

Antonio Moreno Verdejo
y Adela María Villegas Escobar

Matemáticas electorales

CLAVES PARA INTERPRETAR SONDEOS Y ELECCIONES



© ANTONIO MORENO VERDEJO
Y ADELA MARÍA VILLEGAS ESCOBAR, 2017

© FEDERACIÓN ESPAÑOLA DE SOCIEDADES DE PROFESORES
DE MATEMÁTICAS (FESPM), 2017
SERVICIO DE PUBLICACIONES
AVDA. DE LA MANCHA S/N
02006 ALBACETE
WWW.FESPM.ES

© INSTITUTO DE CIENCIAS MATEMÁTICAS (ICMAT), 2017
NICOLÁS CABRERA, Nº 13-15
CAMPUS DE CANTOBLANCO, UAM
28049 MADRID
WWW.ICMAT.ES

© LOS LIBROS DE LA CATARATA, 2017
FUENCARRAL, 70
28004 MADRID
TEL. 91 532 20 77
FAX. 91 532 43 34
WWW.CATARATA.ORG

MATEMÁTICAS ELECTORALES.
CLAVES PARA INTERPRETAR SONDEOS Y ELECCIONES

ISBN: 978-84-9097-363-9
DEPÓSITO LEGAL: M-28.795-2017
IBIC: PDZ/YQM/JPHF

ESTE LIBRO HA SIDO EDITADO PARA SER DISTRIBUIDO. LA INTENCIÓN DE LOS EDITORES ES QUE SEA UTILIZADO LO MÁS AMPLIAMENTE POSIBLE. QUE SEAN ADQUIRIDOS ORIGINALES PARA PERMITIR LA EDICIÓN DE OTROS NUEVOS Y QUE, DE REPRODUCIR PARTES, SE HAGA CONSTAR EL TÍTULO Y LA AUTORÍA.

ÍNDICE

Introducción. En busca del sistema electoral perfecto 9

Capítulo 1. ¿Cómo se sabe lo que pienso? 23

Capítulo 2. ¿Puede un número limitado de entrevistas representar la opinión de los votantes de un país? 41

Capítulo 3. ¿Cuándo un sistema de elección es democrático? 59

Capítulo 4. ¿Cómo se pueden asignar los escaños de una forma proporcional a los votos? 83

Bibliografía 115

Introducción

En busca del sistema electoral perfecto

El ideal de una democracia representativa es simple: cada individuo ha de tener la misma voz que el resto para configurar una cámara de representantes en la que se delegarán ciertas decisiones políticas. A pesar de que parece simple, no es sencillo conseguirlo. Y no siempre ha sido así a lo largo de la historia. Algunas naciones solucionaban este problema en salas de fumadores o bien establecían comisiones para realizar esta tarea. Algunas, como Francia o Inglaterra, no utilizaban métodos de reparto, sino que los resultados se negociaban. Otras, mientras tanto, comenzaron a desarrollar modelos matemáticos. Para diseñar los diferentes métodos de reparto electorales o los distintos modelos de elección social, las matemáticas son clave porque permiten considerar las distintas preferencias de los electores sin que el beneficio de unos se convierta en el perjuicio de los otros.

En esencia, existen dos tipos de sistemas: el federal y el de representación proporcional, aunque en ocasiones se da una combinación de ambas. En ambos casos, el porcentaje de votos que reciben las candidaturas determina de manera proporcional el número de escaños que les son asignados en el órgano correspondiente. En un sistema federal, el estado recibe los escaños de acuerdo a su población, mientras que en los de representación proporcional es el partido político el que recibe

los escaños según los votos recibidos. Por ejemplo, Estados Unidos tiene un sistema federal, España e Israel uno de representación proporcional y Alemania y Suiza una combinación de ambos que relaciona el territorio con los partidos. La diferencia entre ambos sistemas se entenderá mejor si pensamos en los intereses que defiende fundamentalmente cada diputado electo. En un sistema federal habría diputados de Galicia que pertenecen a distintos partidos y en el sistema de reparto proporcional actual hay diputados de un partido que han sido elegidos en Galicia.

Características generales

Proporcionalidad

La idea de representación proporcional es tan antigua como el concepto de democracia. Los griegos la aplicaban ya en el siglo V a. C. y en principio, es sencilla. Cuanto mayor sea la población o el número de votos, mayor será el número de escaños recibidos. Pero en realidad el concepto de proporcionalidad no es suficiente para resolver el problema.

Reparto de las partes decimales

La transformación de votos a escaños rara vez produce números enteros, es decir, los escaños que corresponden a cada partido según la proporción de votos recibidos no son necesariamente exactos. Así pues, dado que evidentemente los representantes son indivisibles, la dificultad está en la asignación de los restos decimales.

Este problema ha inquietado a matemáticos y políticos durante cientos de años, ya que la manera en la que se redondeen los decimales hará variar los resultados. En EE UU, por ejemplo, en el año 1970 la cuota exacta, es decir, el número de escaños o diputados que correspondían, por ejemplo, al colegio electoral o circunscripción de Oregón en la Cámara de

Representantes (que en aquel entonces tenía 435 escaños) era 4,5. En ese caso, ¿se le asignan cuatro o cinco escaños? Si el número se hubiese redondeado a cinco, dando así a los ciudadanos de Oregón “una voz más alta” de lo que le corresponde matemáticamente, otro estado, póngase uno con 4,45 de cuota, conseguiría solo cuatro escaños y por lo tanto menos representación de la que le corresponde.

Circunscripciones y censo

La siguiente pregunta que se plantea, en el caso federal, es si la representación de la población debe distribuirse respecto a los votantes solamente o al número total de personas de cada estado. A lo largo de la historia se han tomado diferentes posturas. Por ejemplo, en EE UU en el año 1787 se tomó en consideración a las poblaciones representativas como la base de la distribución. La población representativa de cada estado se obtenía contando el total de personas en cada uno, descontando un 40% de los esclavos y excluyendo todos los indios no sujetos a impuestos. A pesar del número de votantes, se eligió una base menos ambigua que reflejaba más estrechamente la distribución de las propiedades. El tratamiento especial de los esclavos representó un balance peculiar entre los intereses del Sur y del Norte.

Las bases de la distribución, o el criterio para considerar la cantidad de población entre la que se reparte el número de los representantes que serán elegidos, han cambiado, por supuesto, en el curso de la historia de EE UU y volverán a cambiar. En 1868 se estableció que: “los representantes serán repartidos entre los distintos estados según su número respectivo de habitantes, contando el total de personas en cada estado y excluyendo los indios no sujetos a impuestos”. Los indios empezaron a ser incluidos como parte de la población en el censo de 1940. Y en 1970, por primera vez, parte de la población estadounidense en ultramar fue asignada a sus estados y contabilizada a la hora de hacer la distribución. La nueva definición de población representativa incrementó menos del 1%

el total de la población. Sin embargo, produjo cambios. Por ejemplo, Connecticut perdió un escaño en favor de Oklahoma.

