13 Intervalos de confianza

EJERCICIOS PROPUESTOS

1. **Ejercicio resuelto.**

2. **El peso de los huevos de una granja sigue una ley normal de media desconocida y desviación típica 1,23 gramos. Para estimar la media poblacional se ha tomado una muestra de dos docenas de huevos que han dado un peso total de 1615,2 gramos. Halla un intervalo de confianza, al 96 %, para la media poblacional.**

Dado que la variable *X*: "peso en gramos de huevos de la granja" tiene una distribución $N(\mu,\sigma)=N(\mu;1{,}23)$, la media muestral de las dos docenas de huevos, también sigue una distribución normal con:

$$\overline{X}\sim N\left(\mu;\sigma_{\overline{X}}^2=\frac{1{,}23^2}{24}\right)=N\left(\mu;\sigma_{\overline{X}}^2=0{,}06304\right)$$

Si dos docenas de huevos han pesado 1615,2 gramos, la media de la muestra es $\overline{x}=\frac{1615{,}2}{24}=67{,}3$ g.

Como $1-\alpha=0{,}96$, resulta $\frac{\alpha}{2}=0{,}02$. De las tablas de la *N*(0,1), se obtiene que:

$$P\left(Z<z_{0{,}02}\right)=1-0{,}02=0{,}98\Rightarrow z_{0{,}02}=2{,}055$$

De manera que el intervalo de confianza al 96 % para la media poblacional μ, en gramos, es:

$$IC_{0{,}97}(\mu)=\left(67{,}3-2{,}055\,\frac{1{,}23}{\sqrt{24}};\ 67{,}3+2{,}055\,\frac{1{,}23}{\sqrt{24}}\right)=\left(66{,}784;\ 67{,}816\right)$$

3. **Los precios en euros de un producto se distribuyen según una normal de desviación típica 15. Se ha tomado una muestra de los precios de dicho producto en 9 comercios elegidos al azar en un barrio de una ciudad, y se han encontrado los siguientes precios:**

195, 208, 238, 212, 199, 206, 225, 201, 215

Determina el intervalo de confianza, al 90 %, para el precio medio de este producto.

La variable aleatoria *X*: "precio, en euros, del producto" tiene una distribución $N(\mu,\sigma=15)$.

La muestra aleatoria del precio del producto en 9 tiendas tiene una media muestral de:

$$\overline{x}=\frac{195+208+238+212+199+206+225+201+215}{9}=211\text{ €}$$

Como $1-\alpha=0{,}90$, resulta $\frac{\alpha}{2}=0{,}05$. De las tablas de la *N*(0,1), se obtiene que:

$$P\left(Z<z_{0{,}05}\right)=1-0{,}05=0{,}95\Rightarrow z_{0{,}05}=1{,}645$$

De manera que el intervalo de confianza al 90% para el precio medio del producto μ es:

$$IC_{0{,}90}(\mu)=\left(211-1{,}645\,\frac{15}{\sqrt{9}};\ 211+1{,}645\,\frac{15}{\sqrt{9}}\right)=\left(202{,}78;\ 219{,}23\right)$$

4. **Se supone que el número de telespectadores (en millones) de un programa semanal de televisión se aproxima a una distribución normal de desviación típica 0,5 (millones). Si una muestra aleatoria de 10 semanas proporciona una media muestral de 6,54 millones de telespectadores, calcula un intervalo al 95 % de confianza para la media semanal de telespectadores de ese programa.**

La variable aleatoria *X*: "número, en millones, de telespectadores del programa" tiene una distribución:

$$N(\mu;\sigma = 0{,}5)$$

La muestra aleatoria de 10 semanas tiene una media muestral $\overline{x} = 6{,}54$ millones de telespectadores.

Como $1-\alpha = 0{,}95$, resulta $\frac{\alpha}{2} = 0{,}025$. De las tablas de la *N*(0,1), se obtiene que:

$$P\left(Z < z_{0,025}\right) = 1-0{,}025 = 0{,}975 \Rightarrow z_{0,025} = 1{,}96$$

De manera que el intervalo de confianza al 95% para la media semanal μ, en millones, es:

$$IC_{0,95}(\mu) = \left(6{,}54-1{,}96\frac{0{,}5}{\sqrt{10}};\ 6{,}54+1{,}96\frac{0{,}5}{\sqrt{10}}\right) = (6{,}23;6{,}85)$$

5. **El tiempo de espera para ser atendido en un cierto establecimiento se puede aproximar por una variable aleatoria con distribución normal de media μ desconocida y desviación típica igual a 3 minutos. Se toma una muestra aleatoria simple de tamaño 121.**

a) Determina un intervalo de confianza con un nivel del 95 % para μ, si la media de la muestra es igual a 7 minutos.

b) Calcula la probabilidad de que el valor absoluto de la diferencia entre la media de la muestra y μ sea mayor que 0,5 minutos.

La variable aleatoria *X*: "tiempo de espera, en minutos" tiene una distribución $N(\mu,\sigma = 3)$.

a) La muestra aleatoria de tamaño 121 proporciona una media muestral $\overline{x} = 7$ minutos.

Como $1-\alpha = 0{,}95$, resulta $\frac{\alpha}{2} = 0{,}025$. De las tablas de la *N*(0,1), se obtiene que:

$$P\left(Z < z_{0,025}\right) = 1-0{,}025 = 0{,}975 \Rightarrow z_{0,025} = 1{,}96$$

De manera que el intervalo de confianza al 95% para la media semanal μ, en minutos, es:

$$IC_{0,95}(\mu) = \left(7-1{,}96\frac{3}{\sqrt{121}};\ 7+1{,}96\frac{3}{\sqrt{121}}\right) = (6{,}5;7{,}5)$$

b) La distribución de la media muestral es también normal:

$$\overline{X} \sim N\left(\mu_{\overline{x}},\sigma^2_{\overline{x}}\right) = N\left(\mu,\frac{9}{121}\right)$$

Para calcular la probabilidad de que $\left|\overline{X}-\mu\right| > 0{,}5$, se calcula la probabilidad de que sea menor que 0,5 minutos, y después la del suceso contrario. Por lo que, desarrollando y tipificando a la $Z \sim N(0{,}1)$:

$$P\left(\left|\overline{X}-\mu\right| < 0{,}5\right) = P\left(-0{,}5 < \overline{X}-\mu < 0{,}5\right) = P\left(\frac{-0{,}5}{3} < Z < \frac{0{,}5}{3}\right) = 2\Phi(0{,}17)-1 = 2\cdot 0{,}5675-1 = 0{,}135$$

Y, por último: $P\left(\left|\overline{X}-\mu\right| > 0{,}5\right) = 1-P\left(\left|\overline{X}-\mu\right| < 0{,}5\right) = 1-0{,}135 = 0{,}865$

6. **Ejercicio resuelto.**

7. **En una población se desea conocer la proporción de personas de estatura superior a 190 cm. Para ello se elige una muestra aleatoria de 600 personas, de las que 60 miden más de 190 cm. Para la proporción de personas que miden más de 190 cm en esta población:**

a) Determina un estimador puntual.

b) Calcula un intervalo de confianza al 99 %.

c) Halla el nivel de confianza para el intervalo (0,076 ; 0,124)

Sea X: "número de personas, de las 600, que miden más de 190 cm". La variable X tiene distribución binomial $B(n=600,\ p)$, siendo p la probabilidad de que una persona elegida al azar, en esa población, mida más de 190 cm.

a) Un estimador puntual para la proporción p, es la proporción muestral: $\hat{p} = \frac{60}{600} = 0{,}1$

b) La distribución de la proporción muestral $\hat{p}$ se puede aproximar por la de una distribución normal, dado que el tamaño de la muestra es suficientemente grande:

$$\hat{p} \sim N\left(\mu = p; \sigma^2 = \frac{p(1-p)}{600}\right)$$

Aproximando p por $\hat{p}$ en la estimación de la varianza y teniendo en cuenta que:

$$1-\alpha = 0{,}99 \Rightarrow \frac{\alpha}{2} = 0{,}005 \Rightarrow z_{\alpha/2} = z_{0{,}005} = 2{,}575$$

Se tiene que un intervalo de confianza al 99% para la proporción poblacional p, de personas que miden más de 190 cm, viene dado por:

$$IC_{0{,}99}(p) = \left(0{,}1 - 2{,}575\sqrt{\frac{0{,}1(1-0{,}1)}{600}}; 0{,}1 + 2{,}575\sqrt{\frac{0{,}1(1-0{,}1)}{600}}\right) = (0{,}0685; 0{,}1315)$$

c) Tenemos que:

$$IC_{1-\alpha}(p) = \left(\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}; \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) = (0{,}076; 0{,}124)$$

$$IC_{1-\alpha}(p) = \left(0{,}1 - z_{\frac{\alpha}{2}}\sqrt{\frac{0{,}1(1-0{,}1)}{600}}; 0{,}1 + z_{\frac{\alpha}{2}}\sqrt{\frac{0{,}1(1-0{,}1)}{600}}\right) = (0{,}076; 0{,}124)$$

Despejando $z_{\frac{\alpha}{2}}$ de cualquiera de estas dos ecuaciones:

$$0{,}1 - z_{\frac{\alpha}{2}}\sqrt{\frac{0{,}1(1-0{,}1)}{600}} = 0{,}076 \ ; \ 0{,}1 + z_{\frac{\alpha}{2}}\sqrt{\frac{0{,}1(1-0{,}1)}{600}} = 0{,}124$$

Se obtiene que $z_{\frac{\alpha}{2}} = 1{,}9596 \simeq 1{,}96$, por lo que $P\left(Z < z_{\frac{\alpha}{2}}\right) = 0{,}975 = 1 - \frac{\alpha}{2}$, con lo finalmente $\frac{\alpha}{2} = 0{,}025$.

Como $\alpha = 0{,}05$ entonces $100(1-\alpha) = 100(1-0{,}05) = 100 \cdot 0{,}95 = 95$.

Así pues, el nivel de confianza será del 95 %.

8. Un fabricante de automóviles ha realizado un estudio de mercado en un determinado municipio y ha encontrado que de una muestra aleatoria de 500 turismos, 80 de ellos tienen motor diésel.

Con un nivel de confianza del 94 %, determina el intervalo de confianza para la proporción de turismos que tiene motor diésel en este municipio.

Sea *X*: "número de turismos, de los 500, con motor diesel". La variable *X* tiene distribución binomial $B(n=500,\ p)$, siendo p la probabilidad de que un turismo elegido al azar, en ese municipio, tenga motor diesel.

Un estimador puntual para la proporción p, es la proporción muestral:

$$\hat{p}=\frac{80}{500}=0{,}16$$

La distribución de la proporción muestral $\hat{p}$ se puede aproximar por la de una distribución normal, dado que el tamaño de la muestra es suficientemente grande:

$$\hat{p}\sim N\left(\mu=p;\sigma^2=\frac{p(1-p)}{500}\right)$$

Aproximando p por $\hat{p}=0{,}16$ en la estimación de la varianza, y teniendo en cuenta que:

$$1-\alpha=0{,}94\Rightarrow\frac{\alpha}{2}=0{,}03\Rightarrow z_{\frac{\alpha}{2}}=z_{0{,}03}=1{,}88$$

Se tiene que un intervalo de confianza al 94 % para la proporción poblacional p, de turismos con motor diésel, viene dado por:

$$IC_{0{,}94}(p)=\left(0{,}16-1{,}88\sqrt{\frac{0{,}16(1-0{,}16)}{500}};0{,}16+1{,}88\sqrt{\frac{0{,}16(1-0{,}16)}{500}}\right)=(0{,}1292;0{,}1908)$$

9 y 10. Ejercicios resueltos.

11. En un estudio sobre gorriones se sabe que la distancia que recorren volando en una pasada, en busca de alimento, sigue una distribución normal tanto en los machos como en las hembras. Las desviaciones típicas poblacionales son de 80 y 75 m respectivamente. Con el fin de estimar la diferencia de medias de distancias recorridas, se toma una muestra de 40 machos y 35 hembras y se determinan las medias muestrales que son, respectivamente, 230 y 140 m. Halla un intervalo de confianza al 95 % para la diferencia de medias poblacionales.

Se consideran las variables aleatorias:

X: "distancia, en metros, que recorre el macho de gorrión en una pasada". $X\sim N(\mu_X,\sigma_X=80)$.

Y: "distancia, en metros, que recorre la hembra de gorrión en una pasada". $Y\sim N(\mu_Y,\sigma_Y=75)$.

Se toman muestras aleatorias de $n=40$ machos y de $m=35$ hembras. Las distribuciones de las medias muestrales son, respectivamente, las siguientes distribuciones normales:

$$\overline{X}\sim N\left(\mu_{\overline{X}},\sigma_{\overline{X}}=\frac{80}{\sqrt{40}}\right)\ ;\ \overline{Y}\sim N\left(\mu_{\overline{Y}},\sigma_{\overline{Y}}=\frac{75}{\sqrt{35}}\right)$$

De modo que, la distribución de la diferencia entre la media muestral de los machos y la media muestral de las hembras, es una distribución normal, con la media y la varianza que se indican a continuación:

$$\overline{X}-\overline{Y}\sim N\left(\mu_{\overline{X}-\overline{Y}},\sigma^2_{\overline{X}-\overline{Y}}\right)=N\left(\mu_X-\mu_Y,\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}\right)=N\left(\mu_X-\mu_Y,\frac{80^2}{40}+\frac{75^2}{35}\right)=N(\mu_X-\mu_Y;320{,}714)$$

Las muestras seleccionadas proporcionan los estimadores de las medias poblacionales: $\overline{x}=230$ m ; $\overline{y}=140$ m

Por tanto, teniendo en cuenta que $1-\alpha=0{,}95\Rightarrow\frac{\alpha}{2}=0{,}025\Rightarrow z_{\frac{\alpha}{2}}=z_{0{,}025}=1{,}96$, un intervalo de confianza al 95 %, para la diferencia entre la media poblacional de los machos y la de las hembras, es:

$$IC_{0{,}95}(\mu_X-\mu_Y)=\left(230-140-1{,}96\sqrt{320{,}714}\ ;\ 230-140+1{,}96\sqrt{320{,}714}\right)=(54{,}9;125{,}1)$$

12. El contenido medio en azúcar de una muestra aleatoria de 50 envases de una bebida refrescante de la marca *A* es 248 g, con una cuasidesviación típica de 7,2 g, en tanto que el contenido medio en una muestra de 40 envases del mismo tamaño de la marca *B* es de 320 g, con una cuasidesviación típica de 4,26 g. Obtén un intervalo de confianza al 97 % para la diferencia de contenido medio en azúcar de las marcas *A* y *B*. A este nivel de confianza, ¿puede decirse que los envases de la marca B tienen mayor contenido en azúcar que la marca *A*?

Se consideran las variables aleatorias:

X: "contenido, en gramos, de azúcar en la bebida marca *A*". $X \sim N(\mu_X, \sigma_X)$.

Y: "contenido, en gramos, de azúcar en la bebida marca *B*". $Y \sim N(\mu_Y, \sigma_Y)$.

Se toman muestras aleatorias de $n = 50$ envases de la marca *A* y de $m = 40$ envases de la marca *B*. Como los tamaños muestrales son suficientemente grandes, las varianzas poblacionales pueden estimarse por las cuasivarianzas muestrales.

Además, la distribución de la diferencia de la media muestral del contenido en azúcar de la marca *A* menos la media muestral del contenido en azúcar de la marca *B* puede aproximarse por una distribución normal con la media y varianza que se indican:

$$\overline{X} - \overline{Y} \sim N\left(\mu_{\overline{X}-\overline{Y}}, \sigma^2_{\overline{X}-\overline{Y}}\right) = N\left(\mu_X - \mu_Y, \frac{7{,}2^2}{50} + \frac{4{,}26^2}{40}\right)$$

Las muestras seleccionadas proporcionan los estimadores de las medias poblacionales: $\overline{x} = 248g$; $\overline{y} = 320m$

Por tanto, teniendo en cuenta que $1 - \alpha = 0{,}97 \Rightarrow \frac{\alpha}{2} = 0{,}015 \Rightarrow z_{\frac{\alpha}{2}} = z_{0{,}015} = 2{,}17$, un intervalo de confianza al 97 %, para la diferencia entre la media poblacional de la marca *A* y la de la marca *B* es:

$$IC_{0{,}97}(\mu_X - \mu_Y) = \left(248 - 320 - 2{,}17\sqrt{\frac{7{,}2^2}{50} + \frac{4{,}26^2}{40}} ; 248 - 320 + 2{,}17\sqrt{\frac{7{,}2^2}{50} + \frac{4{,}26^2}{40}}\right) = (-74{,}65 ; -69{,}35)$$

Dado que el intervalo de confianza para la diferencia entre la media de *A* y la media de *B*, tiene los extremos negativos, y, por tanto, no contiene al cero, puede decirse que el contenido medio de azúcar de los envases de la marca B es significativamente más alto que el de la marca A.

