede tratar de reunir todos esos datos y analizarlos en conjunto. Para llevar a cabo un metaanálisis es necesario seleccionar únicamente los estudios que superen ciertos umbrales de calidad, que sean compatibles entre sí y que estén relacionados con la hipótesis de trabajo establecida para el metaanálisis. Además, el análisis estadístico debe tener en cuenta las incertidumbres de los datos originales, que probablemente serán distintas en cada estudio.
- 2) Para llevar a cabo un análisis secundario, que puede consistir en usar los datos de una investigación anterior con un objetivo distinto (por ejemplo, para contrastar una hipótesis diferente), o para el mismo objetivo original, pero con distinta técnica de análisis, más sofisticada o más moderna que la anterior. También puede llevarse a cabo simplemente para comprobar que los análisis de estudios anteriores eran correctos. Recordemos que la reproducibilidad de las investigaciones en todas sus fases, incluida la de análisis de datos, es un requisito esencial en la ciencia.

Aunque pueden llevarse a cabo investigaciones exclusivamente secundarias, por los motivos que se acaban de exponer, en realidad estas aparecen también en las investigaciones primarias en la fase de elaboración del estado de la cuestión (*state-of-the-art*), que consiste en una revisión bibliográfica acerca de lo que ya se conoce sobre el tema de estudio. Esa fase se puede considerar por tanto como una investigación secundaria encajada dentro de la principal.

Además de para la revisión bibliográfica, los datos secundarios se pueden combinar con los nuevos datos recogidos en una investigación primaria para comparar los resultados obtenidos en otros lugares o en otras épocas (análisis históricos), que serían ejemplos de análisis secundarios como los descritos antes.



Licencia: dominio público. Fuente: Pixabay.

# Capítulo 21.

## Presentación pública de la investigación

Es habitual difundir públicamente los resultados de una investigación en reuniones científicas como congresos, conferencias, *workshops* (reuniones especializadas), etc. Una investigación realizada en un centro educativo puede difundirse también en un aula o salón de actos del propio centro, para darla a conocer a otros miembros de la comunidad educativa.

La presentación pública se lleva a cabo principalmente mediante la exhibición de un póster o cartel científico o a través de una exposición oral acompañada por la proyección de una presentación. Estas actividades complementan a la memoria de la investigación o al artículo científico, que normalmente son leídos solamente por personas especialmente interesadas y de manera individual. La presentación pública requiere unas habilidades específicas por parte de los autores de la investigación, que es necesario conocer y entrenar.

En el Certamen Jóvenes Investigadores de la Secretaría General de Universidades se estipula que los participantes que resulten finalistas, y que por tanto sean seleccionados para asistir al congreso del Certamen, deben preparar una exposición oral y un póster. En otros congresos científicos los trabajos se suelen presentar en uno o en otro formato, no necesariamente en ambos a la vez.



Licencia: dominio público. Fuente: Pixabay.

## 21.1. El póster científico

El póster o cartel científico resume el contenido de una investigación de una manera visual, donde las imágenes de cualquier tipo (fotografías, dibujos, gráficas, esquemas, etc.) adquieren aún más importancia que en la memoria escrita. Se expone en alguna sala del recinto del congreso y permite a todos los asistentes, incluyendo el jurado en el caso de un certamen, informarse con más calma acerca de la investigación presentada. Aunque el póster contiene solamente un resumen de la investigación, invita a entablar conversación con los autores para profundizar en el contenido y resolver dudas.

El póster tiene unas grandes dimensiones, precisamente porque su objetivo es atraer la atención desde cierta distancia y permitir que pueda ser leído por varias personas a la vez mientras conversan sobre su contenido entre ellas o con los autores. El tamaño habitual es el de DIN-A0 (118,9 cm de alto por 84,1 cm de ancho, es decir, aproximadamente cuatro veces el alto y el ancho de un folio o de una DIN-A4). En el Certamen Jóvenes Investigadores el póster ha de ser de ese tamaño, confeccionado en orientación vertical y con los textos redactados en castellano.

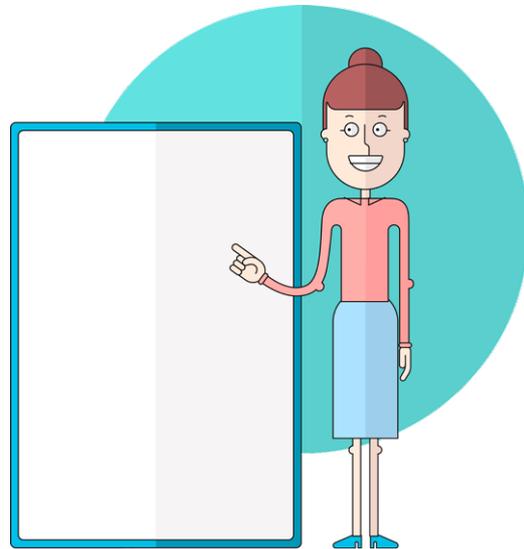
Se pueden encontrar en internet compilaciones de pósters científicos, como Eposter, que se pueden explorar para obtener ideas sobre el diseño. También son muy interesantes los ejemplos de pósters sencillos sobre grandes hitos de la ciencia, incluidos en el sitio web sobre la elaboración de pósters académicos de la Universidad de York<sup>31</sup>.

Existen diversas aplicaciones para confeccionarlos, como Word o Power Point de Microsoft, Pages o Keynote de Apple o las correspondientes versiones en código abierto de OpenOffice, entre otras. El tamaño del documento por defecto suele ser DIN-A4 en los procesadores de texto (Word y similares) y el de transparencia en los programas de presentaciones (Power Point y similares), por lo que el primer paso es cambiarlo al de DIN-A0 en orientación vertical (si no existe la opción predefinida se tendrán que introducir las dimensiones a mano, 118,9 x 84,1 cm). También existen aplicaciones en internet, con versión gratuita, que contienen diseños preestablecidos para pósters, como Piktochart o Canva<sup>32</sup>.

Es difícil trabajar visualizando el póster al completo en la pantalla del ordenador, porque su tamaño estará muy reducido, así que habrá que hacer uso de la función zoom e ir confeccionándolo por partes. De vez en cuando conviene llevar el zoom al tamaño real, 100 %, para comprobar que la resolución de las figuras es adecuada y no se ven borrosas en el tamaño al que van a estar impresas.

Para imprimir el póster lo más conveniente es acudir a tiendas o empresas especializadas que utilizan un plóter o trazador y que permiten elegir el acabado (brillo, mate, plastificado), la calidad y rigidez del papel, etc. Los colores serán los que hayamos usado en nuestro diseño, teniendo en cuenta que algunos tonos pueden variar ligeramente de la pantalla del ordenador a la versión impresa.

Para transportar un póster de estas dimensiones se suele introducir enroscado en un tubo rígido de plástico o cartón. Al sacarlo, conviene



Licencia: dominio público. Fuente: Pixabay.

<sup>31</sup> Enlaces relacionados:

Eposter: <http://www.eposters.net>

Pósters académicos (York Univ.): <http://sites.google.com/a/york.ac.uk/academic-posters/>

<sup>32</sup> Enlaces relacionados:

Piktochart: <http://www.piktochart.com>

Canva: <http://www.canva.com>

enroscarlo temporalmente en la dirección opuesta o estirarlo de algún modo para que pierda la forma adquirida tras permanecer en el tubo. Para colgarlo en el lugar establecido por la organización del congreso se puede usar cinta adhesiva de doble cara (lo más práctico), masilla adhesiva, chinchetas..., material que normalmente es facilitado por la organización.

## Estructura

El contenido del póster se puede estructurar en tres partes: encabezamiento, cuerpo principal y bloque final.

### – Encabezamiento.

Los pósters deben poder ser localizados con facilidad por los asistentes al congreso, por lo que el título debe tener un tamaño de fuente suficientemente grande, que se lea bien a una distancia de dos metros o más. Este tamaño no debe ser inferior a los 36 puntos, pero es mejor que sea bastante mayor, y es aconsejable que vaya en negrita.

Si el póster está relacionado con una memoria de investigación, su título debe ser el mismo, para identificar adecuadamente el trabajo. Únicamente en el caso de un título muy largo se puede usar para el póster una versión más corta. El título puede ir acompañado de un subtítulo que especifique un poco más el tema del trabajo, o que contenga la parte del título original de la investigación que, por su extensión, haya sido recortada. También puede aprovecharse el subtítulo para formular el tema del trabajo de una manera más atractiva y que invite a acercarse al póster, por ejemplo en forma de pregunta. El subtítulo debe tener un tamaño de fuente menor que el título y colocarse justo debajo, de manera parecida a los títulos y subtítulos en las noticias de los periódicos.

Tras el título y, si lo hay, el subtítulo, deben incluirse los nombres de todos los autores, y debajo de ellos su filiación, es decir, el centro en el que se llevó a cabo la investigación. En el caso del Certamen Jóvenes Investigadores, se refiere al centro educativo en el que cursaban sus estudios o la asociación juvenil a la que pertenecían cuando se realizó la investigación, indicando al menos el nombre, la localidad y la provincia, y debe incluirse el nombre del profesor tutor. También se pueden añadir direcciones de contacto (habitualmente el correo electrónico), si los autores están abiertos a continuar la relación con las personas que se interesan por su investigación. Toda esta información debe llevar un tamaño de fuente menor que el del título y el del subtítulo. Las filiaciones a veces se acompañan del logotipo del centro, de un tamaño acorde con el de la fuente empleada. Además, es recomendable incluir el nombre o logotipo del congreso en el que se presenta el póster, que puede colocarse encima del título (si se le puede dar una forma alargada en horizontal) o en alguna de las esquinas superiores del póster.

