Soluciones - Muestreo e intervalos de confianza

Curso: 2 Bachillerato Ccss

Solución del ejercicio 1

a. Muestreo aleatorio simple: todos los vecinos tienen la misma probabilidad de ser elegidos y se eligen por sorteo.
b. Muestreo sistemático: el salto es $k = \dfrac{48\,000}{400} = 120$; tras el arranque aleatorio se toma uno de cada 120.
c. Muestreo estratificado con afijación proporcional: los estratos son los distritos y cada uno aporta según su peso en la población.
d. Muestreo por conglomerados: se sortean grupos completos (las comunidades) y se estudian todos sus miembros.

Solución del ejercicio 2

a. Total: $300 + 280 + 260 + 220 + 140 = 1\,200$ alumnos. Fracción de muestreo: $\dfrac{60}{1\,200} = 0{,}05$, es decir, el 5 % de cada curso.
b. 1.º ESO: $300 \cdot 0{,}05 = 15$; 2.º ESO: $280 \cdot 0{,}05 = 14$; 3.º ESO: $260 \cdot 0{,}05 = 13$; 4.º ESO: $220 \cdot 0{,}05 = 11$; Bachillerato: $140 \cdot 0{,}05 = 7$.
c. $15 + 14 + 13 + 11 + 7 = 60$ alumnos.
d. Garantiza que todos los cursos están representados en la misma proporción que en el instituto; con un sorteo simple podría salir, por azar, casi ningún alumno de Bachillerato.

Solución del ejercicio 3

a. Si la población es normal, la media muestral también lo es, con la misma media y desviación típica dividida entre $\sqrt{n}$:
$\bar{X} \sim N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right) = N\left(50, \dfrac{12}{\sqrt{36}}\right) = N(50;\, 2)$.
b. Tipificamos con la desviación de la media muestral: $z = \dfrac{53 - 50}{2} = 1{,}5$.
$P(\bar{X} > 53) = P(Z > 1{,}5) = 1 - 0{,}9332 = 0{,}0668$.
c. $z_1 = \dfrac{48 - 50}{2} = -1$ y $z_2 = \dfrac{52 - 50}{2} = 1$. $P(-1$.
d. Al promediar 36 datos, los valores altos y bajos se compensan, y la media varía menos que cada dato: la dispersión se divide por $\sqrt{36} = 6$. Cuanto mayor es la

muestra, más se concentra $\bar{X}$ alrededor de μ.

Solución del ejercicio 4

a. Para usar la aproximación normal se necesita $np \geq 5$ y $nq \geq 5$: aquí $np = 150 \cdot 0{,}4 = 60$ y $nq = 90$. Entonces $\hat{p}$ sigue aproximadamente $N\left(p, \sqrt{\frac{pq}{n}}\right) = N\left(0{,}4;\ \sqrt{\frac{0{,}4 \cdot 0{,}6}{150}}\right) = N(0{,}4;\ 0{,}04)$.

b. Tipificamos: $z = \frac{0{,}46 - 0{,}4}{0{,}04} = 1{,}5$.
$P(\hat{p} > 0{,}46) = P(Z > 1{,}5) = 1 - 0{,}9332 = 0{,}0668$. Solo un 6,7 % de las muestras daría más del 46 %.

c. $z = \frac{0{,}36 - 0{,}4}{0{,}04} = -1$. Por simetría, $P(Z$: aproximadamente una de cada seis muestras daría menos del 36 %.

Solución del ejercicio 5

a. El área α que queda fuera se reparte por igual entre las dos colas, $\frac{\alpha}{2}$ en cada una. Por tanto $P(Z$.

b. 90 %: $\alpha = 0{,}10$, buscamos $0{,}95$ y obtenemos $z = 1{,}645$. 95 %: $\alpha = 0{,}05$, buscamos $0{,}975$ y obtenemos $z = 1{,}96$.

c. 99 %: buscamos $0{,}995$, que da $z = 2{,}575$. 98 %: buscamos $0{,}99$, que da $z \approx 2{,}33$.

Solución del ejercicio 6

a. $E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 1{,}96 \cdot \frac{4}{\sqrt{64}} = 1{,}96 \cdot 0{,}5 = 0{,}98$ minutos.

b. $(24{,}5 - 0{,}98;\ 24{,}5 + 0{,}98) = (23{,}52;\ 25{,}48)$ minutos.

c. Con una confianza del 95 %, el tiempo medio de todos los empleados está entre 23,52 y 25,48 minutos. El 95 % de las muestras posibles produciría un intervalo que contiene la media real.