13 a 15. Ejercicios resueltos.

16. Para estimar la proporción de balances contables incorrectos de un banco, se seleccionan aleatoriamente 200 balances, y se encuentra que 19 de ellos son incorrectos.

a) Obtén un intervalo de confianza, al 95 %, para la proporción de balances incorrectos.

b) Calcula el tamaño muestral para que el error al estimar la proporción no sea superior a 0,02 al 99 % de confianza.

La variable aleatoria *X*: "número de balances incorrectos del banco, de los 200 elegidos," tiene una distribución binomial $B(n=200, p)$, con la probabilidad *p*, de que un balance elegido al azar sea incorrecto, desconocida.

Dado que el tamaño de la muestra, $n = 200$, es suficientemente grande, la distribución de la proporción muestral de balances incorrectos se puede aproximar por una distribución normal:

$$\hat{p} \sim N\left(\mu = p, \sigma^2 = \frac{p(1-p)}{200}\right)$$

a) Aproximando p por $\hat{p} = \frac{19}{200} = 0{,}095$, en la estimación de la varianza, y teniendo en cuenta que $1-\alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{\frac{\alpha}{2}} = z_{0{,}025} = 1{,}96$, se tiene que un intervalo de confianza al 95 %, para la proporción poblacional p, viene dado por:

$$IC_{0{,}95}(p) = \left(0{,}095 - 1{,}96\sqrt{\frac{0{,}095(1-0{,}095)}{200}}; 0{,}095 + 1{,}96\sqrt{\frac{0{,}095(1-0{,}095)}{200}}\right) = (0{,}0544; 0{,}1356)$$

b) El error cometido al estimar la proporción *p* mediante un intervalo de confianza es:

$$E = z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Como $1-\alpha = 0{,}99 \Rightarrow \frac{\alpha}{2} = 0{,}005 \Rightarrow z_{0{,}005} = 2{,}575$. Además $\hat{p} = 0{,}095$ y el error debe ser menor o igual que 0,02, por lo que se tiene que:

$$z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le 0{,}02 \Rightarrow n \ge 2{,}575^2\,\frac{0{,}095(1-0{,}095)}{0{,}02^2} \Rightarrow n \ge 1425{,}17$$

Por tanto, el tamaño de la muestra debería ser, al menos, de 1426 balances, para asegurar las condiciones exigidas.

17. En 169 poblaciones distintas en el territorio nacional, se ha encuestado a agentes inmobiliarios sobre el precio de la vivienda, resultando que el precio medio por metro cuadrado es de 1764 €, con una desviación típica de 258 €.

a) Estima el precio medio poblacional con un 97 % de confianza.

b) ¿De qué tamaño tendría que ser la muestra para hacer dicha estimación con un error menor de 30 €, con una confianza del 97 %?

La distribución de la variable *X*: "precio, en euros, del metro cuadrado de vivienda" tiene media μ y desviación típica σ, ambas desconocidas.

Para estimar ambos parámetros se selecciona una muestra de 169 viviendas. Al ser la muestra suficientemente grande, la distribución de la media muestral se puede aproximar por la de una distribución normal, con media μ y varianza $\frac{\sigma^2}{n}$, en la que σ se puede aproximar por la desviación típica de la muestra. Es decir:

$$\overline{X} \sim N\left(\mu, \sigma_{\overline{X}}^2 = \frac{258^2}{n}\right)$$

a) Como $1-\alpha = 0{,}97 \Rightarrow \frac{\alpha}{2} = 0{,}015 \Rightarrow z_{0{,}015} = 2{,}17$, un intervalo de confianza al 97% para la media poblacional μ es, por tanto:

$$IC_{0{,}97}(\mu) = \left(1764 - 2{,}17\frac{258}{\sqrt{169}}, 1764 + 2{,}17\frac{258}{\sqrt{169}}\right) = (1720{,}93\ ;\ 1807{,}07)$$

b) Si $z_{0{,}015} = 2{,}17$ y la desviación típica σ se puede estimar por 258 euros, para que el error *E* cometido en la estimación de la media poblacional sea inferior a 30 euros, se debe cumplir que:

$$\text{Error} < 30 \Rightarrow E = z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} < 30 \Rightarrow 2{,}17\frac{258}{\sqrt{n}} < 30 \Rightarrow n > \left(\frac{2{,}17 \cdot 258}{30}\right)^2 = 348{,}27$$

Por tanto, la muestra debería tener un tamaño de, al menos, 349 viviendas para que se satisfagan las condiciones propuestas.

18. La dirección de un centro educativo, está interesada en conocer la opinión de los estudiantes de todos los cursos a partir de primero de ESO, sobre determinados servicios que presta el centro.

Diseña una ficha técnica que sirva para la realización tal encuesta.

La ficha técnica debe contar, al menos, con los siguientes apartados:

1. Objetivos de la investigación: conocer la opinión de los estudiantes (a partir de 1.º de ESO) sobre algunos servicios que presta el centro.
2. Descripción del universo o población objeto del estudio: todos los estudiantes matriculados en el centro, de 1.º ESO en adelante.
3. Método de selección de la muestra: estratificada, por ejemplo, por curso (1.º ESO, 2.º ESO, etc.) y dentro de cada estrato una muestra aleatoria proporcional al tamaño del estrato.

   Tamaño muestral: en total se encuestará a 300 alumnos.
4. El nivel de confianza será del 95 % y se admite un error de $\pm 5{,}7$ % bajo supuesto de máxima indeterminación.
5. Descripción de cómo se ha llevado a cabo el trabajo de campo: mediante la realización de un cuestionario anónimo que los estudiantes llevan a su casa y devuelven en sobre cerrado sin identificar.
6. Período de recogida de la información: del 5 al 10 de junio del presente curso.
7. Organismos, empresas o personas responsables del informe: departamento de Orientación.

19. La Consejería de Turismo de una Comunidad Autónoma está interesada en conocer la opinión de los turistas mayores de 18 años que visitan una ciudad de esta Comunidad.

Diseña una ficha técnica que recoja las características necesarias que debe tener ese estudio estadístico.

La ficha técnica debe contener, al menos, los siguientes apartados:

1. Objetivos de la investigación: conocer la opinión de los turistas que visitan la ciudad acerca de los servicios y la información que presta la Oficina de Turismo.
2. Descripción del universo o población objeto del estudio: todos los visitantes de la ciudad en un año determinado, mayores de 18 años que buscan información en las oficinas de turismo.
3. Método de selección de la muestra: muestra aleatoria por cuotas de edad y sexo y estratificada mensualmente en función del número de visitantes esperados.

   Tamaño muestral: muestra de tamaño 1500 personas.
4. El nivel de confianza será del 90 % y el error con el que se presentan los resultados de $\pm 2,12\ \%$ bajo supuesto de máxima indeterminación.
5. Descripción de cómo se ha llevado a cabo el trabajo de campo: mediante encuestas que se entregan a los visitantes elegidos y que estos depositan, de forma anónima, en buzones habilitados al efecto.
6. Período de recogida de la información: todo el año en curso.
7. Organismos, empresas o personas responsables del informe: empresa contratada por la Consejería de turismo mediante concurso público.

20 a 27. Ejercicios resueltos.

EJERCICIOS

Intervalo de confianza para la media poblacional

28. Se ha obtenido una muestra de diez valores de una variable con distribución normal de media μ desconocida y varianza $\sigma^2 = 3$. Los valores observados son:

2,1; 2,5; 1,6; 2,4; 2,8; 2,0; 1,9; 1,2; 2,9; 3,2.

Construye un intervalo de confianza para la media con nivel de confianza del:

a) 90 % **b)** 95 % **c)** 99 %

La variable aleatoria *X* tiene una distribución $N\left(\mu, \sigma = \sqrt{3}\right)$.

La muestra aleatoria tiene una media muestral de:

$$\bar{x} = \frac{2{,}1 + 2{,}5 + 1{,}6 + 2{,}4 + 2{,}8 + 2{,}0 + 1{,}9 + 1{,}2 + 2{,}9 + 3{,}2}{10} = 2{,}26$$

a) De $1 - \alpha = 0{,}90$, resulta que $\frac{\alpha}{2} = 0{,}05$; y de las tablas de la N(0,1), se obtiene que:

$$P\left(Z < z_{0,05}\right) = 1 - 0{,}05 = 0{,}95 \Rightarrow z_{0,05} = 1{,}645$$

Por lo que el intervalo de confianza al 90% para la media μ es:

$$IC_{0,90}(\mu) = \left(2{,}26 - 1{,}645\frac{\sqrt{3}}{\sqrt{10}};\ 2{,}26 + 1{,}645\frac{\sqrt{3}}{\sqrt{10}}\right) = (1{,}36\ ;\ 3{,}16)$$

b) De $1 - \alpha = 0{,}95$, resulta que $\frac{\alpha}{2} = 0{,}025$; y de las tablas de la **N**(0,1), se obtiene que:

$$P\left(Z < z_{0,025}\right) = 1 - 0{,}025 = 0{,}975 \Rightarrow z_{0,025} = 1{,}96$$

Por lo que el intervalo de confianza al 95% para la media μ es:

$$IC_{0,90}(\mu) = \left(2{,}26 - 1{,}96\frac{\sqrt{3}}{\sqrt{10}};\ 2{,}26 + 1{,}96\frac{\sqrt{3}}{\sqrt{10}}\right) = (1{,}19;\ 3{,}33)$$

c) De $1 - \alpha = 0{,}99$, resulta que $\frac{\alpha}{2} = 0{,}005$; y de las tablas de la N(0,1), se obtiene que:

$$P\left(Z < z_{0,005}\right) = 1 - 0{,}005 = 0{,}995 \Rightarrow z_{0,005} = 2{,}575$$

Por lo que el intervalo de confianza al 99% para la media μ es:

$$IC_{0,90}(\mu) = \left(2{,}26 - 2{,}575\ \frac{\sqrt{3}}{\sqrt{10}};\ 2{,}26 + 2{,}575\ \frac{\sqrt{3}}{\sqrt{10}}\right) = (0{,}85;\ 3{,}67)$$

29. La cantidad de refresco que se sirve en cada vaso a la entrada de unos cines está normalmente distribuida con una desviación típica de 15 mL. Se han medido las cantidades en los vasos de los 25 asistentes de una determinada sesión que compraron un refresco y se ha obtenido un promedio de 200,8 mL. Fijado un nivel de confianza del 90 %, calcula el intervalo de confianza para la media de la cantidad de refresco que se sirve en cada vaso. Detalla los pasos realizados para obtener los resultados.

Se considera la variable aleatoria X: "cantidad de refresco que se sirve en cada vaso, en mL".

La distribución de la variable X es normal, de media desconocida: $X \sim N(\mu, \sigma = 15)$

Una muestra aleatoria de tamaño $n = 25$, proporciona una media muestral $\overline{x} = 200{,}8$ mL. Si el nivel de confianza es del 90 % entonces:

$$1 - \alpha = 0{,}9 \Rightarrow \frac{\alpha}{2} = 0{,}05 \Rightarrow z_{0{,}05} = 1{,}645$$

De manera que el intervalo de confianza correspondiente para el contenido medio de refresco, μ, que se sirve en cada vaso es:

$$IC_{0{,}9}(\mu) = \left(200{,}8 - 1{,}645\,\frac{15}{\sqrt{25}}\ ;\ 200{,}8 + 1{,}645\,\frac{15}{\sqrt{25}}\right) = (195{,}865\ ;\ 205{,}735)$$

30. La edad a la que obtienen el permiso de conducir los habitantes de una determinada población es una variable aleatoria que se puede aproximar por una distribución normal de media desconocida y desviación típica de 4 años. Para estimar la media se elige aleatoriamente una muestra de 100 habitantes de dicha población.

a) ¿Cuál es la varianza de la distribución muestral?

b) Si la media muestral es 24 años, halla un intervalo de confianza al 90 % para la media poblacional.

La variable aleatoria X: "edad, en años, a la que obtienen el permiso de conducir" tiene una distribución normal:

$$X \sim N(\mu, \sigma = 4)$$

a) La media muestral tiene, también, una distribución normal $\overline{X} \sim N\left(\mu_{\overline{X}}, \sigma^2_{\overline{X}}\right)$ cuya varianza vale:

$$\sigma^2_{\overline{X}} = \frac{\sigma^2}{n} = \frac{16}{100} = 0{,}16$$

b) Como el nivel de confianza es del 90 %: $1 - \alpha = 0{,}9 \Rightarrow \frac{\alpha}{2} = 0{,}05 \Rightarrow z_{0{,}05} = 1{,}645$

Por lo que, un intervalo de confianza al 90 %, para la media de edad poblacional a la que se obtiene el permiso de conducir, es:

$$IC_{0{,}9}(\mu) = \left(24 - 1{,}645\sqrt{0{,}16};\ 24 + 1{,}645\sqrt{0{,}16}\right) = (23{,}34;\ 24{,}66)$$

31. Se desea estimar la presión arterial sistólica de las personas mayores de edad tras la realización de un determinado ejercicio físico. La presión arterial sistólica media de una muestra aleatoria de 50 personas fue de 190 mmHg, con una desviación típica estimada de 26 mmHg. Calcula un intervalo de confianza al 90 % para la presión sistólica media.

Sea X: "presión arterial sistólica, en mmHg, para personas mayores de edad, tras la realización de un ejercicio físico". Sean μ la media de X y σ su desviación típica.

Para estimar ambos parámetros, se elige una muestra aleatoria de $n = 50$ personas que proporciona una media y una desviación típica muestrales de: $\overline{x} = 190$ mmHg ; $\sigma = 26$ mmHg

Como el tamaño de la muestra es suficientemente grande, la distribución de media muestral puede aproximarse por una distribución normal, y como el nivel de confianza es del 90 % tenemos que:

$$1 - \alpha = 0{,}9 \Rightarrow \frac{\alpha}{2} = 0{,}05 \Rightarrow z_{0{,}05} = 1{,}645$$

Un intervalo de confianza al 90 %, para la presión sistólica media poblacional μ, es:

$$IC_{0{,}9}(\mu) = \left(190 - 1{,}645\frac{26}{\sqrt{50}}\ ;\ 190 + 1{,}645\frac{26}{\sqrt{50}}\right) = (183{,}95\ ;\ 196{,}05)$$

32. El consumo bimestral de energía eléctrica de una población de 100 personas se distribuye normalmente con una media de 59 kWh y una desviación típica de 6 kWh. Calcula el intervalo de confianza para la media con un nivel de confianza del 97 %. Detalla los pasos realizados para obtener los resultados.

La variable aleatoria *X*: "consumo bimestral de energía eléctrica, en kWh". La distribución de la variable *X* es normal, de media y varianza desconocidas.

Una muestra aleatoria de $n = 100$ personas, proporciona una media muestral $\overline{x} = 59$ kWh y una desviación típica de $\sigma = 6$ kWh. A un nivel de confianza del 97 % se tiene que:

$$1 - \alpha = 0{,}97 \Rightarrow \frac{\alpha}{2} = 0{,}015 \Rightarrow z_{0{,}015} = 2{,}17$$

De manera que el intervalo de confianza al 97 %, para la media μ, del consumo de energía eléctrica, es:

$$IC_{0{,}97}(\mu) = \left(59 - 2{,}17\frac{6}{\sqrt{100}};\ 59 + 2{,}27\frac{6}{\sqrt{100}}\right) = (57{,}698;\ 60{,}302)$$

33. Para una muestra de 49 pisos de dos habitaciones de una gran ciudad, el alquiler medio resultó igual a 425 €. Tomando una desviación típica igual a 50 €, construir un intervalo de confianza, del 97 %, para la media del alquiler de los pisos de dos habitaciones de esa gran ciudad.

La variable aleatoria *X*: "precio del alquiler, en euros, de pisos de 2 habitaciones".

Se desconoce la media μ de la distribución de *X*, pero se sabe que su desviación típica es $\sigma = 50$ euros.