Por último, en el encabezamiento debe incluirse un resumen o *abstract* del trabajo, que puede ser el mismo que el de la memoria escrita o una versión aún más corta. Esta es la parte más extensa del encabezamiento y puede ocupar varios renglones (los anteriores elementos ocupan habitualmente una o dos líneas cada uno). El tamaño de fuente utilizado debe ser un poco menor que el de los nombres de los autores y sus filiaciones.

### – Cuerpo principal.

Para organizar mejor el cuerpo principal del póster es conveniente dividirlo en apartados, que pueden ser todos los de la memoria escrita o solo algunos de ellos. Los títulos de cada apartado deben tener un tamaño de fuente mayor que el del texto que llevan a continuación, pero menor que los usados en el encabezamiento. Es decir, títulos de secciones como «Introducción», «Motivación», «Hipótesis», «Materiales», etc., deben tener menor tamaño que el título y subtítulo generales del póster.

La descripción del contenido de cada uno de estos apartados es la misma que ya se dio para las secciones de la memoria escrita, pero en el póster la información tiene que estar mucho más

resumida. Puede ser interesante aportarla en forma de listas numeradas o con viñetas, es decir, con párrafos separados que comienzan con un número o con un símbolo (un guion, un punto grueso, etc.). De esta manera el texto se presenta de una forma más esquemática, que se lee con más rapidez y facilidad, y que es especialmente adecuada para listar objetivos, hipótesis, materiales, etc.

El texto de estos apartados nunca debe alargarse a costa de reducir el tamaño de la fuente u ocupando márgenes o separaciones, porque estropearía por completo la claridad y el atractivo del póster. Puede, y de hecho debe, omitirse información menos relevante (que puede aparecer en la memoria escrita, comentarse en la presentación oral o introducirse en las conversaciones frente al póster). Parte de los detalles no incluidos en el texto pueden aparecer en las figuras que los acompañen, aunque allí sean menos evidentes. Las figuras son fundamentales en el póster, y convendría que acompañaran al texto en la mayor parte de los apartados. Su tamaño ha de ser considerable, y por eso es muy importante que tengan muy buena resolución. Algunas pueden ser más decorativas y otras contener más información, pero todas contribuyen al atractivo visual del póster, siempre que no se abuse de su cantidad ni se repitan innecesariamente figuras similares. Por ejemplo, si se han recogido muchas medidas o datos en diversas circunstancias, y se han representado en gráficas distintas, quizá una o dos de ellas puedan resultar suficientemente representativas. Todas las figuras del póster deben llevar un pie de texto que explique brevemente su contenido, y deben estar relacionadas con alguno de los textos principales situados cerca de ellas.

– Bloque final.

En la parte inferior del póster se pueden incluir las referencias más importantes consultadas para realizar la investigación. Si algún fragmento del texto del póster está muy relacionado con alguna fuente o la reproduce literalmente (en cuyo caso iría entrecomillado), es imprescindible incluir su cita. Se recomienda usar el mismo formato de citación que en la memoria escrita (formato Harvard). Entre las referencias se puede citar el trabajo de los propios autores en el que se desarrolla lo expuesto en el póster, si está publicado de algún modo (como artículo, libro, etc.) o si está disponible en internet.

En este bloque también se pueden mencionar algunos agradecimientos e incluir información de contacto de los autores, si no aparecía ya en el encabezamiento; una manera interesante de hacer esto último es incluyendo un código QR que enlace con alguna página web de los autores o de su centro, especialmente si contiene información ampliada sobre el trabajo presentado.



Fuente: elaboración propia.

**Claves**

Tanto las imágenes como el texto deben tener un tamaño proporcional al del póster, y en el texto deben usarse diversas escalas de fuentes según su importancia y la estructura con que se ha configurado. Así, el título utilizará la fuente de mayor tamaño de todo el póster, seguida de la del subtítulo, la del resumen, la del nombre de los autores, etc.

En general, la información en el póster debe distribuirse espacialmente de la misma manera en que se lee un texto escrito: de arriba a abajo y de izquierda a derecha. En ocasiones puede alterarse

esta pauta si se organiza el póster como una línea temporal, un diagrama de flujo, o algún otro tipo de esquema, pero siempre indicando claramente con flechas cualquier orden de lectura que se desvíe del habitual.

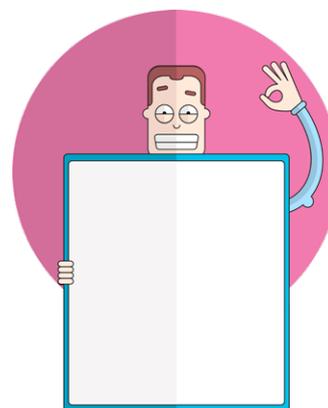
Debido al considerable ancho del póster, los renglones escritos con las fuentes más pequeñas resultan muy largos, lo que dificulta su lectura. Para evitarlo, es conveniente organizar ese texto en columnas, habitualmente dos. El encabezamiento, que contiene el título y alguna otra información principal, quedaría confeccionado con fuentes grandes en una sola columna y el resto del póster, con dos.

El póster no puede estar constituido por un número exagerado de figuras, pero tampoco puede ser un simple texto con letra grande. Lo importante es encontrar un equilibrio adecuado entre estos dos tipos de contenido, que podría rondar en torno a un 50 % de cada uno de ellos, sin olvidar un uso adecuado de los «vacíos»: separaciones entre textos y gráficas, entre diversos apartados del texto, interlineados y separaciones de párrafos, márgenes, etc.

Un póster colorido puede ayudar a atraer la atención y a hacerlo más estético, pero los colores deben seleccionarse y combinarse con sentido común. Pueden usarse colores de fondo diferentes para distinguir unos apartados de otros. El color de los fondos y de los textos que van sobre ellos deben combinarse de tal manera que contrasten lo suficiente, usando por ejemplo colores de fuente oscuros sobre fondos claros; el resultado debe ayudar a la lectura desde cierta distancia, nunca dificultarla.

Junto al póster se puede situar una mesa para exponer algún material interesante utilizado o creado durante la investigación (instrumentos, maquetas, prototipos, etc.). El hecho de poder ver o incluso manejar algunos de los objetos de los que se habla en el póster puede ayudar a comprender mejor su contenido y a hacerlo más ameno.

Además, la presentación en póster se puede acompañar de material impreso para repartir a las personas interesadas que lo visitan. Puede tratarse de una copia en tamaño reducido (DIN-A4) del propio póster, o de un folleto con contenido extraído del mismo, o incluso copias de la memoria escrita o algún otro trabajo propio asociado a la investigación, para el público muy interesado.



Licencia: dominio público. Fuente: Pixabay.

## 21.2. La exposición oral

La exposición oral de una investigación permite difundir su contenido a un amplio número de personas al mismo tiempo. Hoy en día es muy habitual, y se puede decir que casi obligatorio, acompañarla de transparencias o diapositivas, que son documentos digitales con texto e ilustraciones proyectados en la sala desde un ordenador.

Aprovechando el carácter multimedia de las transparencias proyectadas durante la exposición, pueden incluirse animaciones, simulaciones o vídeos sobre los experimentos realizados, pero preferiblemente han de ser breves, porque lo que debe primar es la transmisión de información en directo por parte de los ponentes. Siempre que sea posible, es muy conveniente sustituir esos vídeos por demostraciones reales frente al público.

Tanto la exposición oral como la de a pie de póster pueden acompañarse de los materiales reales usados en la investigación, o creados como resultado de esta: dispositivos experimentales, muestras recogidas, reproducciones, maquetas, prototipos, etc. Como se acaba de comentar, resulta aún mejor si se pueden observar esos materiales en acción a través de demostraciones. Para ello, se puede disponer de una mesa situada frente al póster o en el escenario durante la presentación oral; en este último caso el equipo técnico de la organización puede enfocar una cámara para proyectar la acción en la pantalla

del salón de actos, ya que es difícil que el público capte directamente los detalles de algo que sucede sobre una mesa en el escenario.

Las transparencias pueden elaborarse con aplicaciones específicamente diseñadas para ello, como Power Point de Microsoft, Keynote de Apple o la versión de código abierto Impress de OpenOffice. También existen opciones en internet que pueden producir resultados más dinámicos, como Prezi. Todas ellas ofrecen plantillas prediseñadas, que incluyen combinaciones de colores, fondos, disposición de cuadros con texto, tamaños de fuentes e imágenes, etc.

Además de los prediseños de estas aplicaciones, en internet se pueden consultar ejemplos de transparencias usadas en presentaciones de muy diversos ámbitos profesionales. Por ejemplo, LinkedIn SlideShare es un repositorio que contiene millones de estas presentaciones (e infografías y otro tipo de documentos), que pueden servir como inspiración para el diseño o la estructura<sup>33</sup>.

En las transparencias debe abundar el contenido gráfico (esquemas, gráficas, fotografías, etc.), en equilibrio con el texto, de un modo parecido al del póster. De hecho, puede ser buena idea elaborar el póster y las transparencias a la vez, ya que el contenido es muy similar y el diseño también puede serlo. En particular, las características fundamentales del diseño de los pósters descritas anteriormente son también aplicables a las transparencias. Todos los elementos deben tener el tamaño suficiente como para ser vistos desde una distancia considerable, de lo cual conviene asegurarse personalmente o ensayando con público en una sala grande.