Solución del ejercicio 7

a. $E = 1{,}96 \cdot \frac{6}{\sqrt{100}} = 1{,}96 \cdot 0{,}6 = 1{,}176$. Intervalo: $(17{,}124;\ 19{,}476)$ minutos.

b. $E = 2{,}575 \cdot 0{,}6 = 1{,}545$. Intervalo: $(16{,}755;\ 19{,}845)$ minutos.

c. El del 99 %: para estar más seguros de capturar la media real hay que dar un intervalo más amplio. Más confianza implica menos precisión.

d. No: 20 queda fuera incluso del intervalo al 99 %, cuyo extremo superior es 19,845. Los datos indican que la espera media es menor de lo que afirma la dirección.

Solución del ejercicio 8

a. $\hat{p} = \dfrac{240}{400} = 0{,}6$ y $\hat{q} = 0{,}4$.

b. $E = 1{,}96\sqrt{\dfrac{0{,}6 \cdot 0{,}4}{400}} = 1{,}96 \cdot 0{,}02449 \approx 0{,}0480$.

c. $(0{,}6 - 0{,}048;\ 0{,}6 + 0{,}048) = (0{,}552;\ 0{,}648)$, es decir, entre el 55,2 % y el 64,8 %.

d. Sí: todo el intervalo está por encima de 0,5, así que con una confianza del 95 % la proporción supera el 50 %.

Solución del ejercicio 9

a. El intervalo está centrado en la media muestral: $\bar{x} = \dfrac{12{,}4 + 13{,}6}{2} = 13$ horas.

b. El error máximo es la mitad de la longitud del intervalo: $E = \dfrac{13{,}6 - 12{,}4}{2} = 0{,}6$ horas.

c. Como $E = z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$ y $\sqrt{81} = 9$, queda $0{,}6 = z \cdot \dfrac{3}{9}$, de donde $z = 0{,}6 \cdot 3 = 1{,}8$.

d. $P(Z$, así que $\dfrac{\alpha}{2} = 0{,}0359$ y $\alpha = 0{,}0718$. El nivel de confianza es $1 - \alpha = 0{,}9282$, alrededor del 92,8 %.

Solución del ejercicio 10

a. De $E = z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$ se despeja $n \geq \left(\dfrac{z_{\alpha/2}\,\sigma}{E}\right)^2$.

b. $n \geq \left(\dfrac{1{,}96 \cdot 15}{2}\right)^2 = 14{,}7^2 = 216{,}09$. Hacen falta al menos 217 familias.

c. $n \geq \left(\dfrac{2{,}575 \cdot 15}{2}\right)^2 = 19{,}3125^2 \approx 372{,}97$. Hacen falta al menos 373 familias.

d. Con 216 familias el error sería algo mayor que 2 euros. Para garantizar que $E \leq 2$ hay que llegar al entero siguiente.

Solución del ejercicio 11

a. Se toma $\hat{p} = 0{,}5$, porque el producto $\hat{p}(1-\hat{p})$ alcanza su máximo, $0{,}25$, en ese valor: así el tamaño calculado sirve para cualquier proporción real.
b. $n \geq \dfrac{1{,}96^2 \cdot 0{,}5 \cdot 0{,}5}{0{,}03^2} = \dfrac{0{,}9604}{0{,}0009} \approx 1\,067{,}1$. Al menos 1 068 personas.
c. $n \geq \dfrac{1{,}96^2 \cdot 0{,}2 \cdot 0{,}8}{0{,}03^2} = \dfrac{0{,}614656}{0{,}0009} \approx 682{,}95$. Al menos 683 personas.
d. Con información previa se ahorran $1\,068 - 683 = 385$ encuestas: conocer aproximadamente p abarata mucho el estudio.

Solución del ejercicio 12

a. Al 90 % el valor crítico es $z_{\alpha/2} = 1{,}645$. El error máximo es $E = 1{,}645 \cdot \dfrac{50}{\sqrt{25}} = 1{,}645 \cdot 10 = 16{,}45$ ciclos, y el intervalo, $(820 - 16{,}45;\ 820 + 16{,}45) = (803{,}55;\ 836{,}45)$ ciclos.
b. No es creíble con esta confianza: el valor anunciado, 850, queda fuera del intervalo, por encima del extremo superior. Las baterías parecen durar menos de lo anunciado.
c. Al 99 %, $E = 2{,}575 \cdot 10 = 25{,}75$ y el intervalo es $(794{,}25;\ 845{,}75)$. Aunque es más ancho, tampoco contiene 850, así que la conclusión se mantiene.