Una muestra aleatoria de $n = 49$ pisos proporciona una media muestral $\overline{x} = 425$ €. Como la muestra es suficientemente grande, la distribución de la media muestral puede aproximarse por una distribución normal. Es decir:

$$\overline{X} \sim N\left(\mu_{\overline{X}} = \mu, \sigma^2_{\overline{X}} = \frac{50^2}{49}\right)$$

Teniendo en cuenta que a un nivel de confianza del 97% resulta que:

$$1 - \alpha = 0{,}97 \Rightarrow \frac{\alpha}{2} = 0{,}015 \Rightarrow z_{0{,}015} = 2{,}17$$

Un intervalo de confianza al 97%, para la media del alquiler de pisos de dos habitaciones, es:

$$IC_{0{,}97}(\mu) = \left(425 - 2{,}17\frac{50}{7};\ 425 - 2{,}17\frac{50}{7}\right) = (409{,}5;\ 440{,}5)$$

Intervalo de confianza para la proporción

34. En una muestra de 450 jóvenes, 110 dicen que sus lecturas favoritas son cómics. Para un nivel del 90 %, obtén un intervalo de confianza para la proporción de jóvenes que tienen los comics como sus lecturas favoritas.

La variable *X*: "la lectura favorita del joven es el cómic" es una variable de Bernoulli de parámetro p, probabilidad de que la lectura favorita de un joven elegido al azar sea el cómic.

Si se selecciona una muestra de tamaño $n = 450$, la distribución de la proporción muestral $\hat{p}$ se puede aproximar por la de una normal:

$$\hat{p} \sim N\left(\mu = p, \sigma^2 = \frac{p(1-p)}{450}\right)$$

En este caso, una estimación puntual para p es $\hat{p} = \frac{110}{450} = 0{,}2444$.

Sustituyendo *p* por $\hat{p}$ para estimar la varianza, y teniendo en cuenta que:

$$1 - \alpha = 0{,}90 \Rightarrow \frac{\alpha}{2} = 0{,}05 \Rightarrow z_{0{,}05} = 1{,}645$$

Un intervalo de confianza al 90 % para la proporción p es:

$$IC_{0{,}90}(p) = \left(0{,}2444 - 1{,}645\sqrt{\frac{0{,}2444(1-0{,}2444)}{450}};\ 0{,}2444 + 1{,}645\sqrt{\frac{0{,}2444(1-0{,}2444)}{450}}\right) = (0{,}2111;\ 0{,}2777)$$

35. Para estimar la proporción de personas con sobrepeso en una población se ha tomado una muestra aleatoria simple de tamaño 100 personas, de las cuales 21 tienen sobrepeso. Calcula el intervalo de confianza al 96 % para la proporción de personas con sobrepeso en la población.

Si de la población se elige una persona al azar, la variable *X*: "la persona tiene sobrepeso" es una variable de Bernoulli de parámetro p, probabilidad de que una persona elegida al azar tenga sobrepeso.

Si se selecciona una muestra de tamaño $n = 100$, la distribución de la proporción muestral $\hat{p}$ se puede aproximar por la de una normal:

$$\hat{p} \sim N\left(\mu = p, \sigma^2 = \frac{p(1-p)}{100}\right)$$

En este caso, una estimación puntual para p es $\hat{p} = \frac{21}{100} = 0{,}21$.

Sustituyendo p por $\hat{p}$ para estimar la varianza, y teniendo en cuenta que:

$$1-\alpha = 0{,}96 \Rightarrow \frac{\alpha}{2} = 0{,}02 \Rightarrow z_{0{,}02} = 2{,}055$$

Un intervalo de confianza al 96 % para la proporción p es:

$$IC_{0{,}96}(p) = \left(0{,}21 - 2{,}055\sqrt{\frac{0{,}21(1-0{,}21)}{100}}; 0{,}21 + 2{,}055\sqrt{\frac{0{,}21(1-0{,}21)}{100}}\right) = (0{,}1263;\ 0{,}2937)$$

36. Una compañía aérea tiene contratada una empresa para la recuperación de los equipajes perdidos de sus pasajeros. Para comprobar la eficiencia de la empresa, la compañía desea saber la proporción de equipajes recuperados. Para ello realiza una encuesta a 122 pasajeros que perdieron el equipaje. De ellos, 103 lo recuperaron.

a) ¿Cuál es la estimación de la proporción de equipajes recuperados?

b) Obtén el intervalo de confianza al 99 % para la proporción de equipajes recuperados.

Si de la población (pasajeros que perdieron el equipaje) se elige una persona al azar, la variable *X*: "la persona recuperó el equipaje" es una variable de Bernoulli de parámetro *p*, que es la probabilidad de que una persona de esa población elegida al azar haya recuperado el equipaje.

a) En este caso, si de la muestra de tamaño $n = 122$ pasajeros que perdieron el equipaje, 103 lo recuperaron, una estimación puntual para *p* es:

$$\hat{p} = \frac{103}{122} = 0{,}8443$$

b) Si se selecciona una muestra de tamaño $n = 122$, la distribución de la proporción muestral $\hat{p}$ se puede aproximar por la de una distribución normal:

$$\hat{p} \sim N\left(\mu = p, \sigma^2 = \frac{p(1-p)}{122}\right)$$

Sustituyendo *p* por $\hat{p}$ para estimar la varianza, y teniendo en cuenta que:

$$1-\alpha = 0{,}99 \Rightarrow \frac{\alpha}{2} = 0{,}005 \Rightarrow z_{0{,}005} = 2{,}575$$

Un intervalo de confianza al 99 % para la proporción *p* es:

$$IC_{0{,}99}(p) = \left(0{,}8443 - 2{,}575\sqrt{\frac{0{,}8443(1-0{,}8443)}{122}}; 0{,}8443 + 2{,}575\sqrt{\frac{0{,}8443(1-0{,}8443)}{122}}\right) = (0{,}7598;\ 0{,}9288)$$

37. Una encuesta pregunta a los individuos de una población por la modificación de la ordenanza municipal. Se entrevistó a 2 500 personas de las que 955 se decantaron por mantener la ordenanza municipal como está.

a) ¿Cuál es la estimación puntual del porcentaje de individuos que se inclinan por seguir manteniendo la ordenanza actual del municipio?

b) Obtén un intervalo de confianza al 95 % para el porcentaje de población que prefiere mantener como está la ordenanza municipal.

Si de la población se elige una persona al azar, la variable *X*: "la persona se decanta por mantener la ordenanza municipal como está" es una variable de Bernoulli de parámetro *p*, probabilidad de que una persona de esa población elegida al azar opine que la ordenanza municipal siga como está.

a) En este caso, si de la muestra de $n = 2500$ personas, 955 opinan que la ordenanza municipal siga como está, la estimación de *p* es:

$$\hat{p} = \frac{955}{2500} = 0,382$$

b) Si se selecciona una muestra de tamaño $n = 2\,500$, la distribución de la proporción muestral $\hat{p}$ se puede aproximar por la de una distribución normal:

$$\hat{p} \sim N\left(\mu = p, \sigma^2 = \frac{p(1-p)}{2500}\right)$$

Sustituyendo *p* por $\hat{p}$ para estimar la varianza, y teniendo en cuenta que:

$$1-\alpha = 0,95 \Rightarrow \frac{\alpha}{2} = 0,025 \Rightarrow z_{0,025} = 1,96$$

Un intervalo de confianza al 95 % para la proporción p es:

$$IC_{0,95}(p) = \left(0,382 - 1,96\sqrt{\frac{0,382(1-0,382)}{2500}};\ 0,382 + 1,96\sqrt{\frac{0,382(1-0,382)}{2500}}\right) = (0,363;\ 0,401)$$

Intervalo de confianza para la diferencia de medias

38. *Se supone que el gasto por persona en las rebajas de enero sigue una distribución normal. En la ciudad *A*, es $N(\mu_A, \sigma_A = 40)$ y en la ciudad *B* $N(\mu_B, \sigma_B = 50)$. Para estimar el gasto medio por persona, en la ciudad *A*, se ha seleccionado una muestra aleatoria de 100 personas y se ha obtenido un gasto medio de 225 €. Y para estimar el gasto medio por persona en *B*, se eligió una muestra de 150 personas cuya gasto medio ascendió a 350 €.

a) Construye un intervalo de confianza al 95% para la diferencia del gasto medio por persona en ambas ciudades *A* y *B*, en las rebajas de enero.

b) ¿Se puede afirmar, con un nivel de significación del 5 %, que el gasto medio por persona en *B* es mayor que en *A*?

a) Las distribuciones de las medias muestrales en las ciudades A y B son, respectivamente:

$$\overline{X}_A \sim N\left(\mu_A, \sigma_A = \frac{40}{\sqrt{100}}\right) \quad ; \quad \overline{X}_B \sim N\left(\mu_B, \sigma_B = \frac{50}{\sqrt{150}}\right)$$

Y la distribución de la diferencia de las medias muestrales es también una distribución normal:

$$\overline{X}_A - \overline{X}_B \sim N\left(\mu_{\overline{X}_A - \overline{X}_B} = \mu_A - \mu_B \,;\, \sigma^2_{\overline{X}_A - \overline{X}_B} = \frac{40^2}{100} + \frac{50^2}{150} = 32{,}67\right)$$

Dado que $1 - \alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$, un intervalo de confianza al 95 % para la diferencia del gasto medio en *A* menos el gasto medio en *B* es:

$$IC_{0{,}95}(\mu_A - \mu_B) = \left(225 - 350 - 1{,}96\sqrt{32{,}67};\ 225 - 350 + 1{,}96\sqrt{32{,}67}\right) = (-136{,}20\,;\,-113{,}80)$$

b) Puesto que el intervalo de confianza calculado en el apartado anterior no contiene al cero, ya que ambos extremos son negativos, puede decirse que el gasto medio en *B* es significativamente mayor que el gasto medio en *A* al 95 % de confianza (5 % de significación).

39. Las puntuaciones en las pruebas de acceso en Matemáticas Aplicadas a las Ciencias Sociales siguen una ley normal de media desconocida. En la universidad *A* la desviación típica es 1,8 puntos, mientras que en la universidad *B* la desviación típica es de 2,2 puntos. En una muestra de 36 alumnos de la universidad *A* se ha obtenido una puntuación media de 5,5 puntos y en una muestra de 48 de la universidad *B*, la puntuación media ha sido de 7,2 puntos.

a) Calcula un intervalo de confianza del 90 % para la diferencia de las puntuaciones medias en Matemáticas Aplicadas a las Ciencias Sociales, de las universidades *A* y *B*.

b) ¿Se puede afirmar que las puntuaciones medias son mejores en la universidad *B* que en la *A*, al 90 % de confianza?

a) Las distribuciones de las medias muestrales en las universidades *A* y *B* son, respectivamente:

$$\overline{X}_A \sim N\left(\mu_{\overline{X}_A} = \mu_A, \sigma^2_{\overline{X}_A} = \frac{1{,}8}{\sqrt{36}}\right) \quad ; \quad \overline{X}_B \sim N\left(\mu_{\overline{X}} = \mu_B, \sigma^2_{\overline{X}_B} = \frac{2{,}2}{\sqrt{48}}\right)$$

La distribución de la diferencia de las medias muestrales es también una distribución normal:

$$\overline{X}_A - \overline{X}_B \sim N\left(\mu_{\overline{X}_A - \overline{X}_B} = \mu_A - \mu_B \,;\, \sigma^2_{\overline{X}_A - \overline{X}_B} = \frac{1{,}8^2}{36} + \frac{2{,}2^2}{48} = 0{,}1908\right)$$

Como $1 - \alpha = 0{,}90 \Rightarrow \frac{\alpha}{2} = 0{,}05 \Rightarrow z_{0{,}05} = 1{,}645$, un intervalo de confianza al 90 %, para la diferencia de la puntuación media en MACS, en la universidad *A*, menos la puntuación media en la universidad *B*, es:

$$IC_{0{,}95}(\mu_A - \mu_B) = \left(5{,}5 - 7{,}2 - 1{,}645\sqrt{0{,}1908}\,;\ 5{,}5 - 7{,}2 + 1{,}645\sqrt{0{,}1908}\right) = (-2{,}42;-0{,}98)$$

b) Dado que el intervalo de confianza calculado en el apartado anterior no contiene al cero, ya que ambos extremos son negativos, se puede afirmar que la puntuación media en la universidad *B* es significativamente mayor que en la universidad *A*, al 90 % de confianza (10 % de significación).

Error de la estimación y tamaño muestral

40. El tiempo de espera para ser atendido en la caja de un establecimiento sigue una distribución normal de desviación típica 5 minutos.

a) Calcula el tamaño mínimo de la muestra para estimar, con un nivel de confianza del 95 %, el tiempo medio de espera con un error que no sea superior a medio minuto.

b) ¿Cuál es dicho tamaño mínimo para un nivel de confianza del 99 %?

La variable aleatoria *X*: "tiempo de espera, en minutos, en la caja del establecimiento" tiene distribución normal, con la media μ desconocida y desviación típica 5 minutos, es decir: $X \sim N(\mu, \sigma = 5)$.

a) Si el nivel confianza requerido es del 95 %, se tiene que: $1-\alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$.

Y como el error *E* para estimar el tiempo medio de espera debe ser menor que 0,5 min; entonces:

$$E = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < 0{,}5 \Rightarrow 1{,}96 \frac{5}{\sqrt{n}} < 0{,}5 \Rightarrow n > \left(\frac{1{,}96 \cdot 5}{0{,}5}\right)^2 = 384{,}16$$

Es decir, con las condiciones requeridas, se requiere un tamaño muestral de, al menos, 385 clientes para estimar el tiempo medio de espera en la caja.

b) Si el nivel de confianza se eleva al 99 %, entonces: $1-\alpha = 0{,}99 \Rightarrow \frac{\alpha}{2} = 0{,}005 \Rightarrow z_{0{,}005} = 2{,}575$.

Y como el error *E* para estimar el tiempo medio de espera debe ser menor que 0,5 min; entonces:

$$E = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < 0{,}5 \Rightarrow 2{,}575 \frac{5}{\sqrt{n}} < 0{,}5 \Rightarrow n > \left(\frac{2{,}575 \cdot 5}{0{,}5}\right)^2 = 663{,}1$$

Es decir, con las condiciones requeridas, se requiere un tamaño muestral de, al menos, 664 clientes para estimar el tiempo medio de espera en la caja.

41. Se supone que el número de horas semanales dedicadas al estudio por los estudiantes de una universidad sigue una distribución normal de media desconocida y desviación típica 6 horas. Para estimar la media de horas semanales de estudio se quiere utilizar una muestra de tamaño *n*.

Calcula el valor mínimo de *n* para que, con un nivel de confianza del 99 %, el error en la estimación sea menor de 1 hora.

La variable *X*: "tiempo semanal, en horas, dedicado al estudio" tiene distribución normal $N(\mu, \sigma = 6)$; con μ desconocida.

Se selecciona una muestra de tamaño *n*. Si el nivel de confianza es del 99 %, entonces:

$$1-\alpha = 0{,}99 \Rightarrow \frac{\alpha}{2} = 0{,}005 \Rightarrow z_{0{,}005} = 2{,}575$$

De manera que, para que el error de la estimación *E* sea menor que 1 hora, se debe cumplir que:

$$E = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < 1 \Rightarrow 2{,}575 \frac{6}{\sqrt{n}} < 1 \Rightarrow n > (2{,}575 \cdot 6)^2 = 238{,}7$$

Es decir, debe tomarse una muestra de al menos 239 estudiantes para estimar la media μ con la precisión deseada.

42. Se supone que el precio de un kilo de patatas en una cierta región se puede aproximar por una variable aleatoria con distribución normal de desviación típica igual a 10 céntimos de euro. Una muestra aleatoria simple de tamaño 256 proporciona un precio medio del kilo de patatas igual a 19 céntimos de euro.

a) Determina un intervalo de confianza del 95 % para el precio medio de un kilo de patatas en la región.

b) Se desea aumentar el nivel de confianza al 99 % sin aumentar el error de la estimación ¿Cuál debe ser el tamaño muestral mínimo que ha de observarse?