Es muy importante comprobar antes de la exposición oral que todos los elementos de las transparencias permanecen en su lugar cuando se introducen en el ordenador proporcionado por la organización del congreso. Diferentes versiones de las aplicaciones pueden cambiar la disposición de los elementos en las transparencias, su tamaño o color, o incluso los caracteres de los textos, a veces con resultados desastrosos. Una manera de evitarlo es pasar la presentación a formato pdf, pero hay que tener en cuenta que en ese caso se pierden las animaciones y vídeos que hubiera en la presentación original.



Licencia: dominio público. Fuente: Pixabay.

## Estructura

Es conveniente estructurar la presentación oral con los mismos apartados que la memoria escrita, o al menos los más importantes entre ellos. El único apartado que quizás no tenga cabida es el del resumen o *abstract*, puesto que la audiencia presente va a escuchar la presentación completa y no es necesario adelantar los resultados ni las conclusiones. En la memoria escrita y en el póster este resumen sí es importante, porque los lectores pueden adquirir con él una idea general de la investigación sin necesidad de leer el resto de apartados.

Así, la exposición comienza con la presentación de los ponentes y del centro del que provienen e indicando el título de la investigación. En la transparencia asociada aparece esta información, que se puede acompañar de los logos del centro y del congreso en que se está presentando el trabajo. A continuación, se comenta brevemente la estructura que se dará a la presentación, y en la transparencia asociada se reproduce ese esquema o sumario. Las siguientes partes de la presentación serán la intro-

<sup>33</sup> Enlaces relacionados:

Prezi: <http://prezi.com>

LinkedIn SlideShare: <http://www.slideshare.net>

ducción, algunos comentarios acerca de los antecedentes y el estado de la cuestión, una enumeración de las hipótesis de la investigación y de los objetivos asociados, una descripción de los materiales y métodos empleados, y por último una explicación de los resultados obtenidos, su interpretación y las conclusiones extraídas. Para finalizar, se pueden mostrar brevemente las fuentes más importantes consultadas (pero no leerlas, porque resultaría tedioso), y expresar también brevemente los agradecimientos.

Cada uno de los apartados anteriores irá acompañado de una o varias transparencias, en las que se debe indicar claramente a qué apartado corresponden. Esa información puede figurar en un encabezamiento con un diseño común para todas las transparencias de la presentación, en el que además se puede incluir, por ejemplo, el nombre del congreso o su logo (en pequeño tamaño) y el título de la investigación.

## Claves

Como ocurre con el póster, es muy importante seleccionar para la exposición oral la información más relevante y evitar ser prolijo. Las transparencias no deben quedar muy recargadas, ni haber un número excesivo de ellas, y la exposición no puede realizarse a toda velocidad para que quepan más cosas. Al contrario, debe exponerse a un ritmo normal, que permita vocalizar, e incluso pueden repetirse los aspectos más importantes, que a su vez deben estar reflejados en los textos de las transparencias de forma concisa. De esta manera, cualquier distracción por parte de quien escucha puede subsanarse, al menos en lo fundamental, con un vistazo rápido a la transparencia proyectada en cada momento. En ningún caso se debe intentar incluir en las transparencias una transcripción completa de lo que se dice en la exposición. Al contrario, el texto debe estar resumido o presentarse de forma esquemática, e ir acompañado de ilustraciones.

Durante la exposición el ponente ha de dirigirse al público en todo momento, excepto en los breves instantes en que quiera señalar algún aspecto en la pantalla donde se proyectan las diapositivas. Es importante, por tanto, llevar muy bien preparado lo que se quiere decir, para no tener que consultar notas ni mirar continuamente la pantalla, donde seguramente una parte de lo que se está diciendo también está escrito. Una estrategia consiste en situarse cerca del ordenador desde el que se proyecta la presentación, ya que en su pantalla aparece la diapositiva y puede verse fácilmente sin dejar de dar la cara al público.

La exposición debe realizarse en un tono formal, pero que resulte ameno y atraiga la atención de la audiencia. Tiene que haber sido preparada y ensayada varias veces, si es posible delante de otras personas, pero sin que se note que se sabe de memoria, porque la entonación asociada resulta muy tediosa. Al contrario, debe parecer natural y casi espontánea, aunque no lo sea. El ensayo es fundamental para asegurarse de que la exposición cumple estrictamente el tiempo establecido por la organización, que en el Certamen Jóvenes Investigadores suele ser de 15 minutos. Es mejor acabar antes que pasarse del tiempo, aunque siempre hay un par de minutos de cortesía.

Uno de los aspectos más importantes a la hora de preparar una presentación oral es conocer a qué público irá dirigida, de forma que pueda resultar comprensible y atractiva para la audiencia. Se trata de una tarea difícil, ya que a menudo el público es muy variado. En el congreso del Certamen Jóvenes Investigadores la audiencia de las presentaciones orales está formada principalmente por los propios jóvenes, sus profesores tutores, los organizadores del Certamen y los miembros del jurado. Entre estos últimos, solo uno o dos serán especialistas en el área de la investigación presentada. La variedad de público es por tanto muy amplia, lo que hace difícil adaptarse a todos a la vez.

Lo más recomendable es que la exposición sea accesible a todo el mundo, transmitiendo las ideas principales de la investigación de una manera clara y amena. Se pueden introducir algunos detalles más técnicos dirigidos a los especialistas en el área, pero no demasiados. Si alguna persona, en particular algún miembro del jurado, quiere profundizar en el tema, ya lo habrá hecho leyendo la memoria escrita presentada, y puede preguntar sobre ello en la sesión de pósters. De hecho, uno de los

aspectos que se valoran en la presentación oral es que los autores de la investigación sepan transmitir el contenido y las implicaciones de su trabajo a un público lo más amplio posible. Así, el jurado evaluará la capacidad de los ponentes para llegar a todos los presentes en el auditorio, y no tanto el nivel de especialización o el uso de tecnicismos.

Reflexionar acerca del público receptor también es importante a la hora de redactar la memoria escrita o confeccionar el póster, pero en menor medida. La memoria escrita será leída principalmente por expertos en el área, y al póster se acercarán esos mismos expertos o personas específicamente interesadas en ese tema. Por tanto, en estos dos medios sí cabe un contenido más especializado y un lenguaje más técnico que en la exposición oral. Evidentemente, estas indicaciones no tienen por qué ser válidas en contextos distintos al del Certamen Jóvenes Investigadores. Por ejemplo, en conferencias científicas profesionales donde se reúnen únicamente expertos en un área concreta, la exposición oral tiene que alcanzar un alto nivel de rigor y especialización, acordes con el de la investigación que se está presentando.

Siguiendo en el ámbito del Certamen Jóvenes Investigadores, cuando la investigación ha sido realizada por dos o tres personas es importante que todas participen en la exposición. Lo habitual es que cada una hable durante unos minutos, repartiéndose las distintas secciones. Otra opción es que los autores se turnen en el uso de la palabra cada pocos segundos, encargándose de unas cuantas frases cada uno; aunque esto puede añadir dinamismo a la presentación, también la retrasa, y el cambio continuo de ponente causa distracciones, por lo que no resulta muy aconsejable.

Los vídeos de las exposiciones orales de la XXX edición del Certamen Jóvenes Investigadores pueden verse en el canal de Youtube del Instituto de la Juventud, que es coorganizador del Certamen<sup>34</sup>.



Fuente: elaboración propia.

<sup>34</sup> Enlaces relacionados:  
[http://www.youtube.com/playlist?list=PLIP7RS2XhMsT3iab6109Rtb82\\_d0QhKn5](http://www.youtube.com/playlist?list=PLIP7RS2XhMsT3iab6109Rtb82_d0QhKn5)

# EPÍLOGO



# Epílogo

Llegados a este punto, se ha dedicado un gran esfuerzo a reflexionar sobre la investigación científica desde sus aspectos más abstractos hasta los más prácticos, con la vista puesta en la elaboración de trabajos de investigación en la enseñanza secundaria y en la participación de los estudiantes en certámenes para jóvenes investigadores. Algunos de los contenidos aquí tratados se pueden acabar olvidando con el tiempo, pero hay otros aspectos cuyo impacto seguramente permanecerá: las percepciones, puntos de vista, aptitudes y actitudes adquiridas sobre el conocimiento científico y la investigación. Todo ello forma parte de unas competencias que pueden ser transmitidas al alumnado, muchas veces de manera expresa, pero en otras ocasiones como conocimiento tácito, no verbalizado, casi por imitación; así ocurre con la curiosidad, la cautela, el rigor, la paciencia, el esfuerzo, la crítica constructiva o la creatividad, ingredientes todos ellos esenciales para la actividad científica.

Estaríamos incumpliendo uno de los objetivos principales de este libro si pusiéramos aquí el punto final. Desde mañana mismo podemos comenzar a animar a nuestros jóvenes estudiantes a buscar temas de investigación apasionantes. En este libro se han introducido algunas de las herramientas necesarias para realizar las investigaciones, que requerirán sin duda gran esfuerzo y dedicación por parte de los jóvenes autores y también de sus profesores tutores. No hará falta manejar al detalle todo lo tratado en este libro, pero en ciertos aspectos habrá que profundizar más, según el tipo de investigación que se vaya a realizar.