Distribución de escaños

Cada método matemático conduce a unos resultados, ya que pone el énfasis en ciertas características del reparto. Por ejemplo, que facilite la formación del Gobierno, que se mantenga una cierta representación geográfica, que cada voto valga exactamente lo mismo, etc. En los sistemas de representación proporcional, a veces la intención es sacrificar la imparcialidad y dejar ventaja a los grandes partidos para que puedan formar coaliciones, como sucede en España. Si el discutible fin es mantener la estabilidad política, un partido que se haya fragmentado pierde sus ventajas y se desalienta a los partidos pequeños.

Escoger un método u otro es una cuestión de gran interés, claro, porque la distribución de escaños determina, finalmente, el poder. Conviene que sea lo más justa posible, aunque la idea de justicia es también relativa. La teoría general del reparto, deducida de la experiencia a lo largo de los años y del sentido común, lleva a dos conclusiones. La primera de ellas es que el método no debe disminuir los escaños de un partido cuyos votos aumentan para dárselo a otro partido cuyos votos han disminuido. Es decir, en una nueva elección a un partido le corresponde un número de escaños para un determinado número de votos. La segunda conclusión que se extrae es que si la cámara mantiene el número de escaños, el método no restará escaños a un estado que mantiene la población.

Para evitar estas y otras paradojas, la elección del método debe estar restringida a la familia de los denominados *métodos de divisores*, que incluye casi todos los que se emplean en el mundo. La historia y el sentido común, unidos, sugieren para los métodos de reparto un conjunto de principios o reglas que señalan lo que se debe satisfacer y lo que se debe evitar. Con estos principios se deducen los métodos aceptables, pero la

teoría demuestra, como veremos, que no hay ningún modelo matemático de reparto de escaños que sea perfecto.

Sin embargo, para llegar a esa conclusión muchas fueron las discusiones y, además, la ausencia de un sistema perfecto deja el debate abierto para siempre. La visión internacional de la evolución de los sistemas electorales relativiza la idea de la existencia del método esencialmente justo.

La historia ha mostrado una rica colección de recetas para resolver el problema esencial de dividir proporcionalmente una cámara de representantes. Cada uno de los métodos se basa en una fórmula precisa, y cada uno da un resultado claro para cualquier situación. ¿Cuál de ellos capta mejor el sentido de proporcionalidad? A lo largo de la historia, los más populares han sido el método de Hamilton, el de Lowndes, el de Jefferson, el de Adams, el de Dean, el de Webster, el de Hill... Es inútil tratar de discutir los pros y los contras de los diferentes métodos apelando simplemente a las fórmulas en las que se basan. Cada fórmula tiene su propia lógica y ningún método es completamente convincente.

El tema se complica aún más, puesto que ciertas fórmulas y procedimientos que en apariencia son diferentes dan en realidad el mismo reparto en todos los casos y por lo tanto, representan el mismo método disfrazado. El método de Jefferson fue redescubierto con una forma diferente por Seaton en la década de 1880 y en 1878 por D'Hondt para la representación proporcional. El método de Webster fue propuesto de forma independiente por Sainte-Lagüe en 1910, con una fórmula totalmente diferente. El estudio de sus analogías y diferencias permite categorizar los distintos métodos utilizados históricamente: una primera categoría que incluye los métodos de Hamilton y Lowndes, y una segunda que consiste en los métodos divisores que incluyen los cinco métodos restantes.

Los métodos de tipo Hamilton toman como cálculo principal la llamada cuota, que es el número de escaños o diputados que corresponden a un colegio electoral o circunscripción. Cada uno de ellos comienza por calcular las cuotas y dar a

cada estado la parte entera correspondiente y después reparten los escaños restantes (correspondientes a los decimales que no se han repartido). Aunque más adelante se estudiarán los métodos de reparto con detalle, podemos poner un par de ejemplos: el método de Hamilton da prioridad a los estados que tienen los restos mayores; el método Lowndes ajusta las prioridades dividiendo el resto de cada estado por el número entero de su cuota y asignando los escaños sobrantes a aquellos cuyos restos sean mayores. Este procedimiento matemático responde a la clara lógica de este último enfoque: garantiza un resultado que se mantiene dentro de la cuota. Ninguna asignación se desvía de su cuota más de un escaño y cada estado tiene su cuota redondeada hacia abajo o hacia arriba. Pero también es evidente que existen innumerables maneras de construir un método que se mantenga dentro de la cuota.

La segunda categoría es aquella basada en un divisor común. Los métodos de Jefferson, Webster y Adams son ejemplos de estos métodos divisores. La idea esencial es establecer una proporción entre la población y sus representantes y luego dividir esta proporción o divisor entre las poblaciones de los estados para obtener cocientes. Los cocientes se redondean hacia arriba o hacia abajo hasta el número entero vecino de acuerdo con una regla que depende de cada método divisor. En el capítulo 4 se encuentran ejemplos desarrollados que ayudarán al lector a comprender el método.

Evidentemente, no hay solo cinco métodos divisores. La colocación de los divisores en el intervalo da lugar a un método diferente, por lo que se puede generar el método según las necesidades. Un ejemplo pintoresco obtiene el criterio de reparto a partir de un descubrimiento de los griegos: la razón áurea, la más famosa de todas las proporciones. Esta fue reverenciada en la Antigüedad por la orden de los pitagóricos, que hasta la usaron para construir la estrella de cinco puntas que les servía de insignia. La proporción áurea se halla dividiendo un segmento en dos partes desiguales, de forma que la proporción entre el segmento y la parte mayor sea igual que

la proporción entre la parte mayor y la parte pequeña. Esto sugiere un método divisor completamente diferente, en el que el divisor para conseguir el asiento adicional es la proporción áurea de cada intervalo.

Por todo ello, a lo largo de la historia, determinar el método más “justo” ha requerido formular normas o principios que permitieran evaluar los métodos, y los analizaremos con detalle más adelante. En general, hay que tener en cuenta que la justicia de un método de reparto está relacionada con la agudeza y perspicacia política.

250 años de democracia en EE UU

El país en el que más intensamente se ha cuestionado la elección de un método para resolver el reparto de escaños es EE UU, ya que es una cuestión importante por su influencia en los temas nacionales. La discusión comenzó en la Convención Constitucional en el año 1787, de la que salió la Constitución de los Estados Unidos. La polémica volvió a surgir cuando los resultados del primer censo fueron expuestos en 1791 y los grandes cambios poblacionales en algunos estados rompieron el equilibrio de poder en la Cámara de Representantes y en el colegio electoral. La Constitución recogía que después del primer censo, que se haría dentro de los tres años posteriores a la primera sesión del Congreso, un nuevo censo debía hacerse cada diez años. Así, cada diez años a partir de aquella fecha se ha abierto el debate, e inevitablemente volverá a ocurrir. En muchos otros países se encuentran disposiciones similares para actualizar periódicamente las bases de la representación.