La variable aleatoria X: "precio, en céntimos de euro, del kilo de patatas", tiene distribución normal, con la media μ desconocida y desviación típica 10 céntimos de euro. Es decir:

$$X \sim N(\mu, \sigma = 10)$$

a) Una muestra aleatoria de $n = 256$ kilos, proporciona una media muestral de $\overline{x} = 19$ céntimos/kilo.
Si el nivel de confianza exigido es del 95%, entonces:

$$1 - \alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$$

De manera que un intervalo de confianza al 95 %, para el precio medio de un kilo de patatas, es:

$$IC_{0{,}95}(\mu) = \left(19 - 1{,}96\,\frac{10}{\sqrt{256}}\,;\,19 + 1{,}96\,\frac{10}{\sqrt{256}}\right) = (17{,}775;\,20{,}225)\text{ céntimos}$$

b) Si se aumenta el nivel confianza al 99 %, se tiene que: $1 - \alpha = 0{,}99 \Rightarrow \frac{\alpha}{2} = 0{,}005 \Rightarrow z_{0{,}005} = 2{,}575$; y como el error de estimación del apartado anterior es $E = z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} = 1{,}96\,\frac{10}{\sqrt{256}} = 1{,}225$; Si se desea un nivel de confianza al 99 % con el error anterior E, el tamaño muestral mínimo que debe observarse es:

$$E = 2{,}575\,\frac{10}{\sqrt{n}} \le 1{,}225 \Rightarrow n \ge \left(\frac{2{,}575 \cdot 10}{1{,}225}\right)^2 = 441{,}86$$

Es decir, con la nueva condición del 99 % de confianza, se requiere una muestra aleatoria de al menos 442 kilos de patatas para estimar el precio medio del kilo de patatas en la región, con las condiciones propuestas.

Síntesis

43. En un determinado municipio, los ingresos mensuales de sus habitantes siguen una distribución normal de media μ y desviación típica 200 €. Se seleccionó al azar una muestra de 100 personas cuya media de ingresos mensuales resultó de 1060 €.

a) Para un nivel de confianza del 95 %, calcula un intervalo de confianza para el ingreso medio mensual en ese municipio.

b) Si se toma un nivel de significación de 0,01, calcula el tamaño muestral mínimo necesario para estimar el ingreso medio mensual con un error menor de 30 €.

c) ¿Cuál tendría que ser el tamaño muestral para que, manteniendo el nivel de confianza del apartado anterior, el error se redujese a la mitad?

Sea la variable aleatoria *X*: "ingresos mensuales, en euros, de los habitantes del municipio". La distribución de *X* es normal, con la media μ desconocida:

$$X \sim N(\mu, \sigma = 200)$$

a) Una muestra aleatoria de $n = 100$ personas, proporciona una media muestral de $\overline{x} = 1060$ €. Si el nivel de confianza es del 95 %:

$$1 - \alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$$

De manera que un intervalo de confianza al 95% para el ingreso mensual medio en ese municipio es:

$$IC_{0{,}95}(\mu) = \left(1060 - 1{,}96\,\frac{200}{\sqrt{100}}\,; 1060 + 1{,}96\,\frac{200}{\sqrt{100}}\right) = (1020{,}8; 1099{,}2)$$

b) Si el nivel de significación es del 1 % (nivel de confianza 99 %), entonces:

$$1 - \alpha = 0{,}99 \Rightarrow \frac{\alpha}{2} = 0{,}005 \Rightarrow z_{0{,}005} = 2{,}575$$

El error *E* para estimar el ingreso mensual medio debe ser menor que 30 €, por lo que:

$$E = 2{,}575\,\frac{200}{\sqrt{n}} < 30 \Rightarrow n > \left(\frac{2{,}575 \cdot 200}{30}\right)^2 = 294{,}7$$

Es decir, se requiere un tamaño muestral de, al menos, 295 personas para estimar el ingreso mensual medio con las condiciones propuestas.

c) Para que el error *E* se reduzca, al menos, a la mitad, es decir a $\frac{30}{2} = 15$; el tamaño muestral *n* debe cumplir:

$$E = z_{\frac{\alpha}{2}}\,\frac{\sigma}{\sqrt{n}} \le \frac{30}{2} \Rightarrow 2{,}575\,\frac{200}{\sqrt{n}} \le 15 \Rightarrow n \ge \left(\frac{2{,}575 \cdot 200}{15}\right)^2 = 1178{,}8$$

Por tanto, el tamaño muestral tendría que ser de, al menos, 1179 personas, para que el error se reduzca a la mitad.

44. La estatura de los alumnos de un colegio es una variable aleatoria que tiene una distribución normal de desviación típica 25 cm. Se ha elegido una muestra de 100 alumnos de ese colegio comprobándose que la estatura media es de 170 cm. Calcula:

a) El intervalo de confianza para la estatura media con un nivel de confianza del 99 %.

b) El error máximo de estimación que se ha cometido en el intervalo anterior.

c) El tamaño muestral mínimo necesario para conseguir, con un nivel de confianza del 95 %, un error máximo de 8 cm en la estimación de la estatura media.

La variable aleatoria X: "estatura, en cm, de los alumnos del colegio" tiene distribución normal, con la media μ desconocida y desviación típica 25 cm. Es decir,

$$X \sim N(\mu, \sigma = 25)$$

a) Una muestra aleatoria de $n = 100$ alumnos, proporciona una media muestral de $\bar{x} = 170$ cm. Si el nivel de confianza exigido es del 99 %, entonces:

$$1 - \alpha = 0{,}99 \Rightarrow \frac{\alpha}{2} = 0{,}005 \Rightarrow z_{0,005} = 2{,}575$$

De manera que un intervalo de confianza al 99% para la estatura media de los alumnos del colegio es:

$$IC_{0,99}(\mu) = \left(170 - 2{,}575\,\frac{25}{\sqrt{100}}\,;170 - 2{,}575\,\frac{25}{\sqrt{100}}\right) = (163{,}6;\,176{,}4)$$

b) El error máximo de estimación E que se comete con un nivel de confianza del 99 % es:

$$E = z_{\frac{\alpha}{2}}\,\frac{\sigma}{\sqrt{n}} = 2{,}575\frac{25}{\sqrt{100}} = 6{,}44$$

c) Si el nivel confianza es del 95 %, se tiene que :

$$1 - \alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0,025} = 1{,}96$$

Como el error E para estimar la estatura media debe ser menor que 8 cm, entonces:

$$E = 1{,}96\,\frac{25}{\sqrt{n}} < 8 \Rightarrow n > \left(\frac{1{,}96 \cdot 25}{8}\right)^2 = 37{,}52$$

Es decir, se requiere un tamaño muestral de, al menos, 38 alumnos para estimar la estatura media de los alumnos del colegio con las condiciones establecidas.

45. El tiempo de renovación de un teléfono móvil, expresado en años, se puede aproximar mediante una distribución normal con desviación típica 0,4 años.

a) Se toma una muestra aleatoria simple de 400 usuarios y se obtiene una media muestral igual a 1,75 años. Determina un intervalo de confianza al 95 % para el tiempo medio de renovación de un teléfono móvil.

b) Determina el tamaño muestral mínimo necesario para que el valor absoluto de la diferencia entre la media muestral y la media poblacional sea menor o igual a 0,02 años con un nivel de confianza del 90 %.

La variable aleatoria *X*: "tiempo, en años, de renovación de un teléfono móvil" tiene distribución normal, con la media μ desconocida y desviación típica 0,4 años. Es decir:

$$X \sim N(\mu; \sigma = 0{,}4)$$

a) Una muestra aleatoria de $n = 400$ usuarios, proporciona una media muestral de $\overline{x} = 1{,}75$ años. Si el nivel de confianza exigido es del 95 %, se cumple que:

$$1-\alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$$

De manera que un intervalo de confianza al 95%, para el tiempo medio de renovación de un teléfono móvil, es:

$$IC_{0{,}95}(\mu) = \left(1{,}75 - 1{,}96\,\frac{0{,}4}{\sqrt{400}}\,;\,1{,}75 + 1{,}96\,\frac{0{,}4}{\sqrt{400}}\right) = (1{,}7108; 1{,}7892) \text{ años}$$

b) Si el nivel de confianza es del 90 %, se tiene que :

$$1-\alpha = 0{,}90 \Rightarrow \frac{\alpha}{2} = 0{,}05 \Rightarrow z_{0{,}05} = 1{,}645$$

Como el error *E* para estimar el tiempo medio de renovación debe ser menor o igual que 0,02 años, entonces:

$$E = 1{,}645\,\frac{0{,}4}{\sqrt{n}} \le 0{,}02 \Rightarrow n \ge \left(\frac{1{,}645 \cdot 0{,}4}{0{,}02}\right)^2 = 1082{,}41$$

Es decir, con las condiciones establecidas, se requiere un tamaño muestral de, al menos, 1083 usuarios para estimar el tiempo medio de renovación de un teléfono móvil.

46. En una encuesta realizada en una población, se ha obtenido que 3700 de 4000 jóvenes encuestados tienen reproductor de música en formato MP3. Determina, justificando la respuesta:

a) La estimación puntual que podríamos dar para el porcentaje de jóvenes que poseen reproductor de música en formato MP3.

b) El error máximo que cometeríamos con dicha estimación, con una confianza del 90 %.

a) La variable *X*: "Número de jóvenes, de los 4000, que poseen un MP3" tiene distribución binomial $B(n = 4000, p)$, donde p es la proporción desconocida de jóvenes de la población que poseen un MP3.

Una estimación puntual para p es la proporción muestral, que en este caso es:

$$\hat{p} = \frac{3700}{4000} = 0{,}925$$

Es decir, se estima que el 92,5 % de los jóvenes posee un MP3.

b) Como el tamaño muestral es suficientemente grande, la distribución de la proporción muestral se puede aproximar por una distribución normal, con lo que, si la confianza es del 90 %, se tiene que:

$$1-\alpha = 0{,}90 \Rightarrow \frac{\alpha}{2} = 0{,}05 \Rightarrow z_{0{,}05} = 1{,}645$$

Y, por tanto, el error máximo *E* que se cometería es:

$$E = 1{,}645\sqrt{\frac{0{,}925(1-0{,}925)}{4000}} = 0{,}00685$$

47. La duración en kilómetros de los neumáticos de una cierta marca se puede aproximar por una variable aleatoria con distribución normal de media μ desconocida y desviación típica igual a 3000 kilómetros.

a) Se toma una muestra aleatoria simple de 100 neumáticos y se obtiene una media muestral de 48 000 kilómetros. Determínese un intervalo de confianza con un nivel del 90 % para μ.

b) Calcúlese el tamaño mínimo que debe tener la muestra para que el valor absoluto de la diferencia entre la media de la muestra y μ sea menor o igual a 1000 kilómetros con probabilidad mayor o igual que 0,95.

La variable aleatoria X: "duración, en kilómetros, de los neumáticos" tiene distribución normal, con la media μ desconocida y desviación típica 3000 km. Es decir:

$$X \sim N(\mu, \sigma = 3000)$$

a) Una muestra aleatoria de $n = 100$ neumáticos, proporciona una media muestral de $\bar{x} = 48\,000$ km. Si el nivel de confianza exigido es del 90 %, entonces:

$$1-\alpha = 0{,}90 \Rightarrow \frac{\alpha}{2} = 0{,}05 \Rightarrow z_{0{,}05} = 1{,}645$$

De manera que un intervalo de confianza al 90 % para la duración media μ de los neumáticos es:

$$IC_{0{,}90}(\mu) = \left(48000 - 1{,}645\,\frac{3000}{\sqrt{100}}\,;\ 48000 + 1{,}645\,\frac{3000}{\sqrt{100}}\right) = (47506{,}5;\ 48493{,}5)\ \text{km}$$

b) Si el nivel confianza debe ser del 95 %, se tiene que :

$$1-\alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$$

El tamaño muestral mínimo que debe tener la muestra para que el error E no supere los 1000 km es:

$$E = 1{,}96\,\frac{3000}{\sqrt{n}} \le 1000 \Rightarrow n \ge \left(\frac{1{,}96 \cdot 3000}{1000}\right)^2 = 34{,}5744$$

Es decir, se requiere una muestra aleatoria de, al menos, 35 neumáticos para estimar la duración media de los neumáticos de la marca, con las condiciones propuestas.

48. La antigüedad de los aviones comerciales sigue una distribución normal con una desviación típica de 8,28 años. Se selecciona una muestra de 40 aviones y la media de la antigüedad es de 13,41 años.

a) Obtén un intervalo de confianza del 90 % para la antigüedad media.

b) ¿Cuál debería ser el tamaño mínimo de la muestra para obtener un intervalo de confianza al 95 % con la misma amplitud que el anterior?

La variable X: "antigüedad, en años, de los aviones comerciales" tiene distribución normal $N(\mu; \sigma = 8{,}28)$, con μ desconocida.

Una muestra aleatoria de $n = 40$ aviones, proporciona una media muestral de antigüedad de $\bar{x} = 13{,}41$ años.

a) Si el nivel de confianza exigido es del 90 %, entonces:

$$1-\alpha = 0{,}90 \Rightarrow \frac{\alpha}{2} = 0{,}05 \Rightarrow z_{0{,}05} = 1{,}645$$

De manera que un intervalo de confianza al 90 %, para la antigüedad media μ de los aviones comerciales, es:

$$IC_{0{,}90}(\mu) = \left(13{,}41 - 1{,}645\,\frac{8{,}28}{\sqrt{40}}\,;\ 13{,}41 + 1{,}645\,\frac{8{,}28}{\sqrt{40}}\right) = (11{,}256;\ 15{,}564)$$

b) Si para seleccionar una muestra de tamaño n, el nivel de confianza es del 95 %, es decir:

$$1-\alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$$

El error E que se permite es el mismo del intervalo anterior, es decir la mitad de la longitud del intervalo:

$$E = 1{,}645\,\frac{8{,}28}{40} = 2{,}1536$$

Entonces, debe ser:

$$E = 1{,}96\,\frac{8{,}28}{\sqrt{n}} \le 2{,}1536 \Rightarrow n \ge \left(\frac{1{,}96 \cdot 8{,}28}{2{,}1536}\right)^2 = 56{,}79$$

Es decir, debe tomarse una muestra de al menos 57 aviones comerciales para estimar la antigüedad media μ con las condiciones propuestas.

CUESTIONES

49. Para estimar la proporción *p* de una característica *X* de una población mediante un intervalo de confianza de nivel 100(1 – α) %, ¿cuál es el error máximo a priori que se puede cometer?

El error máximo a priori que se puede cometer, es el error de estimación por intervalos de confianza y coincide con la mitad de la longitud del intervalo de confianza para la proporción *p*:

$$\left(\overline{x} \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{n}}\right)$$

Por tanto, dicho error es $E = z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{n}}$.

50. Un estimador T_1 es un 75 % más eficiente que otro estimador T_2. Si la varianza de T_1 es 6,25, ¿Cuál es la varianza del estimador T_2? El tamaño muestral es el mismo en ambos casos.

Un estimador T_1 es un más eficiente que otro T_2 si su varianza es menor, es decir, si: $\sigma_1^2 < \sigma_2^2$, siendo:

Estimador	T_1	T_2
Varianza	$\sigma_1^2 = 6{,}25$	σ_2^2

Se define, además, la eficiencia relativa entre los estimadores T_1 y T_2 como la razón $\frac{\sigma_1^2}{\sigma_2^2} = 0{,}75$.

Como el estimador T_1 es un 75 % más eficiente el estimador T_2 eso quiere decir que la varianza de T_1 es el 75 % de la varianza de T_2, es decir: $\sigma_1^2 = 0{,}75 \cdot \sigma_2^2$, por tanto $\sigma_2^2 = \frac{\sigma_1^2}{0{,}75} = \frac{6{,}25}{0{,}75} = 8{,}33$.

51. Para una muestra aleatoria se consideran la varianza y la cuasivarianza muestral como estimadores de la varianza de la población. ¿Cuál es la relación entre ambos estimadores?

Se define la varianza muestral como: $\sigma_x^2 = \frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}{n}$, y la cuasivarianza muestral como: $s_x^2 = \frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}{n-1}$.

La relación entre ambos estimadores es la siguiente $n\sigma_x^2 = \left(n-1\right)s_x^2 \Rightarrow s_x^2 = \frac{n}{n-1}\sigma_x^2 \Rightarrow \frac{s_x^2}{\sigma_x^2} = \frac{n}{n-1}$

La principal diferencia entre ambos es que la varianza, de una muestra aleatoria, es un estimador sesgado de la varianza poblacional, mientras que la cuasivarianza, de esa misma muestra, es un estimador insesgado, es decir, su media coincide con la de la varianza, por lo que, cuando se desconoce la varianza, se utiliza, preferentemente, la cuasivarianza muestral como estimador.