Todo esfuerzo en este sentido merecerá la pena, y los resultados pueden sorprender. A veces las cosas son más fáciles de lo que parecen; todo es ponerse, y las dificultades se van superando poco a poco para dar paso a las satisfacciones, como los propios jóvenes investigadores pueden atestiguar<sup>35</sup>. Siempre resulta emocionante comprobar lo que son capaces de lograr, cargados de entusiasmo y curiosidad científica, con la ayuda y apoyo de sus profesores.

---

<sup>35</sup> Enlace relacionado:

Entrevistas a participantes finalistas del Certamen Jóvenes Investigadores:

<http://www.youtube.com/playlist?list=PLlP7RS2XhMsTrlAEij7a1VBZJGjfdVlgW>

Listas de reproducción sobre el Certamen Jóvenes Investigadores en el canal de Youtube del Instituto de la Juventud:

<https://www.youtube.com/user/injuvejuventud>

**PARA SABER MÁS**



# Para saber más

## Parte I

Un libro de divulgación muy conocido e interesante sobre el concepto de ciencia y su método es:

Chalmers, A. (1982): *¿Qué es esa cosa llamada ciencia?* Madrid: Siglo XXI.

Otros libros interesantes a nivel divulgativo son:

Bunge, M. (1989): *La investigación científica: su estrategia y su filosofía*. Barcelona: Ariel. También del mismo autor, varios libros de la Biblioteca Bunge, editorial Laetoli (Pamplona): *La ciencia, su método y su filosofía*; *Las pseudociencias, ¡vaya timo!*; *Pseudociencia e ideología*; *Ciencia, técnica y desarrollo*; etc.

Losee, J. (1997): *Introducción histórica a la filosofía de la ciencia*. Madrid: Alianza Editorial.

Pearl, J. (2018): *The book of why: the new science of cause and effect*. New York: Basic Books.

Ziman, J. (1981): *La credibilidad de la ciencia*. Madrid: Alianza Editorial.

Como libro de texto a nivel universitario sobre fundamentos de la ciencia resulta muy claro y conciso el siguiente (enfocado a las ciencias sociales):

Álvarez Álvarez, J. F., Teira Serrano D., y Zamora Bonilla, J. (2005): *Filosofía de las ciencias sociales*. Madrid: UNED. Disponible (edición 2019) en: <http://e-spacio.uned.es/fez/eserv/bibliuned:500632/Manual.pdf>

También se tratan algunos de los temas de esta parte en los libros de texto de filosofía de Bachillerato. Como ejemplo (de planes de estudio anteriores) se pueden citar los siguientes:

Tejedor Campomanes, C. (1998): *Introducción a la filosofía* (3.º BUP). Madrid: SM.

Barrio, J., y Fullat, O. (1981): *Eidos Filosofía* (3.º BUP). Barcelona: Vicens-Vives.

Las siguientes obras recogen ideas que forman parte de la visión actual del método científico, y que han sido mencionadas en el texto:

Hempel, C. (1988): *La explicación científica. Estudios sobre filosofía de la ciencia*. Buenos Aires: Paidós.

Kuhn, T. S. (1981): *La estructura de las revoluciones científicas*. México: Fondo de Cultura Económica.

Popper, K. (1962): *La lógica de la investigación científica*. Madrid: Tecnos.

— (1967): *El desarrollo del conocimiento científico. Conjeturas y refutaciones*. Buenos Aires: Paidós.

Feyerabend, P. (1975): *Contra el método. Esquema de una teoría anarquista del conocimiento*. Barcelona: Ariel.

— (1981): *Tratado contra el método*. Madrid: Tecnos.

Para profundizar aún más los conocimientos sobre metodología científica y filosofía de la ciencia se puede consultar una amplia lista de bibliografía comentada de autores hispanohablantes y traducciones al castellano elaborada por el profesor de Lógica y Filosofía de la Ciencia Andrés Rivadulla. Disponible en: [http://www.ucm.es/dpto\\_logica/filosofia-de-la-ciencia\\_rivadulla](http://www.ucm.es/dpto_logica/filosofia-de-la-ciencia_rivadulla)

## **Parte II**

El contenido de este bloque puede ampliarse en los siguientes textos:

Campbell, D. T., y Stanley, J. C. (1973): *Diseños experimentales y cuasiexperimentales en la investigación social*. Buenos Aires: Amorrortu editores.

Kerlinger, F. N., y Lee, H. B. (2002): *Investigación del comportamiento*. 4.ª ed. México: McGraw-Hill Interamericana.

Para algunas ciencias específicas:

Alía Miranda, F. (2005): *Técnicas de investigación para historiadores. Las fuentes de la Historia*. Madrid: Editorial Síntesis.

Hammersley, M., y Atkinson, P. (2007): *Etnografía: métodos de investigación*. 2.ª ed. Barcelona: Paidós.

## **Parte III**

El contenido de estadística descriptiva e inferencial recogido en este bloque puede encontrarse, hasta un cierto nivel de desarrollo, en los libros de texto de Matemáticas de Bachillerato, especialmente los de Matemáticas aplicadas a las ciencias sociales.

Para un tratamiento de mayor nivel y enfocado a investigaciones en ciencias sociales pueden consultarse los siguientes textos:

Cea D'Ancona, M. A. (2001): *Metodología cuantitativa. Estrategias y técnicas de investigación social*. Madrid: Editorial Síntesis.

Fernández Díaz, M. J. *et al.* (1999): *Resolución de problemas de estadística aplicada a las ciencias sociales*. Madrid: Editorial Síntesis.

Kerlinger, F. N., y Lee, H. B. (2002): *Investigación del comportamiento*. 4.<sup>a</sup> ed. México: McGraw-Hill Interamericana.

Para áreas como la biología y las ciencias de la salud pueden consultarse los siguientes textos:

Milton, J. S. (2001): *Estadística para biología y ciencias de la salud*. 3.<sup>a</sup> ed. México: McGraw-Hill Interamericana.

Samuels, M. L., Witmer, J. A., y Schaffner, A. A. (2016): *Statistics for Life Sciences*. 5.<sup>a</sup> ed. Reino Unido: Pearson Education.

Para profundizar más en la estadística desde el punto de vista matemático y en su aplicación a ciencias como la física y la química, se puede recurrir a los siguientes libros:

Spiegel, M. R., Schiller, J., y Alu Srinivasan, R. (2010): *Probabilidad y estadística*. 3.<sup>a</sup> ed. Madrid: McGraw-Hill.

Gorgas García, J.; Cardiel López, N., y Zamorano Calvo, J. (2009): *Estadística básica para estudiantes de ciencias*. Disponible en: [http://webs.ucm.es/info/Astrof/users/jaz/ESTADISTICA/libro\\_GCZ2009.pdf](http://webs.ucm.es/info/Astrof/users/jaz/ESTADISTICA/libro_GCZ2009.pdf)

Sánchez del Río, C. (1989): *Análisis de errores*. Madrid: EUDEMA.

A un nivel similar, los siguientes libros ofrecen también explicaciones claras a diversos aspectos de estadística:

Rowntree, D. (1981): *Statistics without tears*. New York: Charles Scribner's Sons.

Behar Gutiérrez, R., y Grima Cintas, P. (2011): *55 respuestas a dudas típicas de estadística*. 2.<sup>a</sup> ed. Madrid: Ediciones Díaz de Santos.

Y a nivel divulgativo, son muy interesantes los siguientes:

Bennett, D. J. (2000): *Aleatoriedad*. Madrid: Alianza Editorial.

Wheelan, C. (2014): *Naked statistics*. New York: W. W. Norton & Company.

Aczel, A. D. (2004): *Chance*. New York: Thunder's Mouth Press.

#### **Parte IV**

A un nivel básico, es interesante el siguiente libro de texto del Telebachillerato de México sobre la metodología de la investigación:

Reyes Corona, M. (2016): *Metodología de la investigación*. Ciudad de México: Secretaría de Educación Pública. Disponible en: <https://libros.conaliteg.gob.mx/TB1IA.htm>

Desde un punto de vista más teórico, el siguiente libro recoge reflexiones y propuestas sobre la enseñanza de la investigación científica a los jóvenes:

Koberski, M. L. (2016): *Cómo iniciar a los jóvenes en la investigación científica*. Córdoba, Argentina: Editorial Brujas.

Como inspiración para jóvenes científicos pueden servir los siguientes:

Ramón y Cajal, S. (2009): *Reglas y consejos sobre investigación científica. Los tónicos de la voluntad*. Madrid: Espasa Libros. También disponible en: <http://archive.org/details/ConsejosParaJovenes-CientificosPorSantiagoRamonYCajal>

Wilson, E. O. (2014): *Cartas a un joven científico*. Barcelona: Debate.

# APÉNDICES



# Apéndice I.

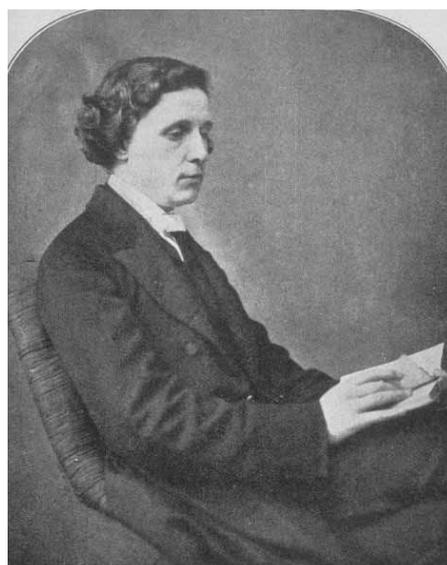
## Actividades para jugar con el método científico

Algunos de los procedimientos que forman parte del método científico pueden ponerse en práctica a través de juegos, que resultan muy apropiados para aclarar y asimilar conceptos. Aquí se describen algunos ejemplos.