La Constitución impuso limitaciones al número de miembros en el Congreso. Así, el número de representantes no excedería de uno por cada 30.000 personas y cada estado tenía asegurado al menos un representante, estableciendo así un límite superior y un límite inferior para la representación. Asignar a cada miembro de una confederación un número

mínimo de escaños, en comparación con su población representativa, es una limitación que está presente en muchos otros países. En Francia cada distrito recibe al menos dos diputados y en Canadá cada provincia tiene garantizados tantos miembros al parlamento como senadores, y el número de senadores varía de provincia a provincia. El convenio adoptado en 1975 para regir la distribución del nuevo Parlamento Europeo incluyó el principio de evitar cualquier reducción en la “cantidad actual de representantes de cualquier estado”, lo que implica que cada uno de los cuatro países más grandes recibiría por lo menos 36 y el más pequeño, Luxemburgo, por lo menos seis. En España, todas las circunscripciones tienen asegurado al menos un escaño.

Otro elemento de debate que apareció de forma recurrente, era el número de representantes que debían asignarse a cada estado. De hecho, el tamaño de la Cámara de Representantes ha crecido de 65 asientos en 1790 a 435 en la actualidad. La Constitución tampoco especificó un método definitivo de reparto de los escaños. Esta falta de claridad ha prolongado el debate sobre cómo asignar los asientos hasta la actualidad.

Tanto el poder de un estado en el Congreso como su influencia en la elección de presidente se determinan por el método de reparto elegido. En sus casi 250 años de democracia, EE UU ha tenido 19 repartos diferentes. Estos métodos se pueden comparar sobre todos los datos disponibles (la población y el número de escaños en cada elección). Además, la experiencia proporciona información sobre el funcionamiento de los métodos en la práctica.

Como ya se ha dicho, en EE UU el sistema es federal, es decir, cada estado obtiene sus escaños de forma proporcional a su población. El problema de la representación justa aparece ya en su Constitución: repartir la representación justamente entre todos los estados. Sin embargo, no expresa cómo se debe hacer. No se establece una norma permanente y precisa que establezca una regla de representación apropiada.

A lo largo de los años la democracia estadounidense ha identificado cuatro elementos esenciales sobre la distribución: la proporcionalidad como ideal, la base de la proporcionalidad en una población representativa, como veíamos al inicio del capítulo. El tercer elemento que afecta a la distribución de escaños es aplicar revisiones periódicas para explicar los cambios en las poblaciones de los estados, y finalmente, considerar que aplicar ciertas limitaciones superiores e inferiores podía anular el ideal puro de proporcionalidad. Estos elementos básicos se encuentran en prácticamente todos los países que tienen un gobierno representativo.

Pero hay otras consideraciones que no están resueltas, como la forma de dar a los grandes estados su justa parte del poder político y, al mismo tiempo, asegurar a los estados pequeños una voz que no se ahogara entre los grandes. Los estados pequeños, temerosos de ser sumergidos por un sistema proporcional, defienden un sistema federal en el que cada estado tuviese un voto igual. Finalmente prosperó la idea de una legislatura compuesta de dos cámaras, una la Cámara de los Representantes, repartida por poblaciones y otra, el Senado, repartido de forma equitativa entre todos los estados, recibiendo cada uno dos escaños. Este tipo de acuerdo entre grandes y pequeños es común en muchos sistemas federales.

Como vemos, no es sencillo encontrar una solución justa. Los efectos de las diferentes fórmulas sobre la composición del Congreso pueden ser dramáticos. Esto se puede observar comparando la primera regla empleada, propuesta por Thomas Jefferson, con la propuesta posteriormente por John Adams (véase capítulo 4). Al aplicar ambas a los datos del censo de 1790, el método de Jefferson daría a Nueva York un total de 41 escaños, mientras que el método de Adams solo le daría 37. En general, los resultados serían diferentes en 27 estados e involucrarían una diferencia de 18 escaños. Las diferencias dependen del método de reparto empleado.

Encontrar el sistema más equitativo tiene una gran importancia. Aunque pueda parecer mínima la diferencia de unos

cuantos escaños aquí o allí, sabemos que los efectos pueden ocasionar consecuencias trascendentales, como por ejemplo el cambio en el resultado de una elección presidencial. De hecho, esta situación ya se presentó. El mal reparto realizado en 1870 fue responsable en 1876 de la elección de Rutherford B. Hayes, a pesar de que su oponente disfrutara del 51,6% de los votos. Además, a veces la variación de escaños no es pequeña, sino que puede haber un cambio significativo a favor de los pequeños estados rurales, perjudicando a grandes estados industriales, o a la inversa, dependiendo del método usado. Es por ello que los políticos, desde Alexander Hamilton en 1792 hasta Arthur Vandenberg en 1941, se han involucrado en diversos conflictos sobre los métodos de reparto, incluso cuando los efectos eran transferir tan solo un escaño.

Ejercicio 1. Consulta la Constitución española y la Ley Orgánica de Régimen Electoral General para responder las siguientes cuestiones:

- a) ¿Cómo se establece la representación de la población?
- b) ¿Está previsto algún mecanismo para actualizar las bases de la representación en función de los cambios en la población?
- c) ¿Existen límites superiores e inferiores para la representación?

Historia de los sistemas electorales en España

La España de los siglos XIX y XX se caracterizaba por una gran inestabilidad institucional y política. La invasión del país por Napoleón en 1808 fue seguida por el colapso del absolutismo de la monarquía borbónica. En 1810 se reunió en Cádiz un parlamento (Cortes) elegido indirectamente por todos los propietarios mayores de 25 años. Este amplio sufragio, con la edad de votación reducida a 21, se confirmó en la constitución de 1812, que estipulaba una monarquía constitucional. En 1814 el rey Fernando VII regresó y restauró la monarquía absoluta. En 1823 el ejército le obligó a reinstaurar la Constitución de 1812, pero fue capaz de restaurar la monarquía absoluta tres años más tarde con ayuda del ejército francés.

Tras la muerte de Fernando VII en el año 1833, su viuda, María Cristina, se convirtió en regente de su hija Isabel II. En 1834 el Estatuto Real estableció un parlamento bicameral con una cámara alta nominada y una cámara baja elegida por un sufragio de propiedad extremadamente restringido. En 1836 un golpe militar llevó a la restauración temporal de la Constitución de 1812 y a la elección de unas Cortes, que en 1832 votaron por una Constitución que proporcionara un sufragio mucho más amplio que el del Estatuto Real, pero menos que un sufragio censitario. En 1840, Espartero sustituyó a María Cristina como regente y tres años más tarde, otro golpe militar lo derrocó y revisó la Constitución de 1837.

En 1868, el régimen fue derrocado por la armada y la reina Isabel dejó el país. Se aprobó una monarquía constitucional con un parlamento bicameral, ambos elegidos por sufragio universal masculino y se nombró rey a Amadeo de Saboya. En 1873 el rey abdicó y se declaró la Primera República. Sin embargo, un golpe militar restauró la monarquía borbónica un año más tarde.

La nueva constitución (1869) establecía una monarquía de poder limitado con un parlamento de dos cámaras. La cámara baja, el Congreso de Diputados, debía ser elegido por un sufragio limitado a hombres mayores de 25 años que cumplieran con los requisitos mínimos de pago de impuestos, con lo que podía votar alrededor del 5% de la población.