PROBLEMAS

52. El tiempo de espera en la cola de un supermercado sigue una distribución normal con media 180 segundos y desviación típica 50 segundos.

a) Se toma una muestra de 64 clientes. Calcula la probabilidad de que la media de la muestra supere los 190 segundos.

b) Calcula el intervalo de confianza al 95 % para la espera media, si la media muestral de los 64 clientes es de 190 segundos de espera.

c) Calcula el tamaño mínimo que deben tener las muestras de manera que el intervalo de confianza, al 90 %, para la espera media tenga longitud 22 segundos (semiamplitud 11 segundos)

Se considera la variable aleatoria X: "tiempo de espera, en segundos, en la cola de un supermercado". La distribución de X es $N(\mu = 180, \sigma = 50)$.

a) La media muestral de una muestra de $n = 64$ clientes tiene la siguiente distribución normal:

$$\overline{X} \sim N\left(\mu_{\overline{X}} = 180;\ \sigma^2_{\overline{X}} = \frac{50^2}{64} = 39{,}0625\right)$$

Con lo que la probabilidad de que la media muestral supere los 190 segundos es:

$$P\left(\overline{X} > 190\right) = P\left(Z > \frac{190-180}{\sqrt{39{,}0625}}\right) = P(Z > 1{,}6) = 1 - \Phi(1{,}6) = 1 - 0{,}9452 = 0{,}0548$$

b) Si la media de la muestra es de $\overline{x} = 190$ segundos, y el nivel de confianza es del 95%, con lo que:

$$1 - \alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$$

Un intervalo de confianza al 95 % para el tiempo de espera medio en ese supermercado es:

$$IC_{0{,}95}(\mu) = \left(190 - 1{,}96\frac{50}{8}\,;190 + 1{,}96\,\frac{50}{8}\right) = (177{,}75\ ;\ 202{,}25) \text{ minutos}$$

c) Si la confianza es del 90 %:

$$1 - \alpha = 0{,}90 \Rightarrow \frac{\alpha}{2} = 0{,}005 \Rightarrow z_{0{,}005} = 1{,}645$$

De manera que el tamaño de la muestra para que el error al estimar el tiempo de espera medio sea inferior a 11 segundos debe verificar:

$$n > 1{,}645^2\left(\frac{50}{11}\right)^2 = 55{,}91$$

Es decir, debe seleccionarse una muestra de al menos 56 clientes para cumplir las condiciones propuestas.

53. Las tensiones de ruptura de los cables fabricados por una empresa siguen una distribución normal $N(\mu, \sigma = 120)$. A partir de una muestra de 70 cables se ha obtenido una tensión media de ruptura de 2100 kilos. Halla un intervalo de confianza al 95 % para la tensión media de ruptura.

¿Qué tamaño deberá tener la muestra para obtener un intervalo de confianza al 99 % con una amplitud igual a la del anterior?

La variable X: "tensión de ruptura, en kilos, de los cables" tiene distribución normal $N(\mu, \sigma = 120)$, con μ desconocida.

Una muestra aleatoria de n = 70 cables, proporciona una media muestral para la tensión de ruptura de $\overline{x} = 2100$ kilos. Si el nivel de confianza exigido es del 95 %:

$$1-\alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$$

De manera que un intervalo de confianza al 95 % para la tensión de ruptura media μ de los cables es:

$$IC_{0{,}95}(\mu) = \left(2100 - 1{,}96\,\frac{120}{\sqrt{70}}\,;\ 2100 + 1{,}96\,\frac{120}{\sqrt{70}}\right) = (2071{,}888\,;\ 2128{,}112)$$

Si para seleccionar una muestra de tamaño n, el nivel de confianza que se exige es del 99 %:

$$1-\alpha = 0{,}99 \Rightarrow \frac{\alpha}{2} = 0{,}005 \Rightarrow z_{0{,}005} = 2{,}575$$

El error E que se permite es el mismo del intervalo de confianza anterior, es decir la mitad de la longitud del intervalo:

$$E = 1{,}96\,\frac{120}{\sqrt{70}} = 28{,}112$$

Por tanto, n debe cumplir:

$$2{,}575\,\frac{120}{\sqrt{n}} \le 28{,}112 \Rightarrow n \ge \left(\frac{2{,}575 \cdot 120}{28{,}112}\right)^2 = 120{,}82$$

Es decir, debe tomarse una muestra de, al menos, 121 cables para estimar la resistencia media μ a la ruptura con las condiciones propuestas.

54. La valoración de las instituciones por parte de los ciudadanos se mide en unas unidades ficticias que se pueden denominar como "u". Se sabe que, en el caso de los españoles, dicha valoración sigue una distribución normal con desviación típica 25 u.

a) Se elige una muestra de 100 españoles, dando una media de 180 u. Calcula un intervalo de confianza para la media poblacional de la valoración de las instituciones, con una confianza del 90 %.

b) Si conocemos que la media poblacional es 182 u, calcula la probabilidad de que una muestra de tamaño 100 tenga media inferior a 180 u.

Se considera la variable aleatoria *X*: "valoración, en unidades u, de las instituciones por parte de los ciudadanos". La variable *X* tiene una distribución normal:

$$X \sim N(\mu, \sigma = 25)$$

a) Una muestra de $n = 100$, ciudadanos españoles, proporciona una media muestral $\overline{x} = 180$ u. Si el nivel de confianza del intervalo es del 90 %, entonces:

$$1 - \alpha = 0{,}90 \Rightarrow \frac{\alpha}{2} = 0{,}05 \Rightarrow z_{0,05} = 1{,}645$$

De modo que el intervalo de confianza para la valoración media poblacional μ es:

$$IC_{0,90}(\mu) = \left(180 - 1{,}645\,\frac{25}{\sqrt{100}}\,;\,180 + 1{,}645\,\frac{25}{\sqrt{100}}\right) = (175{,}89;\,184{,}11)$$

b) En este caso, $X \sim N(\mu = 182, \sigma = 25)$; con lo que la media muestral de una muestra de tamaño $n = 100$, tiene una distribución normal:

$$\overline{X} \sim N\left(\mu_{\overline{X}} = 182, \sigma^2_{\overline{X}} = \frac{25^2}{100}\right)$$

Por lo que, la probabilidad de que la media muestral sea inferior a 180 u, es:

$$P\left(\overline{X} < 180\right) = P\left(Z < \frac{180 - 182}{\frac{25}{10}}\right) = P(Z < -0{,}8) = 1 - \Phi(0{,}8) = 1 - 0{,}7881 = 0{,}2119$$

55. La altura de los edificios de una ciudad sigue una distribución normal de media desconocida y desviación típica 20 m.

a) Calcula el tamaño mínimo que ha de tener una muestra aleatoria de dichos edificios para que el error cometido al estimar la altura media sea inferior a 2 m, con un nivel de confianza del 97 %.

b) Halla el intervalo de confianza al 95 % para la media de la altura de los edificios si una muestra aleatoria del tamaño calculado en el apartado anterior, proporcionó una media muestral de 25,5 metros.

Sea la variable aleatoria *X*: "altura, en metros, de los edificios de la ciudad". La distribución de *X* es normal de media desconocida μ y desviación típica $\sigma = 20$ m.

a) Si el nivel de confianza es del 97 %, entonces: $1-\alpha = 0{,}97 \Rightarrow \frac{\alpha}{2} = 0{,}015 \Rightarrow z_{0{,}015} = 2{,}17$

Y como el error cometido *E* para estimar la altura media debe ser menor que 2, resulta:

$$E = 2{,}17\,\frac{20}{\sqrt{n}} < 2 \Rightarrow n > \left(\frac{2{,}17\cdot 20}{2}\right)^2 = 470{,}89$$

Luego el tamaño mínimo de la muestra de edificios debe ser, al menos, de 471, para estimar la altura media con las condiciones requeridas.

b) Con una muestra aleatoria de $n = 471$ edificios, el intervalo de confianza al 95 % para la altura media de los edificios de la población es:

$$1-\alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$$

$$IC_{0{,}95}(\mu) = \left(25{,}5 - 1{,}96\frac{20}{\sqrt{471}}\,;\ 25{,}5 + 1{,}96\frac{20}{\sqrt{471}}\right) = (23{,}69;\ 27{,}31)$$

56. Se supone que el gasto que hacen los individuos de una determinada población en regalos de Navidad se puede aproximar por una variable aleatoria con distribución normal de media μ y desviación típica igual a 45 €.

a) Se toma una muestra aleatoria y se obtiene el intervalo de confianza (251,6; 271,2) para μ, con un nivel de confianza del 95 %. Calcúlese la media muestral y el tamaño de la muestra elegida.

b) Se toma una muestra aleatoria de tamaño 64 para estimar µ. Calcula el error máximo cometido por esa estimación con un nivel de confianza del 90 %.

La variable aleatoria *X*: "gasto, en euros, en regalos de Navidad", se puede aproximar por una distribución normal $N(\mu, \sigma = 45)$, con μ desconocida.

a) Dada una estimación puntual $\bar{x}$ para μ ,un intervalo de confianza al 95 % tiene la forma:

$$IC_{1-\alpha}(\mu) = \left(\bar{x} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\,,\ \bar{x} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$

De modo que la media muestral $\bar{x}$, es el punto medio del intervalo de confianza. Por tanto:

$$\bar{x} = \frac{251{,}6 + 271{,}2}{2} = 261{,}4$$

Si la confianza es del 95 %, entonces: $1-\alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$

Igualando, por ejemplo, el extremo superior del intervalo dado; 271,2; al del intervalo de confianza genérico al 95 %, se obtiene el tamaño de la muestra:

$$261{,}4 + 1{,}96\,\frac{45}{\sqrt{n}} = 271{,}2 \Rightarrow n = \left(\frac{1{,}96\cdot 45}{9{,}8}\right)^2 = 81$$

NOTA: el mismo resultado se obtiene si se igualan los extremos inferiores.

b) Si el tamaño de la muestra es de $n = 64$ y el nivel de confianza del 90 %, con lo que:

$$1-\alpha = 0{,}90 \Rightarrow \frac{\alpha}{2} = 0{,}05 \Rightarrow z_{0{,}05} = 1{,}645$$

El error máximo *E* que se cometería al estimar μ es:

$$E = 1{,}645\,\frac{45}{8} = 9{,}253 \text{ euros}$$

57. La proporción de mujeres de una población portadoras de hemofilia es desconocida. Para estimarla, se elige una muestra aleatoria de 500 mujeres entre las que se encontraron 80 portadoras de la enfermedad.

a) Calcula un intervalo del 95 % de confianza para la proporción de mujeres portadoras de hemofilia de esa población.

b) Suponiendo que aún no se tomó la muestra y queremos hacer una estimación cometiendo un error no superior al 2 %, con 95 % de confianza, ¿de qué tamaño debería ser la muestra?

Se considera la variable X: "número de mujeres, de las 500, que son portadoras de hemofilia". La distribución de X es binomial $B(n = 500, p)$, con p, proporción de mujeres hemofílicas, desconocida.

En la muestra aleatoria de $n = 500$ mujeres, se encontró que 80 eran portadoras de la enfermedad. Por lo que una estimación puntual de p es:

$$\hat{p} = \frac{80}{500} = 0{,}16$$

a) Dado que n es suficientemente grande, la distribución de la proporción muestral puede aproximarse por la de una normal. Para un nivel de confianza del 95 % resulta:

$$1 - \alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$$

De manera que un intervalo de confianza al 95 % para la proporción de p de mujeres hemofílicas en la población es:

$$IC_{0{,}95}(p) = \left(0{,}16 - 1{,}96\sqrt{\frac{0{,}16(1-0{,}16)}{500}};\ 0{,}16 + 1{,}96\sqrt{\frac{0{,}16(1-0{,}16)}{500}}\right) = (0{,}1279;\ 0{,}1921)$$

b) Para seleccionar una muestra de tamaño n, bajo supuesto de normalidad, si el nivel de confianza que se exige es del 95 %, se tiene que, nuevamente que $z_{0{,}025} = 1{,}96$.

Entonces, como el error E que se tolera no debe ser mayor de 0,02 (es decir, 2 %):

$$E = 1{,}96\sqrt{\frac{0{,}16(1-0{,}16)}{n}} \le 0{,}02 \Rightarrow n \ge \left(\frac{1{,}96}{0{,}02}\right)^2 0{,}16(1-0{,}16) = 1290{,}78$$

Es decir, debe tomarse una muestra de, al menos, 1291 mujeres para estimar la proporción p de mujeres hemofílicas en la población con las condiciones propuestas.

58. Una muestra de 2000 familias es seleccionada aleatoriamente en cierta ciudad. Se comprueba que 300 de ellas disponen de acceso a internet desde su domicilio. Determina justificando la respuesta:

a) El intervalo de confianza al 99 % para el porcentaje de familias de esa ciudad que disponen de acceso a internet desde su domicilio.

b) El error máximo que se comete, con una confianza del 99 %, si se estima que dicho porcentaje es un 15 %.

Se considera la variable *X*: "la familia tiene acceso a internet desde su domicilio", con distribución de probabilidad $\text{Ber}(p)$, siendo p la probabilidad de que, en dicha ciudad, una familia elegida al azar tenga acceso a internet desde su domicilio.

En una muestra aleatoria de 2000 familias, 300 de ellas disponen de acceso a internet desde el domicilio. Por tanto una estimación puntual de p es:

$$\hat{p} = \frac{300}{2000} = 0{,}15$$

a) Dado que la muestra es suficientemente grande, la distribución de la proporción muestral puede aproximarse por una distribución normal, es decir:

$$\hat{p} \sim N\left(p; \sigma^2 = \frac{0{,}15(1-0{,}15)}{2000} = 0{,}000\,063\,75\right)$$

Si el intervalo de confianza es al 99 %, entonces: $1-\alpha = 0{,}99 \Rightarrow \frac{\alpha}{2} = 0{,}005 \Rightarrow z_{0{,}005} = 2{,}575$

De modo que un intervalo de confianza al 99 % para la proporción *p* es:

$$IC_{0{,}99}(p) = \left(0{,}15 - 2{,}575\sqrt{0{,}000\,063\,75}\,;\, 0{,}15 + 2{,}575\sqrt{0{,}000\,063\,75}\right) = (0{,}1294; 0{,}1706)$$

Es decir, un intervalo de confianza al 99 % para el porcentaje de familias de esa ciudad que tiene acceso a internet desde su domicilio es $(12{,}94\ \%\,;\, 17{,}06\ \%)$

b) Como $\hat{p} = 0{,}15$; que es el mismo valor que en el apartado anterior, el error máximo *E* que se cometería, con una confianza del 99 %, es la semilongitud del intervalo calculado en el dicho apartado, es decir:

$$E = 2{,}575\sqrt{0{,}000\,063\,75} = 0{,}0206$$

Es decir, aproximadamente 2,06 %.

NOTA: En el apartado b), el error máximo *E* cometido en la estimación por intervalo también se puede obtener directamente de la semilongitud del intervalo calculado en el apartado a). Es decir:

$$E = \frac{0{,}1706 - 0{,}1294}{2} = 0{,}0206$$

59. Un fabricante de automóviles ha realizado un estudio de mercado, en un determinado municipio, tomando una nuestra de 500 turismos, y ha encontrado que 80 tienen motor diésel. Para un nivel de confianza del 94 %:

a) Determina el intervalo de confianza de la proporción de turismos que tiene motor diésel en este municipio.

b) ¿Cuál es el error máximo de la estimación?

Se considera la variable *X*: "el turismo tiene motor diésel", con distribución de probabilidad $\text{Ber}(p)$, siendo p la probabilidad de que, en dicho municipio, un automóvil elegido al azar tenga motor diésel.

En una muestra aleatoria de 500 turismos, 80 de ellos tienen motor diésel. Por tanto una estimación puntual de p es:

$$\hat{p} = \frac{80}{500} = 0{,}16$$

a) Dado que la muestra es suficientemente grande, la distribución de la proporción muestral puede aproximarse por una distribución normal, es decir:

$$\hat{p} \sim N\left(\mu = p; \sigma^2 = \frac{0{,}16(1-0{,}16)}{500} = 0{,}000\,268\,8\right)$$

Si el intervalo de confianza es al 94 %, entonces: $1-\alpha = 0{,}94 \Rightarrow \frac{\alpha}{2} = 0{,}03 \Rightarrow z_{0{,}03} = 1{,}88$

De modo que un intervalo de confianza al 94 % viene dado por:

$$IC_{0{,}94}(p) = \left(0{,}16 - 1{,}88 \Rightarrow \sqrt{0{,}0002688}\,;\, 0{,}16 + 1{,}88 \Rightarrow \sqrt{0{,}0002688}\right) = (0{,}1292\,;\, 0{,}1908)$$

Es decir, un intervalo de confianza al 94 % para el porcentaje de familias de esa ciudad que tiene acceso a internet desde su domicilio es $(12{,}92\ \%;\ 19{,}08\ \%)$

b) El error máximo *E* que se cometería al estimar la proporción p, con una confianza del 99 %, es la semilongitud del intervalo calculado en el apartado anterior, es decir:

$$E = 1{,}88\sqrt{0{,}000\,268\,8} = 0{,}0308$$

Por tanto, es, aproximadamente 3,08 %.