### Actividad 1: El juego de la lógica

El escritor británico Charles Dodgson (1832-1898), autor, entre otras novelas, de la famosa *Alicia en el País de las Maravillas*, solía firmar con el seudónimo Lewis Carroll. Además de escritor era matemático, lógico y clérigo. Escribió varios libros sobre lógica, entre ellos *Lógica simbólica* y *El juego de la lógica*, en los que propone varios ejercicios o juegos lógicos, además de métodos sistemáticos para resolverlos.

Algunos de los enunciados de esos juegos no se corresponden con nuestra experiencia de la realidad, como por ejemplo «ningún bogavante es irracional», pero su estructura lógica permite combinarlos con otros enunciados y obtener conclusiones lógicamente válidas (aunque sean tan absurdas como las premisas de las que se deducen). Recordemos que la lógica es un procedimiento formal, donde lo único que importa son las consecuencias que se pueden extraer de las premisas, independientemente de que su contenido sea verdadero, falso o absurdo.



Charles Dodgson (Lewis Carroll).  
Licencia: dominio público. Fuente: Wikimedia Commons, Google LIFE Photo Archive.

En este juego se propone obtener la conclusión lógica que se puede deducir a partir de un conjunto de premisas que aparecen en juegos lógicos de Lewis Carroll, como en el siguiente ejemplo:

- a) Ningún país que haya sido explorado está infestado de dragones.
- b) Los países inexplorados son fascinantes.

¿Cuál es la conclusión lógica? Según la premisa a, todos los países explorados están libres de dragones, de modo que, si existe algún país infestado de dragones, permanece inexplorado por el momento. Según la premisa b, los países que permanecen inexplorados son fascinantes. Así que los países infestados de dragones (no sabemos si habrá alguno), permanecen inexplorados y por tanto son fascinantes. La conclusión lógica es por tanto:

Conclusión: Todos los países infestados de dragones son fascinantes.

También se puede llegar a la conclusión equivalente, aunque más enrevesada, de que «no hay ningún país infestado de dragones que no sea fascinante».

### **Juego 1**

- a) Todos los canarios bien alimentados cantan con potencia.
- b) Ningún canario se siente melancólico si canta con potencia.

¿Cuál es la conclusión lógica?

### **Juego 2**

- a) Ningún pájaro, excepto los pavos reales, se pavonea de su cola.
- b) Algunos pájaros que se pavonean de sus colas no saben cantar.

¿Cuál es la conclusión lógica?

### **Juego 3**

- a) Ningún ánade baila el vals.
- b) Ningún oficial declina nunca una invitación a bailar el vals.
- c) Todas mis aves de corral son ánades.

¿Cuál es la conclusión lógica?

Los razonamientos lógicos forman parte de nuestro sentido común, y por tanto pueden resultar evidentes. Con este juego se puede comprobar que esta afirmación es un poco exagerada, porque a veces aplicar la lógica no es tan sencillo. Aunque la deducción lógica no es el único proceso involucrado en el método científico, sí es una parte importante, y a la hora de desarrollar una investigación no se pueden cometer fallos de razonamiento. Afortunadamente, los razonamientos no suelen ser muy complejos, en el sentido de que no suelen involucrar un gran número de premisas entrelazadas, pero nunca viene mal estar bien entrenados.

Para saber más: Carroll, L. (1994) *El juego de la lógica*. Madrid: Alianza Editorial.

## **Actividad 2: El juego de la falsación de hipótesis**

El psicólogo británico Peter Wason (1924-2003) elaboró una sencilla prueba, la tarea de selección o test de las cuatro tarjetas, para comprobar nuestra capacidad de razonamiento deductivo. Dado que el método científico se basa en buena parte en la correcta aplicación de las reglas de deducción lógica en el marco del contraste de hipótesis, se puede emplear esta prueba como una simulación de actividad científica.

En el juego se cuenta con cuatro tarjetas, que tienen una letra por una cara y un número por la otra. Solo se muestra una cara de cada tarjeta, generalmente dos caras con números y dos caras con letras. La prueba consiste en determinar qué tarjetas es necesario girar para comprobar si la siguiente hipótesis es cierta: «Si en una cara de la tarjeta hay una vocal, entonces en la otra cara hay un número par». El juego se gana cuando se giran únicamente las tarjetas que sirven para comprobar la hipótesis, y no se giran las que no sirven.



Licencia: dominio público.  
Fuente: Pixabay y elaboración propia.

Las cuatro tarjetas pueden representar el área de estudio de una ciencia, por ejemplo, el mundo físico, y la hipótesis establecida podría ser una regla o ley acerca de su funcionamiento. A partir de la hipótesis, aplicando un razonamiento deductivo, se pueden hacer predicciones sobre lo que habrá en la cara que no se muestra de cada tarjeta. Girar una tarjeta para ver la otra cara equivale a hacer un experimento, y lo observado ha de compararse con la predicción.

Supongamos que nos muestran las siguientes caras de cuatro tarjetas: [4] [A] [7] [Z].

Recordemos que la hipótesis que se quiere contrastar es:

«Si en una cara de la tarjeta hay una vocal, entonces en la otra cara hay un número par»

En este juego se propone contestar las siguientes preguntas:

¿Qué tarjetas habría que girar para demostrar si la hipótesis es verdadera o falsa? ¿Qué tarjetas no habría que girar, porque no afectan al contraste de la hipótesis?

Para resolver el juego puede resultar de ayuda analizar una por una las predicciones que se deducen de la hipótesis para cada una de las tarjetas:

- ¿Qué predice la hipótesis sobre la otra cara de la tarjeta [A] (vocal)? ¿Es necesario dar la vuelta a esta tarjeta para contrastar la hipótesis?
- ¿Qué predice la hipótesis sobre la otra cara de la tarjeta [7] (número impar)? ¿Es necesario dar la vuelta a esta tarjeta para contrastar la hipótesis?
- ¿Qué predice la hipótesis sobre la otra cara de la tarjeta [Z] (consonante)? ¿Es necesario dar la vuelta a esta tarjeta para contrastar la hipótesis?
- ¿Qué predice la hipótesis sobre la otra cara de la tarjeta [4] (número par)? ¿Es necesario dar la vuelta a esta tarjeta para contrastar la hipótesis?

En general, una hipótesis condicional del tipo «si se da el hecho X, entonces se produce el hecho Y» es falsa solo cuando el hecho X es cierto (es decir, se produce) y el hecho Y es falso (es decir, no ocurre). Por eso en este juego será necesario comprobar únicamente las tarjetas que muestran el hecho X (para ver si en la otra cara presentan o no el hecho Y) y las tarjetas que muestran algo que no es Y (para ver si en la otra cara presentan o no el hecho X).

Al finalizar el juego conviene reflexionar sobre su utilidad para explicar y practicar el procedimiento de falsación de hipótesis de Popper: desde las conclusiones que cabe deducir de la hipótesis de partida hasta el diseño y selección de las observaciones o experimentos que deben llevarse a cabo.

Los resultados de la prueba de las cuatro tarjetas realizados por Wason en 1966, y los de numerosos estudios similares llevados a cabo desde entonces, muestran que más del 75 % de los adultos de

cualquier nivel educativo, incluyendo el universitario, lo fallan, es decir, no son capaces de seleccionar única y exclusivamente las tarjetas necesarias para contrastar la hipótesis. El origen de esta confusión se encuentra en el razonamiento lógico de la afirmación del consecuente, que es incorrecto o falaz.

Aunque usar números pares e impares y letras de tipo vocal y consonante es bastante abstracto, los aciertos en la prueba son igualmente bajos cuando se emplean conceptos mucho más familiares, como comidas y bebidas, y además con hipótesis condicionales habituales, como por ejemplo «Si comes carne roja, bebes vino tinto». En este ejemplo las tarjetas podrían mostrar: [chuletón] [vino tinto] [lenguado] [vino blanco].

Curiosamente, los resultados cambian radicalmente (hasta un 75 % de aciertos) cuando la hipótesis tiene un contenido que se puede denominar de «contrato social»: acciones que se pueden llevar a cabo o no, desde el punto de vista legal o normativo, dependiendo de que se cumplan ciertos requisitos. Por ejemplo, para comprobar si se cumple la ley de que los menores de edad no pueden beber alcohol, se quiere contrastar la siguiente hipótesis en un bar: «Si una persona está bebiendo alcohol, entonces es mayor de edad». El juego de las cuatro tarjetas en este caso podría consistir en: [cerveza] [20 años] [refresco] [16 años].

De todo esto se podría deducir que los humanos no estamos naturalmente preparados para el método científico, en particular para la falsación de hipótesis, pero sí estamos dotados de unas robustas estructuras mentales para las relaciones sociales. Pero esto es un tema abierto en la psicología actual.

Con estos nuevos ejemplos, se puede realizar esta prueba a algunas personas de nuestro entorno para ver la diferencia en los resultados según el contenido de la hipótesis. Se pueden usar pósit y escribir en una de las caras de cada uno de ellos las etiquetas: [chuletón] [vino tinto] [lenguado] [vino blanco] para la primera prueba, y las etiquetas: [cerveza] [20 años] [refresco] [16 años] para la segunda prueba. Solo es necesario escribir una cara de los pósit, porque este juego no consiste en dar realmente la vuelta a las tarjetas, sino en seleccionar las que sería necesario girar.