En 1890 el sufragio masculino fue restaurado. Los diputados fueron elegidos por mayoría, 307 en circunscripciones uninominales y 88 en circunscripciones plurinominales, que devolvían entre tres y ocho diputados. El sistema de partidos de la restauración estaba dominado por el partido liberal y el partido conservador. Debido a un acuerdo entre ambos partidos, ambos alternaban el cargo. Esta disposición, llamada “turno pacífico”, fue implementada por el Ministerio de Interior junto con los jefes políticos locales, los caciques.

Desde 1881 hasta 1923 ningún gobierno perdió unas elecciones generales, que siempre ganaban el liberal o conservador.

Aunque los partidos dinásticos siguieron siendo los dominantes en las Cortes, los partidos de la oposición pudieron ganar la representación parlamentaria en las zonas rurales, donde su apoyo local era en ocasiones abrumador. El apoyo a los partidos de la oposición crecía cada vez más en las ciudades donde el gobierno se mostraba incapaz de manejar las elecciones de forma efectiva. La división dentro del partido liberal y del partido conservador, así como la derrota del ejército español en Marruecos, condujo al pronunciamiento de 1923 y a la dictadura de Primo de Rivera (1923-1930), en la que se prohibieron los partidos políticos.

En agosto de 1930, los opositores republicanos al régimen se aliaron con algunos de los que fueron líderes de los antiguos partidos dinásticos en un pacto para derrocar la monarquía. En abril del año siguiente, las elecciones municipales arrojaron como resultado que la alianza entre republicanos y socialistas había ganado una abrumadora mayoría en las grandes ciudades. De este modo, el rey Alfonso XIII abandonó el país y se formó un gobierno provisional que convocó elecciones para Cortes. Los miembros de las Cortes fueron elegidos por sufragio universal masculino, con una edad mínima para votar de 23 años en 61 circunscripciones plurinominales. Los territorios del norte de África, Ceuta y Melilla, eran circunscripciones uninominales. En las circunscripciones plurinominales se utilizó un sistema de voto limitado.

Los partidos o alianzas de partidos que ganaron por mayoría, en el momento en que consiguieran al menos el 20% de los votos, obtenían tantos escaños en la circunscripción como los electores habían votado. Si ningún partido o alianza alcanzaba este umbral se realizaba una segunda vuelta una semana más tarde en la que la mayoría era suficiente para la elección de candidatos individuales.

Las Cortes eligieron una abrumadora mayoría republicana y fue promulgada la Segunda República española en 1931. Estipulaba un parlamento unicameral elegido por sufragio universal masculino y el umbral de la primera ronda se elevó

del 20 al 40%. La participación en la segunda ronda se limitó a los partidos que ganaban el 8% de la votación en la primera. En un sistema de partidos muy fragmentado, los que pudieron formar amplias alianzas electorales se vieron favorecidos. En 1933 las divisiones electorales entre los partidos republicanos y socialistas permitieron a los radicales y partidos de la derecha política ganar la mayoría en las Cortes. En 1936 la alianza del Frente Popular de Comunistas, Socialistas, Izquierda Republicana e Izquierda Catalana obtuvo 280 de los 474 escaños en las Cortes con el 48% de los votos.

Tras la guerra civil, se estableció un régimen de partido único con el general Franco como jefe de Estado. En 1976, las cortes franquistas aprobaron una ley de reforma política que permitía la elección de unas cortes bicamerales por sufragio universal para redactar una nueva constitución. Los partidos políticos fueron legalizados y un parlamento constitucional fue elegido en 1977. La Constitución fue aprobada por referéndum en 1978, estipulando una monarquía constitucional con un parlamento de dos cámaras compuesto por un Senado y un Congreso de Diputados.

La ley electoral de la Asamblea Constituyente promulgó una ley provisional que establecía el sufragio universal de los adultos con una edad mínima de 21 años. Para el Congreso de los Diputados, cada una de las 50 provincias forma una circunscripción única que devuelve al menos tres diputados y un máximo de 33. Ceuta y Melilla cada una retorna un diputado. Los electores votan por una lista de partidos y no pueden expresar una preferencia por los candidatos individuales. En las circunscripciones plurinominales, los diputados son elegidos por el método D'Hondt. En Ceuta y Melilla la elección es por mayoría. Un umbral a nivel de circunscripción limita la representación a los partidos que obtienen al menos el 3% de los votos. Por último, la constitución redujo la edad de votación a los 18 años, lo que supuso el inicio del sistema electoral actual.

Capítulo 1

¿Cómo se sabe lo que pienso?

La forma más sencilla y evidente de conocer la opinión de un grupo de personas es preguntárselo directamente. De esta forma, durante el siglo XIX, las sociedades curiosas por el posible resultado de unas elecciones reñidas realizaban unos simulacros del acto electoral antes del escrutinio. Estas consultas no tenían rigor alguno, pero gustaban a los ciudadanos. Los primeros en realizar un sondeo local de la opinión electoral de sus lectores fueron los diarios *Harrisburg Pennsylvanian* y *Raleigh Star* en los años veinte. Los lectores completaban un formulario incluido en el periódico con el nombre de la persona a la que pensaban votar y luego lo enviaban por correo a la redacción. Otra veces, las encuestas las realizaban enviando cuestionarios por correo a sus lectores suscritos.

El interés por este tipo de sondeos creció y fue necesaria la intervención de las matemáticas para dar rigor y credibilidad a las afirmaciones realizadas a partir de la recogida de información. La comunidad de matemáticos dedicados a la estadística desarrolló métodos científicos apropiados para conocer características de la población a partir de un número de muestras seleccionadas adecuadamente. Contar con las herramientas matemáticas precisas para analizar datos impulsó el uso generalizado de este tipo de consultas, que adquirieron especial notoriedad a finales del periodo de entreguerras. Y a su

vez, el consiguiente desarrollo de las matemáticas asociadas al análisis de los datos obtenidos en las encuestas hizo que estas se convirtieran en una importante técnica de investigación.

Como consecuencia, los gobiernos se preocuparán por conocer las sociedades que gobiernan, los medios de comunicación descubrirán el interés de sus seguidores por los pronósticos de las elecciones y las grandes empresas querrán saber los gustos de los consumidores antes de poner un producto en venta.

En el siglo XX, el sondeo más conocido fue el de la revista americana *Literary Digest*, que en las elecciones estadounidenses de 1936, tras gastar mucho dinero en enviar cuestionarios por correo y analizar los datos de aproximadamente dos millones de encuestas, predijo erróneamente quién las ganaría. Sin embargo, Gallup, Crossley y Roper, con una encuesta realizada a menos de 5.000 personas, aventuraron que Franklin D. Roosevelt sería el vencedor y acertaron. Estos tres investigadores habían creado una metodología para la realización de encuestas de opinión pública que utilizaba muestreos representativos, seleccionando muestras que contenían las características de toda la población. Para ello, clasificaban a los entrevistados según su nivel socioeconómico, sexo y rango de edad. A pesar de este éxito inicial, no será hasta después de la segunda guerra mundial cuando la encuesta de opinión comience a ser la técnica predominante para recoger información.