60. Se hace una encuesta sobre el nivel de conocimientos generales de los estudiantes de bachillerato de una determinada ciudad. Para ello se ha elegido una muestra aleatoria de 9 de esos estudiantes, a los que se les ha realizado un examen. Las calificaciones obtenidas han sido las siguientes:

7,8 6,5 5,4 7,1 5,0 8,3 5,6 6,6 6,2

Se supone que la variable objeto de estudio sigue una distribución normal de desviación típica 1. Determina un intervalo de confianza al 98 % para la media de las calificaciones del examen.

La variable aleatoria *X*. "calificaciones del examen" tiene distribución normal $N(\mu, \sigma = 1)$ de media desconocida.

Para estimar μ se elige una muestra de tamaño $n = 9$, que proporciona una media muestral:

$$\bar{x} = \frac{7{,}8 + 6{,}5 + 5{,}4 + 7{,}1 + 5{,}0 + 8{,}3 + 5{,}6 + 6{,}6 + 6{,}2}{9} = 6{,}5$$

Que es una estimación puntual de la media μ.

Si la confianza es del 98 %, entonces: $1-\alpha = 0{,}98 \Rightarrow \frac{\alpha}{2} = 0{,}01 \Rightarrow z_{0{,}01} = 2{,}33$

De manera que un intervalo de confianza al 98 % para la media de las calificaciones μ viene dado por:

$$IC_{0{,}98}(\mu) = \left(6{,}5 - 2{,}33\,\frac{1}{\sqrt{9}}\,;\, 6{,}5 + 2{,}33\,\frac{1}{\sqrt{9}}\right) = (5{,}723\,;\, 7{,}277)$$

61. Se considera una variable aleatoria con distribución normal de media μ y desviación típica igual a 210. Se toma una muestra aleatoria simple de 64 elementos.

a) Calcula la probabilidad de que el valor absoluto de la diferencia entre la media muestral y la media μ de la población sea mayor o igual que 22.

b) Determina un intervalo de confianza del 99 % para μ, si la media muestral es igual a 1532.

Se considera la variable aleatoria $X \sim N(\mu, \sigma = 210)$.

Si se selecciona una muestra aleatoria simple de tamaño n = 64, la media muestral $\overline{X}$, tiene también distribución normal:

$$\overline{X} \sim N\left(\mu_{\overline{X}} = \mu;\ \sigma^2_{\overline{X}} = \frac{210^2}{64} = 689{,}0625\right)$$

a) Para calcular la probabilidad pedida, calculamos previamente la del suceso contrario, es decir:

$$P\left(\left|\overline{X} - \mu\right| \geq 22\right) = 1 - P\left(\left|\overline{X} - \mu\right| < 22\right)$$

Desarrollando y tipificando a la $Z \sim N(0,1)$:

$$P\left(\left|\overline{X} - \mu\right| < 22\right) = P\left(-22 < \overline{X} - \mu < 22\right) = P\left(\frac{-22}{\sqrt{689{,}0625}} < \frac{\overline{X} - \mu}{\sqrt{689{,}0625}} < \frac{22}{\sqrt{689{,}0625}}\right) =$$

$$= P(-0{,}84 < Z < 0{,}84) = 2\Phi(0{,}84) - 1 = 2 \cdot 0{,}7995 - 1 = 0{,}5991$$

Con lo que, finalmente, se obtiene que:

$$P\left(\left|\overline{X} - \mu\right| \geq 22\right) = 1 - P\left(\left|\overline{X} - \mu\right| < 22\right) = 1 - 0{,}5991 = 0{,}4009$$

b) Como la confianza es del 99 % entonces: $1 - \alpha = 0{,}99 \Rightarrow \frac{\alpha}{2} = 0{,}005 \Rightarrow z_{0{,}005} = 2{,}575$. Si $\overline{x} =$ 1532 es una estimación puntual de la media μ, un intervalo de confianza al 99 % para la media poblacional es:

$$IC_{0{,}99}(\mu) = \left(1532 - 2{,}575\sqrt{689{,}0625}\,; 1532 + 2{,}575\sqrt{689{,}0625}\right) = (1464{,}41\,; 1599{,}59)$$

62. El tiempo de espera para ser atendido en un restaurante se puede aproximar por una variable aleatoria con distribución normal de media μ desconocida y desviación típica igual a 3 minutos. Se toma una muestra aleatoria simple de tamaño 121.

a) Calcula la probabilidad de que el valor absoluto de la diferencia entre la media de la muestra y la media μ de la población sea mayor que 0,5 minutos.

b) Determina un intervalo de confianza con un nivel del 95 % para μ, si la media de la muestra es igual a 7 minutos.

La distribución de la variable aleatoria *X*: "tiempo de espera, en minutos, para ser atendido en un restaurante", puede aproximarse por la de una distribución normal:

$$X \sim N(\mu, \sigma = 3)$$

La distribución de la media muestral $\overline{X}$ de una muestra aleatoria de tamaño $n = 121$ es también normal:

$$\overline{X} \sim N\left(\mu_{\overline{X}} = \mu; \sigma^2_{\overline{X}} = \frac{3^2}{121} = 0{,}07438\right)$$

a) Para calcular la probabilidad pedida, calculamos previamente la del suceso contrario:

$$P\left(\left|\overline{X} - \mu\right| > 0{,}5\right) = 1 - P\left(\left|\overline{X} - \mu\right| \le 0{,}5\right)$$

Desarrollando y tipificando a la $Z \sim N(0;1)$:

$$P\left(\left|\overline{X} - \mu\right| \le 0{,}5\right) = P\left(-0{,}5 \le \overline{X} - \mu \le 0{,}5\right) = P\left(\frac{-0{,}5}{\sqrt{0{,}07438}} \le \frac{\overline{X} - \mu}{\sqrt{0{,}07438}} \le \frac{0{,}5}{\sqrt{0{,}07438}}\right)$$

$$= P(-1{,}83 \le Z \le 1{,}83) = 2\Phi(1{,}83) - 1 = 2 \cdot 0{,}9664 - 1 = 0{,}9328$$

Por tanto: $P\left(\left|\overline{X} - \mu\right| \ge 0{,}5\right) = 1 - P\left(\left|\overline{X} - \mu\right| \le 0{,}5\right) = 1 - 0{,}9328 = 0{,}0672$

b) Si la confianza es del 95 % entonces: $1 - \alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$; y puesto que $\overline{x} = 7$ es una estimación puntual de la media μ, un intervalo de confianza al 95 % para el tiempo medio de espera en el restaurante es:

$$IC_{0{,}95}(\mu) = \left(7 - 1{,}96\sqrt{0{,}07438}\ ;\ 7 + 1{,}96\sqrt{0{,}07438}\right) = (6{,}47;\ 7{,}53)$$

63. En una encuesta, se ha preguntado a 10 000 estudiantes de Bachillerato sobre su consumo semanal de refrescos y se ha calculado una media de 5 botes con una desviación típica de 2 botes.

a) Determina el intervalo de confianza para la media a un nivel del 95 %.

b) Si se acepta un error máximo de 0,25 botes para la media poblacional, y si queremos un nivel de confianza del 95 %, ¿cuántas personas es necesario entrevistar como mínimo?

Se considera la variable *X*: "número de botes de refrescos consumidos en una semana". Dado que la muestra es suficientemente grande, la distribución de la media muestral $\overline{X}$, se puede aproximar por la siguiente distribución normal:

$$X \sim N\left(\mu_{\overline{X}} = 5, \sigma^2_{\overline{X}} = \frac{4}{10000}\right)$$

a) Un intervalo de confianza para la media μ al 95 % viene dado por:

$$IC_{0{,}95} = \left(5 - 1{,}96\,\frac{2}{100}; 5 + 1{,}96\,\frac{2}{100}\right) = (4{,}96\ ;\ 5{,}04)$$

Para una confianza del 95% resulta: $1 - \alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$

b) Con el mismo nivel de confianza que en el apartado anterior, si el error *E* debe ser menor de 0,25 botes, entonces:

$$E = 1{,}96\,\frac{2}{\sqrt{n}} \le 0{,}25 \Rightarrow n \ge \left(\frac{1{,}96 \cdot 2}{0{,}25}\right)^2 = 245{,}86$$

Se necesita, por tanto, entrevistar, al menos, a 246 personas para asegurar las condiciones establecidas.

64. El peso (en gramos) de las manzanas de un agricultor es aleatorio, con distribución normal de desviación típica igual a 30 g. Queremos construir un intervalo de confianza para la media del peso de las manzanas del agricultor.

a) Determina el tamaño de la muestra para que el intervalo de confianza del 98 % tenga una amplitud menor o igual que 10 g.

b) Si se toma una muestra de tamaño 100, se pesan las 100 manzanas y se calcula su promedio, que es igual a 160 g, construye el intervalo de confianza del 98 % para la media del peso de las manzanas del agricultor.

La variable aleatoria X: "peso, en gramos, de las manzanas" tiene una distribución $N(\mu, \sigma = 30)$, con μ desconocida.

a) La confianza del 98 % implica que: $1-\alpha = 0{,}98 \Rightarrow \frac{\alpha}{2} = 0{,}01 \Rightarrow z_{0,01} = 2{,}33$

Para que la amplitud del intervalo (simétrico respecto a la media muestral) sea menor o igual que 10 g, que el error E debe ser inferior o igual a 5 g, entonces:

$$E = 2{,}33\,\frac{30}{\sqrt{n}} \le 5 \Rightarrow n \ge \left(\frac{2{,}33 \cdot 30}{5}\right)^2 = 195{,}44$$

De manera que el tamaño mínimo de la muestra para que se satisfagan las condiciones propuestas debe ser de 196 manzanas.

b) En este caso, $n = 100$ y la estimación puntual para el peso medio es $\bar{x} = 160$ g. De esta manera, un intervalo de confianza al 98 % para el peso medio de las naranjas, μ es:

$$IC_{0,98} = \left(160 - 2{,}33\,\frac{30}{\sqrt{100}}; 160 + 2{,}33\,\frac{30}{\sqrt{100}}\right) = (153{,}01; 166{,}99)$$

65. La temperatura durante los meses de verano en una ciudad sigue una distribución normal con una desviación típica de 5 °C. Elegida una muestra y con un nivel de confianza del 98 %, se obtiene el intervalo (25 °C, 30 °C). Calcula la media y el tamaño de la muestra elegida. Detalla los pasos realizados para obtener los resultados.

La variable X: "temperatura, en ºC, en los meses de verano" tiene una distribución normal $N(\mu, \sigma = 5)$, con la media μ desconocida.

Como el intervalo de confianza, al 98 %, es (25, 30); la media de la muestra es el punto medio de ese intervalo de confianza, es decir:

$$\bar{x} = \frac{25+30}{2} = 27{,}5\ ºC$$

Para calcular el tamaño de la muestra, se tiene en cuenta que el máximo error cometido E al estimar la media μ mediante un intervalo de confianza al 98 % es la semilongitud del intervalo.

Como $1-\alpha = 0{,}98 \Rightarrow \frac{\alpha}{2} = 0{,}01 \Rightarrow z_{0,01} = 2{,}33$

Igualando el error máximo E cometido con la semilongitud del intervalo, se obtiene:

$$E = 2{,}33\,\frac{5}{\sqrt{n}} = 2{,}5 \Rightarrow n = \left(\frac{2{,}33 \cdot 5}{2{,}5}\right)^2 = 21{,}72$$

En consecuencia, el tamaño de la muestra elegida debe ser, aproximadamente, de 22 días.

66. En una encuesta electoral se ha preguntado, entre otras cosas, por la intención de voto a determinado partido político *A*. La ficha técnica informa que se entrevistó a 1200 personas y que el error es inferior a ±2,7 %. De las 1200 personas, 340 dijeron que votarían al partido *A*. ¿Cuál es el nivel de confianza de esta afirmación?

Si se desea obtener una estimación por intervalo con una precisión de ±1,5 %, al 97 % de confianza, ¿cuál debe ser el número de personas entrevistadas?

Una estimación puntual de la proporción personas que tienen intención de votar al partido *A* es:

$$\hat{p} = \frac{340}{1200} = 0{,}2833$$

Dado que el tamaño de la muestra es suficientemente grande, la distribución de la proporción muestral se puede aproximar por la de una distribución normal. De esta manera, como el error máximo cometido con esta estimación es 0,027. Se tiene que:

$$z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{n}} = 2{,}7\ \% \Rightarrow z_{\frac{\alpha}{2}}\sqrt{\frac{0{,}2833\left(1-0{,}2833\right)}{1200}} = 0{,}027 \Rightarrow z_{\frac{\alpha}{2}} = 2{,}08 \Rightarrow 1-\frac{\alpha}{2} = 0{,}9812 \Rightarrow 1-\alpha = 0{,}9624$$

De manera que el nivel de confianza de esa afirmación es del 96,24 %.

Si, al 97 % de confianza, se desea un intervalo con un error máximo de 0,015 (1,5 %), entonces, como $1-\alpha = 0{,}97 \Rightarrow \frac{\alpha}{2} = 0{,}015 \Rightarrow z_{0{,}015} = 2{,}17$; resulta que:

$$2{,}17\sqrt{\frac{0{,}2833\left(1-0{,}2833\right)}{n}} \le 0{,}015 \Rightarrow n \ge \left(\frac{2{,}17}{0{,}015}\right)^2 0{,}2833\left(1-0{,}2833\right) = 4249{,}33$$

Por tanto, se requieren, al menos, 4250 entrevistas para conseguir las condiciones propuestas.

67. La cantidad de horas que duermen los vecinos de un pueblo se puede aproximar por una distribución normal con una desviación típica de 0,64 horas. Se toma una muestra aleatoria simple y se obtienen los siguientes datos, en horas, que duermen cada noche:

6,9	**7,6**	**6,5**	**6,2**	**7,8**	**7,0**	**5,5**	**7,6**
7,3	**6,6**	**7,1**	**6,9**	**6,7**	**6,5**	**7,2**	**5,8**

a) Calcula la media muestral del número de horas que se duerme cada noche.

b) Determina el nivel de confianza para el cual el intervalo de confianza de la media de horas que se duerme cada noche es (6,65; 7). Detalla los pasos realizados para obtener los resultados.

La variable *X*: "cantidad de horas que duermen los vecinos" tiene distribución $N\left(\mu;\sigma = 0{,}64\right)$.

a) La muestra aleatoria de tamaño $n = 16$, proporciona una estimación puntual de la media μ:

$$\bar{x} = \frac{6{,}9+7{,}6+6{,}5+6{,}2+7{,}8+7{,}0+5{,}5+7{,}6+7{,}3+6{,}6+7{,}1+6{,}9+6{,}7+6{,}5+7{,}2+5{,}8}{16} = 6{,}825 \text{ horas}$$

b) Si el intervalo de confianza para la media μ, de horas que se duerme cada noche, es (6,65; 7), el error máximo *E* que se comete es la semilongitud del intervalo, es decir:

$$E = \frac{7-6{,}65}{2} = 0{,}175$$

Para un nivel de confianza del $100\left(1-\alpha\right)\%$, con n=16, se tiene que:

$$E = z_{\frac{\alpha}{2}}\frac{0{,}64}{\sqrt{16}} = 0{,}175 \Rightarrow z_{\frac{\alpha}{2}} = \frac{0{,}175\cdot 4}{0{,}64} = 1{,}09 \Rightarrow 1-\frac{\alpha}{2} = 0{,}8621 \Rightarrow 1-\alpha = 0{,}7242$$

Por tanto el nivel de confianza es del 72,42 %.

68. Se desea estimar la proporción de individuos con sobrepeso en una población. Para ello se va a tomar una muestra aleatoria simple y se va a determinar, de cada individuo, si tiene sobrepeso o no, y a partir de los resultados se construirá un intervalo de confianza para la proporción de individuos con sobrepeso en la población. El intervalo se hará a un nivel de confianza del 96 %.

a) Si se quiere que el intervalo no tenga una amplitud mayor que 0,1; ¿qué tamaño de la muestra debemos escoger?

b) Se decide tomar una muestra de tamaño 200 individuos, de los cuales 40 tienen sobrepeso. Calcula el intervalo de confianza al 96 % para la proporción de individuos con sobrepeso en la población.