### Actividad 3: El juego del método científico

El juego pautas (*Patterns*) es uno de los cientos que creó el ingeniero norteamericano Sidney Sackson, muchos de los cuales están a la venta como juegos de mesa, y otros que no necesitan materiales sofisticados vienen descritos en varios de sus libros, como *A gamut of games*. La dinámica de Pautas representa con bastante fidelidad el método científico, ya que hace uso de la inducción, de la creación de hipótesis, de la deducción y de la comprobación empírica.

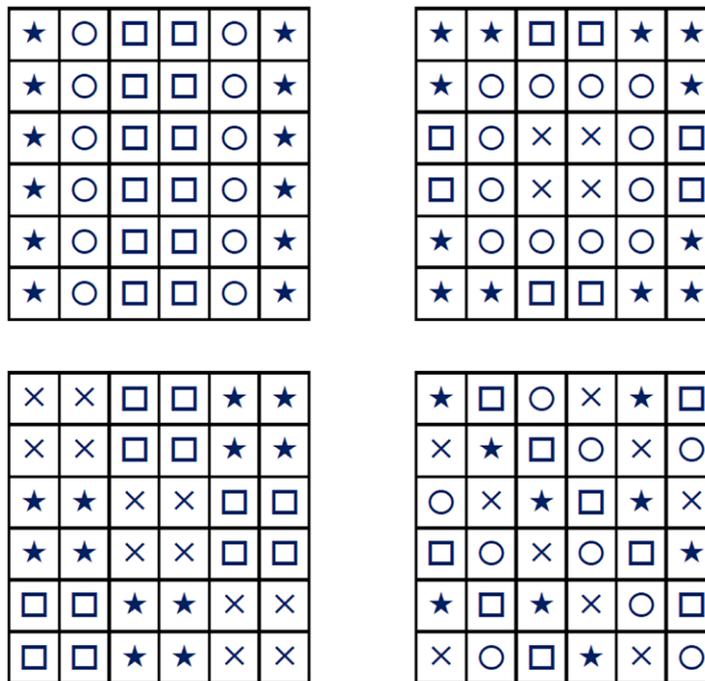
Cada jugador dibuja en un papel un pequeño tablero cuadrado de  $6 \times 6$  casillas. Un jugador actúa como diseñador, y rellena las 36 casillas de su tablero con símbolos de cuatro tipos, por ejemplo aspa, círculo, estrella y cuadrado, manteniéndolo oculto. El resto de jugadores debe adivinar el patrón de símbolos que ha introducido el diseñador en su tablero. Para ello, cada jugador puede pedirle al diseñador en cualquier momento y las veces que quiera (no hay turnos) que le diga qué símbolos ha colocado en algunas casillas, y lo trasladará a su propio tablero inicialmente vacío. Para facilitararlo, las casillas del tablero pueden ser identificadas usando coordenadas de letras (en las filas) y números (en las columnas), como en el famoso juego de hundir barcos.

Cuantas más casillas pregunte un jugador, menos puntuación obtendrá al final. Si tras una serie de preguntas uno de los jugadores cree poder completar el resto de casillas que le faltan para reproducir el tablero completo del diseñador, las rellenará en su tablero con el símbolo entre paréntesis, para distinguir las casillas que han sido deducidas de las que han sido preguntadas.

El sistema de puntuación es como sigue. Cada jugador ordinario (no diseñador) recibe un punto por cada casilla acertada, es decir, en la que ha deducido el símbolo que había puesto el diseñador, y recibe un punto negativo por cada casilla fallada; las casillas que fueron preguntadas no cuentan para la puntuación. Si se retira del juego sin proponer una solución al tablero completo, su puntuación es cero. En cuanto al diseñador, su puntuación es el doble de la diferencia entre la mayor y la menor puntuación obtenida por los jugadores ordinarios. Además, se le restan puntos por cada jugador que ha abandonado

sin proponer solución al tablero: 5 puntos menos por el primer abandono y otros 10 puntos menos por cada abandono adicional. Por ejemplo, en una partida con tres jugadores ordinarios, uno obtiene 8 puntos y los otros dos abandonan y obtienen por tanto 0 puntos; la puntuación del diseñador será el doble de la diferencia entre 8 y 0, es decir, 16, y además hay que restarle 5 puntos por el primer abandono y 10 puntos por el segundo abandono, obteniendo una puntuación final de 1 punto.

Aunque el diseñador puede rellenar su tablero como quiera, las puntuaciones del juego están pensadas para que obtenga mayor puntuación si lo hace con patrones regulares de símbolos, como los de las figuras de ejemplo que aparecen más abajo. Ello permite que los otros jugadores encuentren un patrón de regularidad a partir de las casillas que conocen (porque las han preguntado), y puedan adivinar así los símbolos de las casillas que les faltan. Pero al diseñador tampoco le interesa que el patrón sea muy sencillo, porque si muchos jugadores lo adivinan, todos obtienen puntuación alta y la diferencia entre la mayor y la menor puntuación es muy pequeña. Así, el diseñador debería rellenar su tablero con un patrón lo suficientemente fácil como para ser adivinado en gran parte por algún jugador, pero lo suficientemente difícil como para que no lo consigan todos; y, por supuesto, debe evitar hacerlo tan irregular que los jugadores se vean incapaces de inducir patrones y decidan abandonar.



Fuente: elaboración propia.

En este juego el diseñador actúa como la naturaleza, y el resto de jugadores son científicos. Cada vez que uno de los jugadores pregunta al diseñador por el contenido de una casilla, está realizando un experimento. Con los datos que se van obteniendo de los experimentos, los jugadores formulan por inducción hipótesis acerca del patrón de símbolos que ha empleado el diseñador para rellenar su tablero. Es decir, los científicos crean hipótesis por inducción sobre el funcionamiento de la naturaleza. Una vez que se maneja una hipótesis, se deduce de ella qué símbolo debería aparecer en alguna de las casillas vacías, y se contrasta preguntando al diseñador. Si la predicción es errónea, se ha falsado la hipótesis y hay que cambiarla. Si, por el contrario, la hipótesis funciona en varias comprobaciones, puede darse por buena y proceder a completar lo que queda del tablero para ganar puntos. ¿Podemos estar seguros en algún momento, antes de preguntar todas las casillas, de que nuestra hipótesis es la correcta? La respuesta es que no, porque el patrón puede ser más complicado de lo que pensamos; todo depende de lo regular o lo caótico que haya querido ser nuestro diseñador (o la naturaleza).

Para saber más:

Sackson, S. (1992): *A gamut of games*. New York: Dover Publications.

Gardner, M. (1979): *Circo matemático*. Madrid: Alianza Editorial.

#### Actividad 4: El juego del sudoku para diseños factoriales

En la parte II vimos que en muchos diseños de investigaciones la muestra se divide en distintos grupos, uno para cada categoría (valor discreto) de la variable independiente. En un diseño factorial, que incluye varias variables independientes, el número de grupos necesarios crece muy deprisa: el número de variables multiplicado por el número de categorías en cada una de ellas. En consecuencia, la complejidad de la investigación también aumenta mucho.

Existe la posibilidad de formar menos grupos, siempre que aparezcan en la misma proporción todas las variables y todas sus posibles categorías. Las combinaciones adecuadas pueden construirse a través de los cuadrados latinos o grecolatinos, como se comentó en el capítulo 10. Este modo de proceder recuerda mucho a la resolución de un sudoku, y de hecho son casi lo mismo: el conocido pasatiempo consiste típicamente en un cuadrado de  $9 \times 9$  casillas en las que hay que situar las nueve cifras del 1 al 9 sin repetirlas en ninguna fila ni en ninguna columna (con la dificultad adicional de que en los subcuadrados de tamaño  $3 \times 3$  tampoco puede repetirse ninguna cifra, y además algunas cifras ya vienen dadas y no se pueden cambiar).

En este juego se propone idear una investigación factorial con tres o cuatro variables independientes y varias categorías en cada una de ellas (el mismo número en todas), y construir un diseño a través de un cuadrado latino (en el caso de tres variables) o grecolatino (en el caso de cuatro variables). Un ejemplo con tres variables y tres categorías en cada una de ellas se dio en el capítulo 10. En el caso de tres variables se puede ampliar el juego a más categorías por variable, pero para cuatro variables puede resultar demasiado complicado.

También se puede proporcionar a los participantes las características de la investigación (variables independientes y sus categorías), así como la estructura del cuadrado latino o grecolatino con alguna casilla cumplimentada, de modo que tengan que averiguar el resto de casillas (tanto encabezamientos de las filas y las columnas como casillas interiores). En este caso es necesario construir primero el cuadrado correcto completo para asegurarse que las casillas que se dejan forman parte de un cuadrado válido. Este juego sería el más parecido al sudoku, pero remplazando las cifras por categorías de variables para diseñar de manera óptima una investigación factorial.

El nombre de cuadrado latino fue acuñado por el matemático suizo Leonhard Euler (1707-1783), porque colocaba  $n$  caracteres latinos (a, b, c...) en las  $n \times n$  casillas del cuadrado. Si se distribuyen dos conjuntos de  $n$  elementos, simbolizando el segundo de ellos con caracteres griegos ( $\alpha$ ,  $\beta$ ,  $\gamma$ ...), resulta el cuadrado grecolatino (ver figura).