En España, durante este mismo periodo, el régimen dictatorial retrasará la aparición de las encuestas de opinión en la vida social. En realidad, no negaba la realización de encuestas, pero determinados temas no encontraban facilidades para su publicación. Solo al final de la dictadura aparecieron publicados sondeos sobre algunos temas políticos. En cualquier caso, la publicación de encuestas tiene detractores y defensores sobre su valor científico y sobre la injerencia innecesaria, e incluso peligrosa, de su publicación en periodo electoral. El significado implícito de la consabida frase de muchos políticos, “la mejor encuesta es la del día electoral”, es suficientemente

esclarecedora. Esta polémica ha llevado a que se regule la publicación de sondeos en periodo electoral mediante el artículo 69 de la Ley de Régimen Electoral, para defender al ciudadano de una información manipulada.

¿Qué es una encuesta?

En la actualidad, el interés por conocer la opinión social no solo permanece vigente, sino que incluso parece imprescindible. Los medios de comunicación, además de reflejar los resultados, responden al interés por saber los procesos. Sin embargo, el lector no siempre cuenta con los conocimientos matemáticos para comprender el procedimiento de obtención de datos y la veracidad de los resultados. Ante esta necesidad, conviene empezar por lo que ya intuimos.

Ejercicio 1. A partir de la lectura de los siguientes fragmentos de artículos:

- a) Explica la expresión “cocinar una encuesta”.
- b) Indica las dificultades que encuentran las empresas demoscópicas para la predicción de resultados electorales.
- c) Explica algunas técnicas de recogida de información en los sondeos electorales.
- d) Define “ficha técnica de una encuesta”.

Así se hacen y se cocinan las encuestas electorales: ‘Cada vez lo tenemos más complicado’

Metroscopia, Sigma Dos, TNS, Feedback, Gesop... Es muy posible que todas estas empresas de análisis demoscópico te suenen porque en los últimos meses el interés de los españoles por las encuestas electorales se ha disparado. Cada vez que un medio de comunicación publica un sondeo, las redes sociales bullen, los comentarios en la calle se suceden y los partidos políticos tiemblan... o lo celebran. [...]

Con todo, cada sondeo viene acompañado casi siempre de acusaciones de manipulación y exceso de cocina. [...]

Lo que ocurre ahora, aseguran los expertos, es que su labor es mucho más complicada porque el panorama social, económico y político es muy “volátil”. “Hace 10 años, la estabilidad en el voto en España era muy grande. Se hacía una encuesta tres meses antes de las elecciones y los resultados eran muy similares. Ahora hay movimiento de una semana para otra y una encuesta se queda vieja pronto”, afirma De Elías. [...]

Un proceso subjetivo

[...] La cocina de las encuestas es un proceso subjetivo y los expertos admiten que, con los mismos datos brutos, dos analistas pueden dar finalmente dos resultados distintos. “Es verdad que algunas encuestas en España, a veces, parecen reflejar opiniones que encajan con el medio de comunicación que las publica, sobre todo cuando las elecciones están lejos y no hay una verificación cercana, pero lo más general no es eso, lo más general es que produzcan resultados muy parecidos entre sí”, señala Alberto Penadés, sociólogo y politólogo de la Universidad de Salamanca.

Penadés advierte de que, aunque en España no existe un consejo regulador estricto como en otros países, los llamados institutos de investigación se adhieren a unas reglas profesionales que, “aunque son más bien genéricas”, deberían ser suficientes para descartar fraudes o manipulaciones. [...]

La publicación

La última fase del proceso de una encuesta llega con la publicación de los datos por parte del medio de comunicación. Los expertos aseguran que nunca han recibido presiones de los medios para cambiar los datos, pero subrayan que en el titular y el texto que eligen los periodistas no tienen voz ni voto. “Mi experiencia es que hay mucha más manipulación en los titulares que hace cada medio que en la propia entrega de datos”, incide Michavila. [...]

Los expertos señalan que, con todo, hay algunos elementos que indican que una encuesta es seria y fiable o no tanto. “Siempre tiene que ir acompañada de una ficha técnica y es muy importante que se publiquen las preguntas para que se vea que no hay engaño”, explica José Pablo Ferrándiz, vicepresidente de Metroscopia. En esa ficha tiene que figurar

la fecha en que se hicieron las entrevistas, el universo (a quién han preguntado), el método (telefónico, domiciliario, *online*) y el tamaño de la muestra, que, por lo general, debe ser de al menos 1.000 personas. [...]

Rodrigo Carretero, *El Huffington Post*, 05/11/2014

¿Qué enseñó el 20-D a quienes hacen encuestas? 'Tenemos que tener cuidado'

[...] Algunos van más allá de las bromas y acusan a las empresas demoscópicas de manipular los datos para condicionar el voto de la gente. Y se apoyan en los datos: hubo varias encuestas que incluso situaban a Ciudadanos como segunda fuerza, solo por detrás del PP y con porcentajes de voto de hasta el 23,3%. Finalmente, la formación naranja fue cuarta con un 13,93%. Con Podemos pasó justo al contrario: los sondeos daban a los de Pablo Iglesias resultados sensiblemente peores de los que acabó logrando.

¿Por qué ocurrió eso? De entrada, las empresas demoscópicas niegan cualquier intencionalidad y señalan que se trata de algo más complejo que tiene que ver con aspectos técnicos y políticos. [...]

José Pablo Ferrándiz, investigador principal de Metroscopia, coincide en ese análisis e indica que últimamente los electores retrasan cada vez más la decisión de a quién votar. Eso se une a que los sondeos solo pueden reflejar la tendencia hasta cinco días antes de los comicios, puesto que luego está prohibido publicar encuestas. [...]

Pero la experiencia y los errores del 20-D sí que sirven a las empresas de sondeos para afinar más el tiro. José Ramón Lorente, analista principal de Celeste-Tel, insiste en que Podemos y Ciudadanos fueron los partidos que tenían un mayor porcentaje de voto procedente de nuevos electores y de “votantes intermitentes”, aquellos que no habían ido a las urnas en 2011 pero sí pensaban hacerlo en diciembre. Ese voto, afirma, fue muy difícil de contabilizar y de cuantificar. [...] “La última semana es el momento en el que hay que ampliar el tamaño de la muestra e intensificar la labor de análisis”, Sara Morais, responsable de investigación de GAD3. [...]

Por tanto, esa experiencia de diciembre ayuda en cierto modo a hacer ‘la cocina’ de los datos, un proceso que consiste en transformar las respuestas que las empresas obtienen al hacer las encuestas en una estimación de voto, que es el resultado final que publican los medios. En ese proceso subjetivo tratan, por ejemplo, de asignar a los indecisos un voto basándose en algunas variables como, por ejemplo, lo que recuerdan que votaron en elecciones anteriores, el resultado real de las mismas, la fidelidad tradicional del votante de cada partido... [...] “La clave va a estar en la participación, que es lo más difícil de estimar, pero que es el dato fundamental para dar el dato electoral”, asegura Ferrándiz. [...]