Solución del ejercicio 13

a. Como el peso de cada paquete es normal, la media de 16 paquetes también lo es, con desviación $\dfrac{\sigma}{\sqrt{n}}$: $\bar{X} \sim N\left(1\,000;\ \dfrac{20}{\sqrt{16}}\right) = N(1\,000;\ 5)$.
b. Tipificamos: $z = \dfrac{990 - 1\,000}{5} = -2$. Por simetría, $P(\bar{X}$.
c. $z_1 = -2$ y $z_2 = \dfrac{995 - 1\,000}{5} = -1$. Restamos áreas: $P = \Phi(-1) - \Phi(-2) = 0{,}1587 - 0{,}0228 = 0{,}1359$.
d. Buscamos c con $P(\bar{X}$, es decir, $z = -1{,}645$. Destipificando, $c = 1\,000 - 1{,}645 \cdot 5 = 991{,}775$ g. Si la media de la muestra baja de unos 991,8 g, el inspector detiene la máquina.

Solución del ejercicio 14

a. Si $p = 0{,}35$, con $n = 500$ se cumple $np \geq 5$ y $nq \geq 5$, así que $\hat{p} \sim N\left(0{,}35;\ \sqrt{\dfrac{0{,}35 \cdot 0{,}65}{500}}\right) = N(0{,}35;\ 0{,}02133)$.
b. $z = \dfrac{0{,}38 - 0{,}35}{0{,}02133} \approx 1{,}41$; $P(Z > 1{,}41) = 1 - 0{,}9207 = 0{,}0793$.

c. 0,32 está a la misma distancia de la media por debajo: $z \approx -1{,}41$ y, por simetría, la probabilidad también es $0{,}0793$.

d. $z = \dfrac{0{,}40 - 0{,}35}{0{,}02133} \approx 2{,}34$, y $P(Z > 2{,}34) = 1 - 0{,}9904 = 0{,}0096$. Un resultado así ocurriría menos del 1 % de las veces si nada hubiera cambiado, así que seguramente la intención de voto ha subido.

Solución del ejercicio 15

a. Marcas: 15, 25, 35, 45 y 55.
$\bar{x} = \dfrac{15 \cdot 6 + 25 \cdot 14 + 35 \cdot 18 + 45 \cdot 9 + 55 \cdot 3}{50} = \dfrac{90 + 350 + 630 + 405 + 165}{50} = \dfrac{1\,640}{50} = 32{,}8$ euros.

b. $E = 1{,}96 \cdot \dfrac{8}{\sqrt{50}} = 1{,}96 \cdot 1{,}1314 \approx 2{,}22$ euros. Intervalo: $(30{,}58;\ 35{,}02)$ euros.

c. $n \geq \left(\dfrac{1{,}96 \cdot 8}{1}\right)^2 = 15{,}68^2 = 245{,}86$: al menos 246 estudiantes.

d. $n \geq \left(\dfrac{1{,}96 \cdot 8}{0{,}5}\right)^2 = 31{,}36^2 = 983{,}45$: al menos 984 estudiantes, unas cuatro veces más que para $E = 1$.

e. Como $E = z\dfrac{\sigma}{\sqrt{n}}$, si n se multiplica por 4, $\sqrt{n}$ se multiplica por 2 y el error se divide entre 2. La precisión cuesta cara: cada mitad de error cuadruplica las encuestas.

Solución del ejercicio 16

a. Fracción de muestreo $\dfrac{360}{4\,500} = 0{,}08$. Economía: $2\,000 \cdot 0{,}08 = 160$; Derecho: $1\,500 \cdot 0{,}08 = 120$; Educación: $1\,000 \cdot 0{,}08 = 80$.

b. Economía $\dfrac{96}{160} = 0{,}6$; Derecho $\dfrac{66}{120} = 0{,}55$; Educación $\dfrac{54}{80} = 0{,}675$. Total: $\dfrac{96 + 66 + 54}{360} = \dfrac{216}{360} = 0{,}6$.

c. $E = 1{,}96\sqrt{\dfrac{0{,}6 \cdot 0{,}4}{360}} = 1{,}96 \cdot 0{,}0258 \approx 0{,}0506$. Intervalo: $(0{,}549;\ 0{,}651)$.

d. Sí, con una confianza del 95 %: el extremo inferior, 0,549, ya supera 0,5. Aunque la proporción real fuera la más baja del intervalo, seguiría habiendo mayoría de satisfechos.