Dado que el intervalo debe tener un nivel del confianza del 96 %, entonces: $1-\alpha=0{,}96 \Rightarrow \frac{\alpha}{2}=0{,}02 \Rightarrow z_{0{,}02}=2{,}05$.

a) Si la amplitud del intervalo no debe ser mayor que 0,1; ello quiere decir que el error máximo E permitido es 0,05. Como no se dispone de estimación puntual para la proporción poblacional de individuos con sobrepeso, se toma $\hat{p}=0{,}5$ (que representa la situación más desfavorable). De esta manera:

$$E=2{,}05\sqrt{\frac{0{,}5(1-0{,}5)}{n}}\leq 0{,}05 \Rightarrow n\geq\left(\frac{2{,}05}{0{,}05}\right)^2 0{,}5\cdot 0{,}5=420{,}25$$

Se debe elegir, por tanto una muestra de, al menos, 421 individuos para cumplir las condiciones propuestas.

b) Una estimación puntual $\hat{p}$ de la proporción p de individuos con sobrepeso en la población es:

$$\hat{p}=\frac{40}{200}=0{,}2$$

Con lo que un intervalo de confianza al 96 % para p viene dado por:

$$IC_{0{,}96}(p)=\left(0{,}2-2{,}05\sqrt{\frac{0{,}2(1-0{,}2)}{200}};\ 0{,}2+2{,}05\sqrt{\frac{0{,}2(1-0{,}2)}{200}}\right)=(0{,}142;\ 0{,}258)$$

La muestra es lo suficientemente grande como para poder asumir que la distribución de la proporción muestral es aproximadamente normal.

69. Se desea estudiar la proporción de alumnos repetidores de una región.

a) Se ha tomado una muestra de 200 alumnos, de los cuales 30 son repetidores. Construye un intervalo de confianza, con un nivel del 95 %, para estimar la proporción de alumnos repetidores de la región.

b) ¿Cuál debe ser el tamaño mínimo muestral para poder estimar la proporción de alumnos repetidores de la región con un error máximo de estimación del 2 % si se desea un nivel de confianza del 90 %?

La variable *X*: “el alumno es repetidor” tiene una distribución de Bernoulli $Ber(p)$, con p desconocido. Con los datos de la muestra, una estimación puntual de p es:

$$\hat{p}=\frac{30}{200}=0{,}15$$

a) La distribución de la proporción muestral puede aproximarse por la de una distribución normal ya que el tamaño de la muestra es suficientemente grande.

Como el nivel de confianza es del 95 % se tiene que: $1-\alpha=0{,}95 \Rightarrow \frac{\alpha}{2}=0{,}025 \Rightarrow z_{0{,}025}=1{,}96$

El intervalo de confianza para estimar p viene dado por:

$$IC_{0{,}95}(p)=\left(0{,}15-1{,}96\sqrt{\frac{0{,}15(1-0{,}15)}{200}};\ 0{,}15+1{,}96\sqrt{\frac{0{,}15(1-0{,}15)}{200}}\right)=(0{,}1005;\ 0{,}1995)$$

b) Como el nivel de confianza es del 90 % se tiene que: $1-\alpha=0{,}90 \Rightarrow \frac{\alpha}{2}=0{,}05 \Rightarrow z_{0{,}05}=1{,}645$

Utilizando 0,5 como el caso más desfavorable para estimar la proporción desconocida y que el error máximo E es de 0,02; se tiene que:

$$E=1{,}645\sqrt{\frac{0{,}5\cdot 0{,}5}{n}}\leq 0{,}02 \Rightarrow n\geq\left(\frac{1{,}645}{0{,}02}\right)^2 0{,}5^2=1691{,}3$$

Por lo que el tamaño mínimo de la muestra, para que se satisfagan las condiciones propuestas, debe ser de 1692 alumnos.

70. El mínimo tamaño muestral necesario para estimar la media de una variable aleatoria con distribución normal de desviación típica σ, con un error máximo de 3,29 y un nivel de confianza del 90 %, supera en 7500 unidades al que se necesitaría si el nivel de confianza fuera del 95 % y el error máximo fuera de 7,84.

Exprésense los tamaños muestrales en función de la desviación típica σ y calcula la desviación típica de la población y los tamaños muestrales respectivos.

La variable X tiene una distribución normal: $X \sim N(\mu,\sigma)$

En el primer caso, como el nivel de confianza es del 90 % se tiene que: $1-\alpha=0{,}90 \Rightarrow \frac{\alpha}{2}=0{,}05 \Rightarrow z_{0{,}05}=1{,}645$

Además, si el error máximo E_1 es 3,29 y el tamaño de la muestra es n, se tiene que:

$$E_1=1{,}645\frac{\sigma}{\sqrt{n}}=3{,}29 \Rightarrow n=\left(\frac{1{,}645\sigma}{3{,}29}\right)^2 \Rightarrow n=0{,}25\sigma^2$$

En el segundo caso, como el nivel de confianza es del 95 % se tiene que: $1-\alpha=0{,}95 \Rightarrow \frac{\alpha}{2}=0{,}025 \Rightarrow z_{0{,}025}=1{,}96$

Además, si el error máximo E_2 es 7,84, y el tamaño de la muestra es m, se tiene que:

$$E_2=1{,}96\frac{\sigma}{\sqrt{m}}=7{,}84 \Rightarrow m=\left(\frac{1{,}96\sigma}{7{,}84}\right)^2 \Rightarrow m=0{,}0625\sigma^2$$

Como, $n=m+7500$, simplificando y despejando σ^2 se obtiene:

$$0{,}25\sigma^2=0{,}0625\sigma^2+7500 \Rightarrow \sigma^2=\frac{7500}{0{,}1875}=40000 \Rightarrow \sigma=200$$

Y el tamaño de las muestras es:

$$n=0{,}25\sigma^2=0{,}25\cdot 40000=10000 \qquad m=0{,}0625\cdot 40000=2500$$

71. Una industria conservera envasa latas de sardinas, cuyo peso, en gramos, sigue una distribución normal con media μ y desviación típica 1 g. Para estimar el valor de μ, se selecciona una muestra de 30 latas, que proporciona el intervalo (79,81; 80,49) como estimador de μ.

a) Calcula el peso medio de la muestra de 30 latas.

b) ¿Cuál es el nivel de confianza con el que se ha construido el intervalo?

c) Determina el tamaño muestral mínimo para que con un 99 % de confianza se estime el valor de la media con un error menor que 0,2 g.

La variable X: “Peso, en gramos de las latas de sardinas” tiene una distribución normal $N(\mu,\sigma=1)$.

a) El peso medio de la muestra de 30 latas es el punto medio del intervalo, es decir:

$$\bar{x}=\frac{79{,}81+80{,}49}{2}=80{,}15g$$

b) El nivel de confianza se obtiene igualando el máximo error a la semilongitud del intervalo

$$E=z_{\frac{\alpha}{2}}\frac{1}{\sqrt{30}}=\frac{80{,}49-79{,}81}{2} \Rightarrow z_{\frac{\alpha}{2}}=0{,}34\sqrt{30}=1{,}86 \Rightarrow 1-\frac{\alpha}{2}=0{,}9686 \Rightarrow 1-\alpha=0{,}9371$$

Por lo que el nivel de confianza es del 93,71 %.

c) Si la confianza es del 99 %, entonces: $1-\alpha=0{,}99 \Rightarrow \frac{\alpha}{2}=0{,}005 \Rightarrow z_{0{,}005}=2{,}575$

Como el error de la estimación E debe ser menor de 0,2 gramos, tenemos que:

$$E=2{,}575\frac{1}{\sqrt{n}}<0{,}2 \Rightarrow n>\left(\frac{2{,}575}{0{,}2}\right)^2=165{,}8$$

El tamaño muestral mínimo, para cumplir las condiciones requeridas, es de 166 latas de sardinas.

72. Para estimar la proporción de habitantes que es favorable a la construcción de un centro comercial en un municipio, se ha obtenido el intervalo de confianza (0,31; 0,39), al 94 %.

a) ¿Cuál ha sido el valor de la proporción muestral?

b) Si la muestra aleatoria elegida de esa población para el estudio fue de 500 personas, ¿cuántas de ellas deseaban la construcción del centro comercial?

c) Se desea repetir el estudio para obtener un intervalo de confianza con un error máximo de 0,03 y el mismo nivel de confianza. ¿Cuántas personas, como mínimo, debe tener la nueva muestra aleatoria?

La variable aleatoria *X*: "el habitante seleccionado es favorable a la construcción de un centro comercial" es una variable de Bernoulli $Ber(p)$, siendo p la probabilidad de que un habitante elegido al azar sea favorable a la construcción del centro comercial.

a) El valor de la proporción muestral, que es una estimación puntual de p, es el punto medio del intervalo de confianza dado. Es decir:

$$\hat{p} = \frac{0{,}31 + 0{,}39}{2} = 0{,}35$$

b) De las n = 500 personas de la muestra aleatoria, deseaban la construcción del centro comercial $500 \cdot 0{,}35 = 175$ personas.

c) Si el nivel de confianza es del 94 %, se tiene que: $1 - \alpha = 0{,}94 \Rightarrow \frac{\alpha}{2} = 0{,}03 \Rightarrow z_{0{,}03} = 1{,}88$

Como el máximo error *E* que se permite es 0,03, entonces:

$$E = 1{,}88\sqrt{\frac{0{,}35(1-0{,}35)}{n}} \le 0{,}03 \Rightarrow n \ge \left(\frac{1{,}88}{0{,}03}\right)^2 0{,}35(1-0{,}35) = 893{,}42$$

Por lo que, para que se cumplan las condiciones impuestas, deben formar parte de la muestra, al menos, 894 personas.

73. En un tramo peligroso de una carretera, se sabe que la velocidad a la que circulan los vehículos sigue una distribución normal de media desconocida y desviación típica σ = 15 km/h. Se tomó una muestra aleatoria de 400 vehículos que circulaban por dicho punto peligroso, y se comprobó que la velocidad media de los vehículos de dicha muestra era de 110 km/h.

a) Halla un intervalo de confianza para la media poblacional de la velocidad de circulación en el tramo peligroso, con un nivel de confianza del 95 %.

b) Explica razonadamente el efecto que tendría sobre el intervalo de confianza el aumento o la disminución del nivel de confianza.

La variable aleatoria *X*: "velocidad, en km/h, a la que circulan los vehículos" tiene una distribución normal $N(\mu, \sigma = 15)$, con μ desconocida.

Una muestra aleatoria de $n = 400$ vehículos proporciona una estimación puntual $\bar{x} = 110$ km/h de la media μ.

a) El nivel de confianza del 95 %, implica que: $1 - \alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$; con lo que un intervalo de confianza al 95 % para la media μ es:

$$IC_{0{,}95}(\mu) = \left(110 - 1{,}96\frac{15}{\sqrt{400}}\,;\, 110 + 1{,}96\frac{15}{\sqrt{400}}\right) = (108{,}53;\, 111{,}47)$$

b) Si el nivel de confianza aumenta (disminuye) la amplitud del intervalo de confianza aumenta (disminuye) con lo que la precisión en la estimación disminuye (aumenta). El motivo es que la abscisa z_α, para la que $P(Z < z_\alpha) = 1 - \alpha$, de la distribución normal es mayor (menor) si se aumenta (disminuye) el nivel de confianza. En efecto, si $100(1-\alpha_1)\% < 100(1-\alpha_2)\%$ entonces:

$$1 - \alpha_1 < 1 - \alpha_2 \Rightarrow \frac{\alpha_1}{2} > \frac{\alpha_2}{2} \Rightarrow z_{\frac{\alpha_1}{2}} < z_{\frac{\alpha_2}{2}} \Rightarrow z_{\frac{\alpha_1}{2}}\frac{\sigma}{\sqrt{n}} < z_{\frac{\alpha_2}{2}}\frac{\sigma}{\sqrt{n}}$$

Y, por tanto, la longitud del intervalo correspondiente a la confianza $100(1-\alpha_1)\%$ es menor que la longitud del intervalo correspondiente a la confianza $100(1-\alpha_2)\%$.

74. Un nuevo operador telefónico quiere lanzar en la ciudad una nueva línea de ADSL. Realiza una encuesta entre 520 familias de la ciudad, de las cuales 150 contestan que se cambiarían al nuevo operador.

a) ¿En qué intervalo se encuentra la proporción de familias que cambiaría de operador, con una confianza del 97 %?

b) Haciendo uso de la información muestral inicial, ¿qué tamaño muestral sería necesario para estimar la proporción de familias que se cambiarían de operador, con un error menor del 2 % y una confianza del 95 %?

La variable aleatoria *X*: "la familia seleccionada se cambiaría al nuevo operador" es una variable aleatoria de Bernoulli Ber (p) con p la probabilidad (desconocida) de que una familia elegida al azar cambie de operador.

Una estimación puntual de p, obtenida mediante una encuesta a 520 personas, es:

$$\hat{p}=\frac{150}{520}=0,288$$

a) Si la confianza del intervalo es del 97 %, entonces: $1-\alpha=0,97 \Rightarrow \frac{\alpha}{2}=0,015 \Rightarrow z_{0,015}=2,17$

Y un intervalo de confianza para la proporción p viene dado por:

$$IC_{0,97}(p)=\left(0,288-2,17\sqrt{\frac{0,288(1-0,288)}{520}};\ 0,288+2,17\sqrt{\frac{0,288(1-0,288)}{520}}\right)=(0,245;\ 0,331)$$

b) Con una confianza del 95%, resulta: $1-\alpha=0,95 \Rightarrow \frac{\alpha}{2}=0,025 \Rightarrow z_{0,025}=1,96$

De modo que para que el error cometido E en la estimación por intervalo sea menor de 0,02, debe ser:

$$E=1,96\sqrt{\frac{0,288(1-0,288)}{n}}<0,02 \Rightarrow n>\left(\frac{1,96}{0,02}\right)^2 0,288(1-0,288)=1969,35$$

Por lo que, para estimar la proporción de familias que cambiaría de operador, con las condiciones propuestas, debería elegirse una muestra de, al menos, 1970 familias.

75. En un país en vías de desarrollo, se quiere estimar la proporción de mujeres en su población. El país se compone de cuatro regiones *A*, *B*, *C* y *D* con 1 millón, 2 millones, 2,5 millones y 7 millones de habitantes respectivamente. Se selecciona una muestra aleatoria estratificada del 1 % de la población, con afijación proporcional.

a) ¿Cuántos habitantes de cada una de las regiones hay en la muestra?

b) Si en la muestra de la región *A* hay 5100 mujeres, ¿cuál es la estimación de la proporción de mujeres en esa región?

c) Obtén el intervalo de confianza al 90 % para la estimación de la proporción de mujeres en la región *A*.

Se considera la variable *X*: "la persona elegida es mujer", que es una variable aleatoria de Bernoulli $\text{Ber}(p)$, con p probabilidad de que una persona elegida al azar sea mujer.

El total de la población es de 12,5 millones de habitantes repartidos en cuatro regiones.

a) En la tabla se puede ver el número de habitantes de cada región que se incluyen en la muestra, junto con el total:

	A	B	C	D	Total
Población	1 000 000	2 000 000	2 500 000	7 000 000	12 500 000
Muestra	10 000	20 000	25 000	70 000	125 000

b) El tamaño de la muestra en la región *A* es $n_A=10\,000$. Si en la muestra hay 5100 mujeres, una estimación puntual de la proporción de mujeres en la región A es:

$$\hat{p}_A=\frac{5100}{10\,000}=0,51$$

c) Con un nivel de confianza de 90%, se tiene que: $1-\alpha=0,90 \Rightarrow \frac{\alpha}{2}=0,05 \Rightarrow z_{0,025}=1,645$

De manera que un intervalo de confianza al 90% para la proporción de mujeres en la región *A*, p_A, es:

$$IC_{0,90}(p_A)=\left(0,51-1,645\sqrt{\frac{0,51(1-0,51)}{10000}};\ 0,51+1,645\sqrt{\frac{0,51(1-0,51)}{10000}}\right)=(0,5018;0,5182)$$

76. Se ha tomado una muestra aleatoria de 80 conejos en un criadero industrial. Se ha encontrado que 21 de ellos presentaban una enfermedad que, probablemente, adquirían a través del pienso con que se les alimentaba. Se sabe que la población de conejos en el criadero es de 12 000 unidades.

a) Determina, con una confianza del 92 %, entre qué valores se encuentra el número de conejos enfermos.

b) Haciendo uso de la información muestral inicial, ¿qué número de conejos será necesario estudiar para estimar la proporción de conejos enfermos con un error menor del 7 % y con una confianza del 92 %?