Rodrigo Carretero, *El Huffington Post*, 28/06/2016

El término “encuesta” se refiere al procedimiento utilizado para conseguir información, en forma de datos numéricos, de un grupo de sujetos que se pretende que represente a una población. La matemática permite controlar los márgenes de error de las afirmaciones realizadas a partir de esos datos.

Las encuestas pueden tener diferentes fines según quién las utilice: las empresas las usan para determinar si un producto gusta más o menos que otro a los consumidores; los investigadores en ciencias sociales, para conocer las opiniones de la sociedad y cómo van cambiando a lo largo del tiempo; los gobiernos y las instituciones públicas, para saber cuáles son las preferencias de los ciudadanos, las evaluaciones o los problemas de la sociedad, y poder a partir de ellas tomar decisiones o elaborar una política ajustada a las prioridades de los individuos. Cuando lo que se estudia a través de una encuesta son las ideas de un colectivo acerca de un determinado asunto, se llama “sondeo de opinión”, una técnica estadística de tipo sociológico usada para predecir la respuesta de la sociedad ante unas elecciones, por ejemplo.

Todos los sondeos tienen limitaciones, pues el entrevistado puede no ser sincero en sus respuestas por diferentes motivos; la encuesta puede contener errores en la redacción de las preguntas y otros factores que hacen que la información no

refleje necesariamente la realidad que se desea conocer. Así, para garantizar la fiabilidad y la validez de las encuestas, como técnica de investigación, se requiere que cada una de sus fases se lleve a cabo correctamente.

Fases de una encuesta

La interpretación de los resultados de una encuesta no puede hacerse sin tener en cuenta el proceso mediante el que se han obtenido. Para ello, conviene conocer en cada una de las fases cuáles son los requisitos que se deben cumplir y qué riesgos entrañan.

Las principales fases de una encuesta son las siguientes:

1. Diseño de la encuesta
2. Trabajo de campo
3. Análisis de los datos
4. Análisis de resultados
5. Publicación de resultados

Diseño de la encuesta

La planificación de una encuesta o sondeo de opinión requiere concretar previamente algunas ideas relativas a la población de la que se desea obtener información, el objetivo, los datos concernientes a esta población, la precisión que el resultado requiere (lo que equivale a tomar decisiones sobre el tamaño de la muestra y el modo en que se obtiene) y el propio diseño de la encuesta.

El objetivo de la investigación

La primera tarea en el diseño de una encuesta es definir adecuadamente los objetivos que se desean alcanzar. Dicho de otro modo, determinar los temas sobre los que se desean conocer las opiniones de los ciudadanos. Una manera de plantearlos es realizar un estudio y análisis de las diversas aportaciones

ya existentes. En ocasiones, para que los objetivos de la encuesta sean más claros, se plantea un objetivo general y algunos específicos.

La información que deseamos obtener de una encuesta o sondeo consistirá en un número o un conjunto de números que llamamos *valor real de la población*. Estos corresponderían a los valores de “los ingresos medios por familia”, “la proporción masculina de una actividad” o “la proporción de posibles votantes a cada partido político”. Los valores reales de la población habitualmente son una media de los valores reales para los miembros individuales de la población.

EJEMPLO 1

El Centro de Investigaciones Sociológicas (CIS) realizó una serie de estudios que denominó “Barómetro Autonomico”. Estableció los siguientes objetivos para esos estudios y que ilustrarán las ideas anteriores:

“La denominación de ‘Barómetro Autonomico’ se inició en 2005 (estudio 2.610), continuó en 2010 (estudio 2.829) y en 2012 (estudio 2.956). Estos tres estudios poseen como objetivos comunes:

- Realizar un seguimiento de la valoración que realizan los/as ciudadanos/as de distintos aspectos vinculados al desarrollo y funcionamiento del Estado de las Autonomías.
- Obtener información en torno a la valoración de los/as ciudadanos/as de cada comunidad autónoma sobre: la gestión de sus gobiernos, la actuación de la oposición, la labor de sus Presidentes/as autonómicos/as, o la valoración de diferentes políticas públicas.”

La población de la que se desea extraer la información

La *población*, en ocasiones también se denomina población objetivo, es el conjunto de todos los individuos (personas, animales, objetos, eventos...) que contienen información sobre un fenómeno que se desea estudiar. Los individuos que constituyen la población se denominan *elementos*. Un estudio sobre los

gustos de los turistas tiene elementos reales, los turistas; pero los elementos también pueden ser abstractos como la intención de voto en unas elecciones.

Es especialmente importante que la población esté bien definida. Para ello, la población debe dejar claro cuáles son los elementos que la constituyen. En términos electorales, “los ciudadanos españoles mayores de edad en 2017” podría parecer una población bien definida de los potenciales electores; pero si la consideramos minuciosamente, notaremos que ser español y mayor de edad en el año de la votación incluye algunos hechos que no se tienen en cuenta para un estudio adecuado. ¿En todas las elecciones se requiere ser español? ¿Ser mayor de edad en el año en que se celebran las elecciones significa que estará incorporado al censo electoral el día de las elecciones? ¿Todas las personas mayores de edad tienen derecho al voto? En todos los casos anteriores, la respuesta es negativa. Podríamos entonces mejorar la definición de la población anterior como “ciudadanos con derecho a voto en las elecciones generales del 2017 residentes en España”.

Ejercicio 2. Define la población objeto de estudio de la siguiente investigación:

“Un estudio elaborado por la escuela universitaria Eusa y la Cámara de Comercio pone de manifiesto que los estudiantes, en su mayoría, tienen una visión de su futuro en el mercado laboral que se aleja de la realidad, y que el deseo de ser funcionario sigue superando a la vocación emprendedora entre estos jóvenes. Son alumnos, entre 15 y 18 años, que piensan, casi la mitad, en un 48 por ciento, que tendrán un salario de 1.500 euros, con un contrato indefinido, y la mayoría, casi el 68 por ciento, cree que encontrará un trabajo en menos de un año.”

Natalia Perales, “Casi la mitad de los estudiantes de 15 a 18 años cree que tendrá un salario de 1.500 euros”, *Radio Sevilla*, 16/06/2016.

La muestra

Las grandes poblaciones son difíciles de estudiar. Por esa razón suele seleccionarse un subconjunto de la población al que se denomina *muestra*. Pero cuidado, al elegirla se corre el

riesgo de sesgar la información obtenida a partir de ella. El tamaño de la muestra es una de las causas fundamentales de dicho error como analizaremos más adelante.

También es importante el criterio con el que se elige la muestra. Por ejemplo, si la muestra se obtiene entre personas que han conseguido empleo en los últimos seis meses, ¿tendremos las mismas conclusiones que si obtenemos la muestra de las oficinas de empleo? Así, para que la muestra represente adecuadamente las características de la población, su elección será aleatoria, de manera que cada individuo tendrá la misma probabilidad de ser elegido. De no ser así, las conclusiones estarían sesgadas, desviadas de la realidad.