Solución del ejercicio 17

a. $E_A = 1{,}96 \cdot \dfrac{40}{\sqrt{64}} = 1{,}96 \cdot 5 = 9{,}8$. Intervalo: $(175{,}2;\ 194{,}8)$ minutos.

b. $E_B = 1{,}96 \cdot \dfrac{40}{\sqrt{256}} = 1{,}96 \cdot 2{,}5 = 4{,}9$. Intervalo: $(173{,}1;\ 182{,}9)$ minutos.

c. Porque la muestra de B es cuatro veces mayor: el error se divide entre $\sqrt{4} = 2$.

d. No: los intervalos se solapan entre 175,2 y 182,9 minutos. Un valor como 180 minutos es compatible con los dos estudios; la diferencia entre medias se explica por el azar del muestreo.

e. $n \geq \left(\dfrac{2{,}575 \cdot 40}{2}\right)^2 = 51{,}5^2 = 2\,652{,}25$: al menos 2 653 adolescentes.

Solución del ejercicio 18

a. $\hat{p} = \dfrac{28}{400} = 0{,}07$, es decir, el 7 %.

b. $E = 1{,}96\sqrt{\dfrac{0{,}07 \cdot 0{,}93}{400}} = 1{,}96 \cdot 0{,}012757 \approx 0{,}025$. Intervalo: $(0{,}045;\ 0{,}095)$.

c. No la contradice: 0,05 está dentro del intervalo, aunque cerca del extremo inferior.

d. $\hat{p} \sim N\left(0{,}05;\ \sqrt{\dfrac{0{,}05 \cdot 0{,}95}{400}}\right) = N(0{,}05;\ 0{,}010897)$. $z = \dfrac{0{,}07 - 0{,}05}{0{,}010897} \approx 1{,}84$ y $P(Z \geq 1{,}84) = 1 - 0{,}9671 = 0{,}0329$.

e. El resultado es poco probable (un 3,3 %) si la garantía se cumple, pero el intervalo al 95 % todavía incluye el 5 %. Lo prudente es pedir una muestra mayor antes de reclamar: con más piezas el intervalo se estrecharía y la conclusión sería más clara.

Solución del ejercicio 19

a. $\bar{x} = \dfrac{172{,}04 + 175{,}96}{2} = 174$ cm y $E = \dfrac{175{,}96 - 172{,}04}{2} = 1{,}96$ cm.

b. $1{,}96 = 1{,}96 \cdot \dfrac{10}{\sqrt{n}}$, luego $\sqrt{n} = 10$ y $n = 100$ varones.

c. Longitud 3 significa $E = 1{,}5$: $1{,}5 = z \cdot \dfrac{10}{10}$, $z = 1{,}5$. $P(Z$, $\alpha = 0{,}1336$ y el nivel es del 86,64 %.

d. Hace falta $E = 1$: $n \geq \left(\dfrac{1{,}96 \cdot 10}{1}\right)^2 = 384{,}16$, es decir, una nueva muestra de al menos 385 varones.

Solución del ejercicio 20

a. $p(1-p) = p - p^2 = \frac{1}{4} - \left(p - \frac{1}{2}\right)^2 \leq \frac{1}{4}$, con igualdad solo si $p = \frac{1}{2}$.

b. $\frac{2\,000}{3} \approx 666{,}7$: como máximo 666 entrevistas.

c. En el peor caso $\hat{p} = 0{,}5$: $E = 1{,}96\sqrt{\frac{0{,}25}{666}} = 1{,}96 \cdot 0{,}019375 \approx 0{,}038$, un 3,8 %.

d. $E = 1{,}645 \cdot 0{,}019375 \approx 0{,}0319$, un 3,2 % aproximadamente: se gana precisión a cambio de menos confianza.

e. $n \geq \frac{1{,}96^2 \cdot 0{,}25}{0{,}02^2} = 2\,401$ entrevistas, que cuestan $2\,401 \cdot 3 = 7\,203$ euros.