Además de para diseñar de forma óptima investigaciones con varias variables independientes, estos cuadrados son muy interesantes desde el punto de vista puramente matemático. Uno de los problemas asociados es el de calcular el número de cuadrados grecolatinos distintos de tamaño  $n \times n$  que pueden formarse, recordando que su construcción exige colocar parejas de caracteres latinos y griegos de forma que en cada fila y en cada columna aparezca cada carácter solo una vez y no se repita ninguna pareja en todo el cuadrado.

A $\alpha$	B $\gamma$	C $\beta$
B $\beta$	C $\alpha$	A $\gamma$
C $\gamma$	A $\beta$	B $\alpha$

Cuadrado grecolatino  $3 \times 3$ . Licencia: dominio público. Fuente: Wikimedia Commons.

Euler descubrió métodos para construir cuadrados grecolatinos de lado  $n$  impar o múltiplo de 4, y conjeturó que el resto, los de lado  $n$  par que no fueran múltiplos de 4, no podían existir. Transcurrieron casi 180 años hasta que se pudo demostrar que la conjetura de Euler era falsa, y que realmente sí se pueden construir cuadrados grecolatinos de cualquier tamaño, con dos únicas excepciones: de lado  $n = 2$  (cuya imposibilidad es trivial) y de lado  $n = 6$  (que es el origen del problema de los 36 oficiales de Euler<sup>36</sup>). Estas investigaciones matemáticas fueron de las primeras de la historia cuya resolución se apoyó en cálculos con computadoras digitales, en concreto la UNIVAC, que era usada por el ejército estadounidense y por la NASA.

Para saber más: Gardner, M. (1982): *Nuevos pasatiempos matemáticos*. Madrid: Alianza Editorial.



Leonhard Euler. Licencia: dominio público.  
Fuente: Wikimedia Commons.

### Soluciones al juego de la lógica:

Juego 1: Ningún canario bien alimentado se siente melancólico.

Juego 2: Algunos pavos reales no saben cantar.

Juego 3: Ninguna de mis aves de corral es un oficial.

### Solución al juego de la falsación de hipótesis:

Habría que dar la vuelta a la carta que muestra la A (vocal) y la que muestra el 7 (número impar). La hipótesis sería falsa si detrás de la carta con la A hubiera un número impar o si detrás de la carta con el 7 hubiera una vocal. Es necesario girar estas dos cartas, porque ambas tienen capacidad de invalidar la hipótesis.

Lo que no tiene sentido es girar las cartas con el 4 o con la Z, porque son irrelevantes para invalidar la hipótesis. Detrás del 4, número par, podría haber una consonante, pero la hipótesis no dice nada sobre qué tipo de número va asociado a las consonantes. Por la misma razón, es irrelevante lo que haya detrás de la Z.

<sup>36</sup> Enlace relacionado:

<http://www.ams.org/publicoutreach/feature-column/fcarc-latini1>

# Apéndice II.

## Actividades para reflexionar sobre la ciencia

### Actividad 1: Reflexión sobre las fronteras entre las áreas de la ciencia

La frontera entre las ciencias naturales y las ciencias sociales, ¿es siempre nítida? ¿Qué ocurre, por ejemplo, con la psicología o con la lingüística?

Aunque aquí hemos usado habitualmente la denominación de ciencias sociales, ¿es ese el término más adecuado siempre? ¿se puede trazar alguna diferencia con las ciencias humanas? ¿existe alguna diferencia con las humanidades?

Por su parte, las matemáticas o la lógica estudian estructuras abstractas sin conexión, *a priori*, con el mundo real, pero muchas de ellas acaban teniendo aplicación en la descripción de fenómenos naturales o sociales, y por eso son tan importantes para las demás ciencias. ¿Por qué esto es así? ¿Por qué nuestro mundo es matemático y se comporta de manera lógica? ¿Las matemáticas se inventan, o están más ligadas con la realidad del mundo de lo que pensamos, y por tanto se descubren?

### Actividad 2: Reflexión sobre la naturaleza de la causalidad

El concepto de causa es muy delicado en ciencia. La aplicación del método científico da lugar a leyes que describen de manera concisa las regularidades que surgen entre ciertos fenómenos naturales o sociales. Pero esas regularidades no tienen por qué implicar una conexión causa-efecto. De hecho, ¿cómo se puede comprobar empíricamente esa conexión? ¿cómo se puede demostrar la diferencia entre correlación temporal y causalidad?

Uno de los filósofos que más ha incidido en la imposibilidad de demostrar la causalidad es el escocés David Hume (1711-1776). A propósito de este tema escribió:

No tenemos otra noción de causa y efecto, excepto que ciertos objetos siempre han coincidido, y que en sus apariciones pasadas se han mostrado inseparables. No podemos penetrar en la razón de la conjunción. Sólo observamos la cosa en sí misma, y siempre se da que la constante conjunción de los objetos adquiere la unión en la imaginación.

Así, para Hume, percibimos que un hecho sucede a otro, pero no que exista alguna condición necesaria y suficiente entre ambos; no podemos afirmar que un hecho causó otro, solo sabemos con seguridad que un hecho está correlacionado con otro (en conjunción). Según Hume solo podemos confiar en el conocimiento que adquirimos a través de nuestros sentidos, y por eso considera que la relación causa-efecto no es una característica real de la naturaleza, sino únicamente un mecanismo psicológico: tanto el ser humano como otros animales poseen la tendencia instintiva a creer en la relación causa-efecto. La causalidad se sustenta así en la esperanza de que ciertos hechos se den después que otros.

¿Por qué establecemos la relación causa-efecto de forma intuitiva? ¿Hacerlo favorece nuestra supervivencia? En caso afirmativo, ¿por qué es así, si se supone que esa relación no es una característica real de la naturaleza?

Para saber más: Hume, D. (1888): *A Treatise of Human Nature*. Oxford University Press, Oxford<sup>37</sup>.

### Actividad 3: Reflexión sobre los límites de la ciencia

¿Cuáles son los límites de la ciencia? ¿Pretende describir y explicar una realidad objetiva externa al propio ser humano? ¿Es esa realidad accesible a nuestros sentidos y comprensible por nuestra mente? Si, aparentemente, lo ha sido hasta ahora, ¿existen garantías de que lo siga siendo siempre o puede existir un límite para el avance del conocimiento?

Si creemos que la realidad no existe o es inalcanzable científicamente, ¿qué otras funciones podría desempeñar la ciencia? Dado que es capaz de generar a menudo predicciones correctas sobre la naturaleza, ¿no significa eso que describe adecuadamente una realidad externa a nosotros?

¿Es la ciencia una actividad aislada de todo lo demás, que se puede describir de manera relativamente simple mediante un conjunto de pasos teóricos bien definidos? ¿o se ve influida por factores psicológicos, culturales y sociales? Si esto último ocurre, ¿es deseable? ¿es evitable?

¿Qué distingue la ciencia de la filosofía o de la religión? ¿Son compatibles la ciencia y las creencias?

¿Qué distingue la ciencia de las pseudociencias, incluyendo las pseudoterapias médicas como la homeopatía o el movimiento antivacunas? ¿Es fácil desacreditar las pseudociencias manteniendo una actitud científica? Es conveniente recordar en este punto las ideas de Popper, para quien el espíritu científico no consiste en la búsqueda de la certeza, sino en la permanente crítica y revisión de lo que se da por sabido.



Licencia: dominio público. Fuente: Pixabay.

<sup>37</sup> Enlace relacionado:

A Treatise of Human Nature: <http://archive.org/details/treatiseofhumann002393mbp/page/n367/mode/2up>

# Apéndice III.

## Actividades para elaborar una investigación científica

Al plantear una investigación para un certamen de jóvenes investigadores, y también para otros contextos, conviene reflexionar sobre una serie de aspectos que ayudarán a perfilar el diseño y el desarrollo de la investigación, a redactar la memoria final y a revisar el trabajo una vez finalizado. Para ello, en estas actividades se formulan una serie de preguntas organizadas siguiendo los contenidos tratados en las cuatro partes de este libro y particularizadas en gran medida a las características y especificaciones del Certamen Jóvenes Investigadores de la Secretaría General de Universidades, que se trató con detalle en el capítulo 18.

### Actividad 1: Planteamiento general de la investigación

Una vez elegido un tema de investigación, se debe reflexionar sobre los siguientes aspectos:

- a) ¿En qué sentido el tema elegido es innovador, original y representa un avance del conocimiento en su área? ¿Es adecuado para las capacidades y medios de los autores? (Conviene repasar las orientaciones sobre los temas de investigación dadas casi al final del capítulo 18).
- b) ¿A qué objetivo general de la ciencia (uno o varios) se dirige principalmente y de qué manera? (En referencia a lo explicado en el capítulo 1).
- c) ¿En qué ciencias se enmarca principalmente la investigación y de qué tipo son? (En referencia a lo explicado en el capítulo 1).
- d) ¿La investigación se basa en la observación o en la experimentación? ¿De qué manera? (En referencia a lo explicado en el capítulo 2).
- e) ¿Podría requerir algún procedimiento de división o clasificación? ¿De qué entidades? (En referencia a lo explicado en el capítulo 2).