Sea la variable aleatoria *X*: "el conejo elegido presenta la enfermedad". La variable *X* tiene distribución de Bernoulli $\text{Ber}(p)$, con p, la proporción de conejos enfermos, desconocida

Mediante una muestra aleatoria de $n = 80$ conejos se obtiene una estimación puntual de p:

$$\hat{p} = \frac{21}{80} = 0{,}2625$$

a) Si la confianza es del 92 %, entonces: $1-\alpha = 0{,}92 \Rightarrow \frac{\alpha}{2} = 0{,}04 \Rightarrow z_{0{,}04} = 1{,}75$

De manera que un intervalo de confianza al 92% para la proporción p viene dado por:

$$IC_{0{,}92}(p) = \left(0{,}2625 - 1{,}75\sqrt{\frac{0{,}2625(1-0{,}2625)}{80}}\,;\ 0{,}2625 + 1{,}75\sqrt{\frac{0{,}2625(1-0{,}2625)}{80}}\right) = (0{,}1764;\ 0{,}3486)$$

Esto quiere decir que, con una confianza del 92 %, entre el 17,64 % y el 34,86 % de los conejos del criadero presentan la enfermedad. Aplicando estos porcentajes al total de 12 000 conejos del criadero, se obtiene que el número de conejos enfermos está comprendido, al 92 % de confianza, en el siguiente intervalo:

$$(0{,}1764 \cdot 12\,000;\ 0{,}3486 \cdot 12\,000) = (2116{,}8;\ 4183{,}2)$$

b) Si la confianza debe ser ahora del 92 %, entonces: $1-\alpha = 0{,}92 \Rightarrow \frac{\alpha}{2} = 0{,}04 \Rightarrow z_{0{,}04} = 1{,}75$

Como se pide un error *E* inferior al 7 %:

$$E = 1{,}75\sqrt{\frac{0{,}2625(1-0{,}2625)}{n}} < 0{,}07 \Rightarrow n > \left(\frac{1{,}75}{0{,}07}\right)^2 0{,}2625(1-0{,}2625) = 121$$

De modo que para que se verifiquen las condiciones propuestas la muestra debe estar compuesta de, al menos, 121 conejos.

77. El saldo en cuenta a fin de año de los clientes de una cierta entidad bancaria se puede aproximar por una variable aleatoria con distribución normal de desviación típica igual a 400 €. Con el fin de estimar la media del saldo en cuenta a fin de año para los clientes de dicha entidad, se elige una muestra aleatoria simple de 100 clientes.

a) ¿Cuál es el nivel máximo de confianza de la estimación si se sabe que el valor absoluto de la diferencia entre la media muestral y la media poblacional es menor o igual que 66 €?

b) Calcula el tamaño mínimo necesario de la muestra que ha de observarse para que el valor absoluto de la diferencia entre la media muestral y la media poblacional sea menor o igual que 40 €, con un nivel de confianza del 95 %.

La variable aleatoria X: "saldo en cuenta, en euros, a fin de año de los clientes" tiene una distribución de probabilidad que se puede aproximar por la de una distribución normal $N(\mu, \sigma = 400)$, con la media μ desconocida.

Se elige una muestra aleatoria de $n = 100$ clientes. La distribución de la media muestral es:

$$\overline{X} \sim N\left(\mu_{\overline{X}} = \mu, \sigma_{\overline{X}}^2 = \frac{400^2}{100} = 1600\right)$$

a) Se sabe que el error E (diferencia en valor absoluto entre la media muestral y la poblacional) es menor o igual que 66 €. Por tanto:

$$E = z_{\frac{\alpha}{2}} \frac{400}{\sqrt{100}} \le 66 \Rightarrow z_{\frac{\alpha}{2}} \le \frac{66}{40} = 1{,}65 \Rightarrow 1 - \frac{\alpha}{2} = 0{,}9505 \Rightarrow 1 - \alpha = 0{,}9011$$

El nivel de confianza es, por tanto del 90,11 %

b) Si el nivel de confianza es del 95 %, entonces: $1 - \alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$

Como el error E de la estimación por intervalo debe ser menor o igual de 40 €, se tiene que:

$$E = 1{,}96 \frac{400}{\sqrt{n}} \le 40 \Rightarrow n \ge \left(\frac{1{,}96 \cdot 400}{40}\right)^2 = 384{,}16$$

Es decir, para que error cometido en la estimación sea menor o igual que 40 € con un 95 % de confianza se deben tomar una muestra de, al menos, 385 clientes.

AUTOEVALUACIÓN

Comprueba qué has aprendido

1. ***El peso de las barras de pan que se producen en un horno sigue una distribución normal con desviación típica 33 gramos. Una muestra aleatoria de 100 barras proporciona un peso medio de 252 gramos:**

   a) Calcula un intervalo de confianza al 96 % para el peso medio de las barras.

   b) Si se quiere estimar el precio medio con un error menor de 30 gramos, ¿cuál debería ser el tamaño muestral?

   c) Si se quiere aumentar la confianza al 98 %, ¿cuánto hay que aumentar el tamaño muestral para que el error se mantenga?

La variable X: "peso, en gramos, de las barras de pan" tiene una distribución normal $N(\mu, \sigma = 33)$, la media μ desconocida.

Con una muestra aleatoria de $n = 100$ barras se obtiene un peso medio de $\overline{x} = 252$ g.

a) Si la confianza del intervalo es del 96 %, entonces: $1 - \alpha = 0{,}96 \Rightarrow \frac{\alpha}{2} = 0{,}02 \Rightarrow z_{0,02} = 2{,}054$

De modo que un intervalo de confianza al 96 % para la media μ viene dado por:

$$IC_{0,96}(\mu) = \left(252 - 2{,}054\,\frac{33}{\sqrt{100}}\ ;\ 252 + 2{,}054\,\frac{33}{\sqrt{100}}\right) = (245{,}222\ ;\ 258{,}778)$$

b) Si se quiere que el error cometido E, con una confianza del 96 %, al estimar el precio medio μ sea menor de 30 gramos:

$$E = 2{,}054\,\frac{33}{\sqrt{n}} < 30 \Rightarrow n > \left(\frac{2{,}054 \cdot 33}{30}\right)^2 = 5{,}105$$

Se precisa una muestra de, al menos, 6 barras de pan para estimar el peso medio con las condiciones propuestas.

c) Si la confianza del nuevo intervalo es del 98 %, entonces: $1 - \alpha = 0{,}98 \Rightarrow \frac{\alpha}{2} = 0{,}01 \Rightarrow z_{0,01} = 2{,}33$

Como se mantiene el error $E = 5{,}105$; entonces:

$$E = z_{0,01}\,\frac{\sigma}{\sqrt{n}} = 30 \Rightarrow 2{,}33\,\frac{33}{\sqrt{n}} = 30 \Rightarrow n > \left(\frac{2{,}33 \cdot 33}{30}\right)^2 = 6{,}56$$

La nueva muestra deberá tener en total, al menos, 7 barras; es decir 7 – 6 = 1 barras más.

2. En una encuesta se afirma que, con una confianza del 95 %, la proporción de fumadores está comprendida entre el 28,2 % y el 34,4 %. Determina:

a) La proporción estimada con la muestra.

b) El error cometido en la estimación.

c) El tamaño de la muestra utilizada.

a) La proporción estimada con la muestra es el punto medio del intervalo de confianza dado. Es decir:

$$\hat{p} = \frac{0,282 + 0,344}{2} = 0,313$$

Lo que representa el 31,3 %.

b) El error cometido *E* en la estimación es la semiamplitud del intervalo:

$$E = \frac{0,344 - 0,282}{2} = 0,031$$

c) Si la confianza es del 95%, entonces: $1 - \alpha = 0,95 \Rightarrow \frac{\alpha}{2} = 0,025 \Rightarrow z_{0,025} = 1,96$

Igualando el error *E*, calculado en el apartado anterior, con el que se comete aproximando la distribución de la proporción muestral por una distribución normal, se obtiene:

$$0,031 = 1,96\sqrt{\frac{0,313(1-0,313)}{n}} \Rightarrow n = \left(\frac{1,96}{0,031}\right)^2 0,313(1-0,313) = 859,58$$

Se ha utilizado, pues, una muestra aleatoria de 860 personas.

3. Se supone que la vida útil de un dispositivo electrónico que fabrica una determinada empresa sigue una distribución normal de media desconocida y desviación típica 60 horas. Para estimar la vida media se quiere utilizar una muestra de tamaño *n*. Calcula el valor mínimo de *n* tal que, con un nivel de confianza del 99 %, el error de la estimación sea menor que 10 horas.

La variable aleatoria X: "tiempo, en horas, de duración del dispositivo electrónico" tiene una distribución de probabilidad $N(\mu, \sigma = 60)$ con la media μ desconocida.

Con un nivel de confianza del 99 %, se tiene que:

$$1 - \alpha = 0,99 \Rightarrow \frac{\alpha}{2} = 0,005 \Rightarrow z_{0,005} = 2,575$$

De modo que, si se quiere que el error *E*, cometido en la estimación de μ, sea menor que 10 horas:

$$E = 2,575\,\frac{60}{\sqrt{n}} < 10 \Rightarrow n > \left(\frac{2,575 \cdot 60}{10}\right)^2 = 238,7$$

El tamaño mínimo de la muestra debe ser 239 bombillas.

4. **El nivel de colesterol en sangre se supone que sigue una distribución normal. En una muestra de 300 personas de una población *A*, la media muestral del colesterol en sangre fue de 189 mmg/dL, con una desviación típica de 40 mmg/dL. En otra población *B*, el análisis de sangre de 180 personas, proporcionó una media muestral de 168 mmg/dL con una desviación típica de 50 mmg/dL. Halla un intervalo de confianza al 90 % para la diferencia de las medias del colesterol entre ambas poblaciones.**

Se considera la variable *X*: "nivel de colesterol en sangre, en mmg/dL". La distribución probabilidad de *X* es normal.

Tanto la media como la desviación típica de *X* se estiman en cada una de las poblaciones *A* y *B* con una muestra suficientemente grande.

$n_A = 300 \qquad \bar{x}_A = 189 \qquad \sigma_A = 40$

$n_B = 180 \qquad \bar{x}_B = 168 \qquad \sigma_B = 50$

Una confianza del 90%, implica que: $1-\alpha = 0{,}9 \Rightarrow \frac{\alpha}{2} = 0{,}05 \Rightarrow z_{0{,}05} = 1{,}645$

Por tanto, un intervalo de confianza al 90 %, para la diferencia de medias de *A* menos *B*, viene dado por:

$$IC_{0{,}9}(\mu_A - \mu_B) = \left(189 - 168 - 1{,}645\sqrt{\frac{40^2}{300}+\frac{50^2}{180}};\ 189 - 168 + 1{,}645\sqrt{\frac{40^2}{300}+\frac{50^2}{180}}\right) = (13{,}79;\ 28{,}21)$$

Relaciona y contesta

Elige la única respuesta correcta en cada caso

1. **Como resultado de una encuesta se asegura que, con un nivel de confianza del 95 %, el porcentaje de ciudadanos que quiere un carril bici está entre el 59 % y el 62 %. El tamaño mínimo de la muestra ha sido:**

   A. 4004 personas

   B. 4080 personas

   C. 4086 personas

   D. No se puede saber con esos datos.

   La proporción estimada con la muestra es el punto medio del intervalo de confianza dado, pues:

$$\hat{p} = \frac{0,59 + 0,62}{2} = 0,605$$

   Si la confianza es del 95 %, entonces: $1 - \alpha = 0,95 \Rightarrow \frac{\alpha}{2} = 0,025 \Rightarrow z_{0,025} = 1,96$

   Como el error cometido E en la estimación es la semiamplitud del intervalo:

$$E = 1,96\sqrt{\frac{0,605(1-0,605)}{n}} = \frac{0,62 - 0,59}{2} = 0,015 \Rightarrow n = \left(\frac{1,96}{0,015}\right)^2 0,605(1-0,605) = 4080,2$$

   La respuesta, por tanto, es la B.

2. **Una encuesta realizada a 100 hogares, se obtiene como resultado que cada familia tira al año una media de 40 kg de plástico al contenedor de basura orgánica. Si la cuasivarianza muestral es 25, el límite superior del intervalo de confianza al 90 % para la estimación de la media poblacional es:**

   A. 50,00 kg

   B. 42,19 kg

   C. 40,82 kg

   D. 41,29 kg

   Utilizamos la cuasivarianza muestral $s^2 = 25$, como estimador de la varianza poblacional σ^2.

   La muestra aleatoria es de $n = 100$ hogares.

   Si la confianza es del 90%, entonces: $1 - \alpha = 0,90 \Rightarrow \frac{\alpha}{2} = 0,05 \Rightarrow z_{0,05} = 1,645$

   El límite superior del intervalo de confianza se calcula de la siguiente manera:

$$\overline{x} + z_{0,05}\frac{s}{\sqrt{n}} = 40 + 1,645\frac{5}{10} = 40,82$$

   La respuesta, por tanto, es la C.

3. **Para estimar el precio medio de los libros de bolsillo, se elige una muestra aleatoria de 64 libros. El error que se comete en la estimación por intervalo al 99 % de confianza es:**

   A. 0,64375

   B. 0,64735

   C. 0,64537

   D. No se puede calcular.

   La solución es la D, pues el error E se calcula mediante la fórmula $E = z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$, y tenemos los datos $z_{\frac{\alpha}{2}}$ (que se obtiene de la confianza del 99 %) y $n = 64$, pero falta la desviación típica σ.

Señala, en cada caso, las respuestas correctas

4. En la ficha técnica de una encuesta se ha borrado el número de entrevistas realizadas y no se dispone del nivel de confianza, pero se sabe que el error de estimación es ±3,2 % bajo supuesto de máxima indeterminación ($p = 0{,}5$). El nivel de significación y el tamaño de la muestra pueden ser:

A. 95 %; $n = 1661$

B. 95 %; $n = 938$

C. 99 %; $n = 938$

D. 99 %; $n = 1661$

Para el nivel de significación del 95 %: $1-\alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}025 \Rightarrow z_{0{,}025} = 1{,}96$

Para el nivel de significación del 99 %: $1-\alpha = 0{,}99 \Rightarrow \frac{\alpha}{2} = 0{,}005 \Rightarrow z_{0{,}005} = 2{,}575$

El error E se calcula mediante la fórmula: $E = z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$

Para la respuesta A: $E = z_{0{,}025}\sqrt{\frac{p(1-p)}{n}} = 1{,}96\sqrt{\frac{0{,}5(1-0{,}5)}{1661}} = 0{,}024$

Para la respuesta B: $E = z_{0{,}025}\sqrt{\frac{p(1-p)}{n}} = 1{,}96\sqrt{\frac{0{,}5(1-0{,}5)}{938}} = 0{,}032$

Para la respuesta C: $E = z_{0{,}005}\sqrt{\frac{p(1-p)}{n}} = 2{,}575\sqrt{\frac{0{,}5(1-0{,}5)}{938}} = 0{,}042$

Para la respuesta D: $E = z_{0{,}005}\sqrt{\frac{p(1-p)}{n}} = 2{,}575\sqrt{\frac{0{,}5(1-0{,}5)}{1619}} = 0{,}032$

La respuestas correctas son, por tanto, la B y la D.

Señala el dato innecesario para contestar

5. Para estimar un intervalo de confianza para la diferencia de medias de dos poblaciones independientes se necesita:

A. El tamaño de la muestra en cada población.

B. La media muestral en cada población.

C. El nivel de confianza.

D. La longitud del intervalo

Un intervalo de confianza para la diferencia de las medias se calcula mediante la fórmula:

$$\left(\overline{X}-\overline{Y}-z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_X{}^2}{n}+\frac{\sigma_Y{}^2}{m}}\ ,\ \overline{X}-\overline{Y}+z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_X{}^2}{n}+\frac{\sigma_Y{}^2}{m}}\right) = \left(\overline{X}-\overline{Y}\pm z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_X{}^2}{n}+\frac{\sigma_Y{}^2}{m}}\right).$$

Donde:

- $\overline{X}$ e $\overline{Y}$ son las medias muestrales de cada población.
- $\sigma_X{}^2$ y $\sigma_Y{}^2$ son las varianzas de las poblaciones.
- n y m son los tamaños de las muestras de cada población.
- $z_{\frac{\alpha}{2}}$ se obtiene a partir del nivel de confianza.

A simple vista parece que se necesitan casi todos los datos, pero hay que observar que la longitud del intervalo es el doble del error de estimación $E = z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_X{}^2}{n}+\frac{\sigma_Y{}^2}{m}}$. Es decir, que bastaría con la longitud del intervalo (dato D) y con las medias muestrales $\overline{X}$ e $\overline{Y}$ (dato B) para poder calcular el intervalo de confianza.

La solución, pues, es que, en ese caso no serían necesarios los datos A, C.