No solo se necesitan muestras cuando la población es demasiado grande. Hay otras situaciones, como cuando el análisis de los elementos conlleva su destrucción, que también lo requieren. Por ejemplo, al estudiar la calidad de un producto. Si un fabricante de bizcochos desea conocer el nivel de calidad de su producto, no tendría sentido abrir todos los paquetes, sino que bastaría con probar algunos, una muestra, para inferir el nivel de calidad de toda la producción con un cierto nivel de confianza. También son necesarias las muestras cuando los elementos de la población no se pueden controlar con facilidad, como ocurre, por ejemplo, en el estudio de poblaciones de animales en un determinado entorno. El estudio de muestras es también una solución en el caso de que se necesite rapidez en el conocimiento de la información.

En definitiva, el uso de muestras tiene como ventajas la posibilidad de reducir los gastos del estudio, la rapidez para obtener la información y además permite hacer estudios de poblaciones infinitas. Sin embargo, toda muestra, al no coincidir exactamente con las características de la población, lleva implícito un error en los estimadores calculados que puede invalidar cualquier encuesta. Este hecho y los fundamentos matemáticos que lo explican serán los protagonistas de este capítulo y el siguiente.

Antes de decidir el diseño de una encuesta entre las distintas alternativas, debemos determinar cuál de ellos es más eficiente. Esto a su vez implica la existencia de un conjunto de criterios que servirán para distinguir entre diseños buenos y malos. Algunos de los criterios en uso actualmente presentan ciertas limitaciones.

En el caso de que se desee estimar la superficie total de olivos en España o en una comunidad concreta, se dispone de muchos métodos de estimación. Uno de ellos sería solicitar a algunos expertos en el tema que emitan un juicio sobre el número de hectáreas de olivos. Esta técnica se conoce como método Delphi y es interesante cuando carecemos de datos estadísticos fiables. Sin embargo, si una decisión importante depende de la precisión de las estimaciones, un error puede resultar más costoso a largo plazo que un mayor gasto en el modelo de encuesta considerado.

En cualquier caso, el uso de encuestas tampoco garantiza que los resultados elaborados estén libres de sesgos ya que influyen un elevado número de factores que juegan un papel importante en el diseño de la muestra. Por ejemplo, el tamaño de la muestra, la estratificación de la población de estudio o el método de estimación, entre otros.

Así, en ocasiones encontramos investigaciones que usan llamadas telefónicas o cuestionarios *online* con una proporción de respuestas inadecuadas. La primera limitación de este tipo de encuestas es el propio listado de teléfonos o correos electrónicos. La segunda limitación está en que la proporción de personas que no responden (tasa de no respuesta) puede ser alta, y no se puede asumir que estas tienen características similares a las de quienes respondieron.

Estas limitaciones de las encuestas telefónicas u *online* han dado fuerza a lo que se llama método de encuestas de campo. Aquí los entrevistadores seleccionan áreas del país y gestionan el tiempo y el esfuerzo en establecer controles para asegurarse

de que obtienen una muestra que es una buena “fotografía” de la población. Sin embargo, a menudo encontramos procedimientos en los cuales el entrevistador tiene una considerable autonomía, de forma que su criterio personal puede influir en la selección de la muestra.

Las distintas formas de seleccionar a los individuos que se van a encuestar se denominan técnicas de muestreo y las estudiaremos con más profundidad en el próximo capítulo.

Diseño del cuestionario y trabajo de campo

El *diseño del cuestionario* con el que se recoge la información tiene una gran influencia en el resultado de la encuesta. Los cuestionarios contienen un conjunto de preguntas cuyo fin es recopilar las opiniones de los ciudadanos respecto a diferentes temas sobre los que se desea investigar. Incluyen datos sobre los entrevistados que para la investigación pueden ser pertinentes, como sexo, edad, profesión... El diseño de las preguntas se realiza con un cuidado especial, puesto que todas ellas han de ser comprendidas por los entrevistados. Además, es aconsejable establecer tanto el número de preguntas como el orden en el que van a aparecer. Estas preguntas han de ser simples y directas, intentando siempre evitar sesgos.

El trabajo de campo consiste en distribuir el cuestionario a las personas que conforman la muestra, siguiendo rutas aleatorias y contactando con los ciudadanos, ya sea en sus domicilios, por teléfono o internet. Hay factores como la duración de la encuesta, el momento temporal en que se realiza o la competencia de los entrevistadores que pueden afectar a los resultados de la encuesta. Para intentar evitarlo, las encuestas suelen incluir el espacio y el tiempo en el que se deben realizar los cuestionarios y además, instrucciones muy precisas para los entrevistadores.

EJEMPLO 2

Aquí podemos ver parte de un cuestionario elaborado por el Centro de Investigaciones Sociológicas para un sondeo electoral:

P.6 Como Ud. seguramente sabe, el próximo 26 de junio se celebrarán elecciones generales, es decir, al Parlamento español. ¿Piensa Ud. ir a votar en estas elecciones? **(MOSTRAR TARJETA A)**

- Sí, con toda seguridad	1	
- Probablemente sí	2	(51)
- Probablemente no	3	
- No, con toda seguridad	4	→ Pasar a P.7
- (NO LEER) Aún no lo tiene decidido. .	8	
- N.C.	9	

P.6a ¿Y podría decirme si ya tiene Ud. decidido su voto?

- Sí, lo tiene decidido	1	→ Pasar a P.7
- No, no lo ha decidido aún	2	(52)
- N.C.	9	→ Pasar a P.7

P.6b ¿Y entre qué partidos u opciones duda Ud.?
(ENTREVISTADOR/A: ESPERE RESPUESTA ESPONTÁNEA Y MARQUE LAS DOS OPCIONES QUE SEÑALE LA PERSONA ENTREVISTADA, CADA UNA DE ELLAS EN UNA COLUMNA).

Análisis de los datos

Finalizado el trabajo de campo, se procesa toda la información recopilada a partir de las respuestas de los cuestionarios, la cual se recoge en una hoja de cálculo para su posterior análisis, situando en las filas de la hoja los números correspondientes a los cuestionarios y en las columnas, las preguntas que se presentan en el cuestionario.

Publicación de los resultados

La ley obliga a publicar la ficha técnica de cada encuesta con el fin de garantizar el derecho de las personas que quieren conocer sus características científicas y la interpretación correcta de los datos que se publican. Esta ficha técnica incluye al menos los siguientes datos técnicos: los responsables de la realización de la encuesta, la entidad que se encarga de hacerla, el sistema de muestreo, el tamaño de la muestra, el

error muestral, el nivel de representatividad, el procedimiento de selección de los encuestados, la fecha de realización del trabajo de campo, el texto íntegro de las cuestiones planteadas y el número de personas que no sabe o no contesta a las cuestiones planteadas.