- f) ¿Podría requerir la definición de algún concepto nuevo o hacerlo de manera más precisa que en el lenguaje común? ¿De qué conceptos? (En referencia a lo explicado en el capítulo 2).
- g) ¿Qué tipos de explicación serían las más adecuadas para una investigación de ese tipo y por qué? (En referencia a lo explicado en el capítulo 2 y ampliado en los capítulos 4, 5 y 6).
- h) ¿Se podría hacer uso de la inducción para elaborar hipótesis? ¿Sobre qué tipo de observaciones o medidas? (En referencia a lo explicado en el capítulo 2).
- i) Plantear algunas posibles hipótesis de trabajo relacionadas con el tema de investigación elegido (en referencia a lo explicado en el capítulo 3).
- j) ¿Qué consecuencias experimentalmente comprobables se pueden deducir lógicamente de las hipótesis planteadas? ¿Qué observaciones o experimentos podrían llevarse a cabo para comprobar si las consecuencias deducidas se cumplen? (En referencia a lo explicado en el capítulo 3).

### Actividad 2: Análisis de los principales elementos de una investigación general

- a) ¿Qué tipo de metodología sería la más adecuada para la investigación? (En referencia a lo explicado en el capítulo 8).
- b) ¿Cuáles serían las variables más importantes en la investigación, cómo se relacionarían entre sí, y qué escalas de medida serían más apropiadas para cada una de ellas? (En referencia a lo explicado en el capítulo 9).
- c) ¿Cuál sería el diseño más adecuado para la investigación? (En referencia a lo explicado en el capítulo 10).
- d) ¿Qué técnicas de control de variables podrían resultar adecuadas para la investigación? (En referencia a lo explicado en el capítulo 11).
- e) ¿Cuáles serían las fuentes de invalidez, tanto internas como externas, que más podrían afectar a la investigación y por qué? ¿Cómo podrían mitigarse? (En referencia a lo explicado en el capítulo 12).

### Actividad 3: Análisis de los principales elementos de una investigación cuantitativa

- a) ¿Qué estadísticos muestrales podrían resultar adecuados para describir los datos numéricos recogidos en la investigación, y a qué variables corresponderían? (En referencia a lo explicado en el capítulo 13). Dar alguna estimación del tamaño de la muestra que podría ser adecuado.
- b) ¿Sería apropiado llevar a cabo algún procedimiento de estadística inferencial? ¿Cuál? (Estimación de parámetros, contraste de hipótesis). ¿A qué variables estaría referido y cuál sería su objetivo? (En referencia a lo explicado en los capítulos 14 y 15).
- c) ¿Sería adecuado asociar incertidumbres a las medidas de la investigación? ¿A cuáles y de qué tipo serían (directas o indirectas)? ¿Cómo se obtendrían o calcularían esas incertidumbres? (En referencia a lo explicado en el capítulo 16).
- d) ¿Qué tipos de representación gráfica de datos serían adecuados en la investigación y para qué variables? (En referencia a lo explicado en el capítulo 16).

### Actividad 4: Elaboración y revisión de la memoria escrita y de los materiales para la presentación pública

La memoria escrita de la investigación presentada al Certamen Jóvenes Investigadores no puede exceder de 10 páginas de texto, junto con otras 10 de un anexo de figuras. En el texto han de incluirse los siguientes apartados (ver su descripción en el capítulo 19):

- Portada.
- Índice.
- Resumen y su traducción al inglés (*abstract*).
- Introducción.
- Antecedentes. Apoyarse en el contenido del capítulo 20 sobre la investigación secundaria y el uso de fuentes documentales.
- Hipótesis y objetivos. Apoyarse en los contenidos de la parte I.
- Materiales y métodos. Apoyarse en los contenidos de las partes II y III.
- Resultados. Apoyarse en los contenidos de las partes II y III.
- Conclusiones.
- Agradecimientos.
- Bibliografía y webgrafía. Apoyarse en el contenido del capítulo 20 para citar correctamente las fuentes.
- Anexo de figuras.

En la fase de elaboración de la memoria de una investigación ya concluida, cabe reflexionar sobre los siguientes aspectos:

- a) ¿Se cuenta con toda la información necesaria para cumplimentar de manera completa y concreta todos los apartados de la memoria?
- b) ¿Qué tipo de fuentes serán necesarias para redactar de forma completa y coherente el apartado de antecedentes? (Si no se han usado ya durante la propia investigación) ¿Dónde se pueden encontrar?
- c) De todos los datos recogidos y procesados durante la investigación, ¿cuáles convendría seleccionar para ser mostrados en la memoria, dadas las restricciones de espacio? ¿De qué manera sería más adecuado hacerlo (texto redactado, listas, tablas o gráficas de datos numéricos, etc.)?
- d) ¿Qué tipo de figuras convendría incluir en el anexo de la memoria? ¿Serían de autoría propia o tomadas de alguna fuente?

Una vez elaborado un borrador casi definitivo de la memoria, es conveniente plantearse las siguientes preguntas:

- a) ¿Reúne la memoria las características de formato requeridas (tamaño y tipo de fuente, espaciado, extensión, etc.)?
- b) ¿Se incluye en cada apartado la información específica que le corresponde, sin mezclarla con la de otros y sin que falten datos importantes?
- c) ¿Se han citado correctamente todas las fuentes empleadas? ¿Se ha atribuido la autoría de las figuras, su fuente y su licencia?
- d) ¿Ha quedado la memoria redactada con el nivel científico adecuado, con estilo formal y de la forma más clara y fluida posible? ¿Resulta también, en la medida de lo posible, informativa e interesante con carácter general? Es recomendable que también lean la memoria y contesten a esta pregunta personas ajenas a la investigación, que deben ser mencionadas en el apartado de agradecimientos en reconocimiento a la lectura crítica realizada.

Un conjunto análogo de cuestiones, ligeramente adaptadas, resultará también apropiado para la elaboración de materiales destinados a la presentación pública de la investigación, como el póster, el conjunto de diapositivas o el guion de la presentación oral, cuyas características se describen en el capítulo 21; en este caso será importante plantearse, además de todo lo anterior, lo siguiente:

- a) ¿Resulta el material atractivo visualmente (diseño, color, imágenes, etc.)? ¿Contiene un equilibrio adecuado entre texto y figuras? ¿Incluye algún aspecto destinado a captar la atención y el interés de la audiencia?

- b) ¿Reúne el material las características técnicas adecuadas para transmitir información a cierta distancia y a varias personas a la vez (tamaño de fuentes, espaciados, colores, intensidad, nitidez, volumen de voz, vocalización, etc.)? De nuevo, es interesante que estos aspectos sean valorados también por personas distintas a las que han elaborado los materiales.

### Actividad 5: Simulación de la evaluación de la memoria

En el Certamen Jóvenes Investigadores las investigaciones son evaluadas en una primera fase a través de la memoria escrita presentada, siguiendo los criterios descritos en el capítulo 18.

Debe notarse que no en todos los casos existe una correspondencia directa entre criterios de valoración y apartados de la memoria. Algunos de ellos deben valorarse según el contenido de varios apartados, o sobre la memoria en su conjunto (por ejemplo, el de claridad en la exposición).

A continuación se ofrece una rúbrica de evaluación de la memoria basada en los criterios de valoración del Certamen Jóvenes Investigadores, que puede servir como guía para realizar una revisión crítica del trabajo realizado hasta el momento.

<b>Grado alcanzado →</b>	<b>Nulo</b>	<b>Bajo</b>	<b>Aceptable</b>	<b>Alto</b>	<b>Excelente</b>	
<b>Ítems de evaluación ↓</b>						<b>Asignación del ítem ↓</b>
Originalidad	0	2.5	5	7.5	10	
Calidad del marco teórico	0	2.5	5	7.5	10	
Correcto planteamiento de hipótesis científicas y objetivos de investigación	0	2.5	5	7.5	10	
Adecuación de la metodología empleada	0	2.5	5	7.5	10	
Claridad en la exposición	0	2.5	5	7.5	10	
Correcta extracción y análisis de conclusiones	0	2.5	5	7.5	10	
Autonomía de los autores en el diseño y desarrollo de la investigación	0	2.5	5	7.5	10	
Grado de utilización de recursos propios por parte de los autores de la investigación	0	2.5	5	7.5	10	
<b>Resultado total de la rúbrica (sobre 80) →</b>						

Este texto pretende ofrecer una visión general del método científico e introducir las técnicas de investigación más importantes en todas las ramas del conocimiento. Está especialmente dirigido al profesorado de Enseñanza Secundaria Obligatoria, Bachillerato y Formación Profesional que desee tutorizar trabajos de investigación científica realizados por sus estudiantes, ya sea porque forme parte del currículo oficial o porque los estudiantes muestren interés. Pretende además servir de guía para la participación en iniciativas como el Certamen Jóvenes Investigadores, convocado anualmente por la Secretaría General de Universidades y el Instituto de la Juventud, cuyo fin es despertar vocaciones científicas en los jóvenes de entre 15 y 20 años mediante la concesión de importantes premios a sus trabajos de investigación.

El contenido pretende refrescar algunos conocimientos y profundizar en otros, discutiendo aspectos como los fundamentos del método científico moderno, los elementos básicos sobre el diseño y la calidad de las investigaciones, algunos métodos específicos para investigaciones cuantitativas y el papel de la enseñanza de la investigación científica en la educación secundaria, tanto en el aula como a través de certámenes para jóvenes investigadores. Estos temas, con las oportunas adaptaciones pedagógicas por parte del profesorado, pueden resultar también muy útiles para los estudiantes de secundaria interesados en la investigación científica.