El tamaño de la muestra

Ya vimos que la concreción del tamaño de la muestra es una de las cuestiones a definir en el diseño de la encuesta. Sin embargo, merece la pena profundizar en una pregunta: ¿por qué una muestra grande es mejor predictor que una muestra pequeña? Con el único fin de argumentar planteando un problema claro, vamos a suponer que el 60% de una población está a favor de la cadena perpetua. Una muestra de dos personas únicamente equivaldría a sacar dos bolas de una urna enorme que contenga un 60% de bolas rojas (a favor de la cadena perpetua) y 40% negras (opuestas a la cadena perpetua). La probabilidad de cada resultado se interpreta como el número de personas a favor o en contra dividido entre el total de la población.

Los resultados posibles de la extracción son los pares ordenados {NN, NR, RN, RR} y la probabilidad asociada a cada uno de ellos es el resultado de multiplicar las probabilidades de cada uno de los sucesos individuales. De esta forma, podemos construir una tabla en la que se reflejen estas probabilidades:

Suceso	% de rojas	Probabilidad
{NN}	0	$0,4 \cdot 0,4 = 0,16$
{NR, RN}	50	$0,4 \cdot 0,6 + 0,6 \cdot 0,4 = 0,48$
{RR}	100	$0,6 \cdot 0,6 = 0,36$

Los resultados se interpretan de la siguiente manera: si encuestamos a un gran número de muestras aleatorias de dos

personas, alrededor del 16% de esas muestras se opondría a la cadena perpetua, otro 36% la defendería y el 48% estaría dividido sobre esta cuestión. Si aumentásemos el tamaño de la muestra, diferentes tablas similares a la anterior reflejarían cómo evoluciona el porcentaje de muestras y nos sirven como indicador razonable para la investigación. Así, para el caso de una muestra de tamaño 20 personas, la tabla quedaría:

Suceso (% de rojas)	Probabilidad (%)
0 - 35	0
35 - 40	2
40 - 45	4
45 - 50	7
50 - 55	12
55 - 60	16
60 - 65	18
65 - 70	17
70 - 75	12
75 - 80	7
80 - 85	3
85 - 90	1
90 - 100	0

Ahora más de tres cuartos de las muestras de tamaño 20 reflejan un porcentaje de bolas rojas entre el 50% y el 75%, es decir, son estimadores razonables del actual porcentaje de la población favorable a la cadena perpetua (60%). El modo en que se distribuyen estas muestras se denomina *distribución muestral*.

Ahora podemos extraer algunas conclusiones cuantitativas del muestreo. En este ejemplo de la cadena perpetua, con el tamaño de muestra $n = 100$, la distribución muestral se asemeja a una curva normal con media $f = 56\%$ (porcentaje a favor de la pena de muerte en una muestra de tamaño 100) y

la dispersión de los datos en torno a esta media, la *desviación estándar*, se obtiene:

$$\sigma = \sqrt{\frac{f(1-f)}{n}} = \sqrt{\frac{0,56 \cdot 0,44}{100}} = 0,496, \text{ casi } 5\%$$

Una de las propiedades de la curva normal es que el 68% de los valores se encuentra en el intervalo probabilístico $[0'56-0'05, 0'56+0'05]$, el 95,44% de los valores en el intervalo $[0'56-2 \cdot 0'05, 0'56+2 \cdot 0'05]$ y el 99,74% de los valores en el intervalo $[0'56-3 \cdot 0'05, 0'56+3 \cdot 0'05]$.

Para nuestro ejemplo, de todas las muestras de tamaño 100, esto significa:

El 68% de las veces que se repita el experimento (o la consulta), la probabilidad de obtener bola roja (o porcentaje de la población favorable a la pena de muerte) estará entre el 51% y el 61%.

El 95% de las veces que repita el experimento, la probabilidad de obtener bola roja estará entre el 46% y el 66%. El 98% de las veces que repita el experimento, la probabilidad de obtener bola roja estará entre el 41% y el 71%.

Para la redacción de las conclusiones, suelen utilizarse dos conceptos clave: la probabilidad de que el valor real de la variable esté en un intervalo concreto, el *nivel de confianza*; y el *margen de error* que indica el error cometido al hacer una estimación o, dicho de otro modo, entre qué cantidades se encuentra el valor de la variable estudiada con el nivel de confianza dado.

La conclusión sobre la población que se deduce de nuestra muestra se escribe de varias maneras:

- a) Tenemos el 95% de confianza de que el porcentaje de votantes favorables a la cadena perpetua está entre el 46% y el 66%.
- b) Con un *nivel de confianza* del 95%, el porcentaje de votantes favorable a la cadena perpetua es del 56%, con un *margen de error* de $\pm 10\%$.

Ejercicio 3. Los elementos de una muestra aleatoria de 1.000 estudiantes de secundaria votaron 740 frente a 260 a favor de suprimir los deberes en casa. ¿Qué se puede inferir sobre los sentimientos de los estudiantes de secundaria sobre este tema?

Ejercicio 4. Supongamos que el Ministerio de Hacienda selecciona aleatoriamente 600 declaraciones de impuestos y encuentra que 240 de estas contienen errores. Con el 95% del nivel de confianza, ¿qué porcentaje de todas las declaraciones contienen errores?

Ejercicio 5. En una muestra de 1.574 votantes, el 60% afirmó que votaría por el candidato Amat para presidente. Los encuestadores indicaron que el error muestral era de 4 puntos. ¿Están en lo cierto?

A medida que aumentamos el número de elementos de la muestra, el error muestral disminuye. Pero el estudio anterior también nos dice que el error muestral no disminuye en la misma proporción a cómo se aumenta el número de elementos de la muestra. Una precisión mayor, exigiría un aumento considerable de la muestra. Por ejemplo, multiplicar por 4 el tamaño de la muestra divide por 2 la desviación estándar.

En las encuestas y sondeos electorales suele utilizarse una cifra entre 1.000 y 2.500 que se considera que representa relativamente bien a la población. Con errores entre 3 y 2 puntos, para un nivel de confianza del 95,5%.

Ejercicio 6. Una encuesta de opinión utiliza una muestra de tamaño 1.500.
a) Para un nivel de confianza del 95%, ¿cuál es el intervalo de confianza para una encuesta en la que se han recibido un 35% de respuestas afirmativas?
b) ¿A cuántas personas tendrían que entrevistar para obtener un margen de error del 1% con un 95% de confianza?

Solución:

- a) Utilizando la expresión anterior, para $f = 0,35$, la desviación estándar es 1,2% y el intervalo de confianza para el 95% es $35\% \pm 2,4\%$.
b) Para tener un nivel de confianza del 95% con un error del 1%, el intervalo sería $35\% \pm 1\%$. Hay que calcular n a partir de la expresión

$$\sqrt{\frac{0,35 \cdot 0,65}{n}} = 0,5\% = 0,005$$

De ahí nos daría como resultado $n = 9.100$ personas encuestadas.

Ejercicio 7. Un reportaje indica que un sondeo de una muestra aleatoria de votantes revela, en el 95% de nivel de confianza, que la candidata Tavares obtendrá entre el 51% y el 59% de los votos.

- a) ¿Qué porcentaje de los votantes de la muestra dijeron que votarían por Tavares?
- b) Tavares está convencida de que ganará las elecciones, según el sondeo, ¿qué piensas tú?
- c) ¿Cuántos votantes había en la muestra?