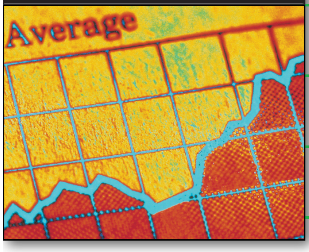
UNIDAD 11 INFERENCIA ESTADÍSTICA



Página 291

Problema 1

- **A partir de una muestra de 500 individuos hemos estimado, con un nivel de confianza del 90%, que la estatura media de los soldados de un cierto reemplazo está comprendida entre 174,3 cm y 175,1 cm.**

  Utilizando el sentido común, y sin realizar ningún tipo de cálculo, responde a las siguientes preguntas:

  a) Imagina que disminuimos el tamaño de la muestra pero queremos que el nivel de confianza se mantenga.

  ¿Cómo influirá este cambio en la longitud del intervalo? ¿Aumentará? ¿Quedará igual? ¿Disminuirá?

  b) Ahora aumentamos el tamaño de la muestra pero queremos que se mantenga la longitud del intervalo.

  ¿Cómo influirá este cambio en el nivel de confianza? ¿Aumentará? ¿Quedará igual? ¿Disminuirá?

  c) Manteniendo el tamaño de la muestra, disminuimos la longitud del intervalo.

  ¿Cómo influirá este cambio en el nivel de confianza? ¿Aumentará? ¿Quedará igual? ¿Disminuirá?

a) Aumentará la longitud del intervalo.

b) Aumentará el nivel de confianza.

c) Disminuirá el nivel de confianza.

Problema 2

- **Reflexionemos sobre cada una de las siguientes experiencias:**

  a) Lanzamos una moneda 10 veces y obtenemos 6 caras.

  b) Lanzamos una moneda 100 veces y obtenemos 60 caras.

  c) Lanzamos una moneda 1000 veces y obtenemos 600 caras.

¿Podemos deducir de alguna de ellas que la moneda es incorrecta? ¿Con cuál de ellas llegamos a esa conclusión con más seguridad? (Responde intuitivamente).

De los apartados b) y c) podemos deducir que la moneda es incorrecta. Con el apartado a) llegamos a esa conclusión con más seguridad.

Página 294

1. **De una variable estadística conocemos la desviación típica, $\sigma = 8$, pero desconocemos la media, μ. Para estimarla, extraemos una muestra de tamaño $n = 60$ cuya media obtenemos: $\bar{x} = 37$. Estima μ mediante un intervalo de confianza del 99%.**

Para un nivel de confianza del 99% tenemos que $z_{\alpha/2} = 2{,}575$.

El intervalo de confianza para μ será:

$$\left(37 - 2{,}575 \cdot \frac{8}{\sqrt{60}};\ 37 + 2{,}575 \cdot \frac{8}{\sqrt{60}}\right); \text{ es decir, } (34{,}34;\ 39{,}66)$$

Por tanto, tenemos una confianza del 99% de que μ esté comprendida entre 34,34 y 39,66.

Página 295

2. **La desviación típica de las estaturas de los soldados es de 5,3 cm.**

¿Qué tamaño ha de tener la muestra para estimar la estatura media, μ, de la población con un error menor de 0,5 cm y con un nivel de confianza del 95%?

Para un nivel de confianza del 95% $(\alpha = 0{,}05)$, tenemos que $z_{\alpha/2} = 1{,}96$. El error máximo admisible es:

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Queremos que $E < 0{,}5$ cm. Despejamos n:

$$1{,}96 \cdot \frac{5{,}3}{\sqrt{n}} < 0{,}5 \quad \rightarrow \quad \sqrt{n} > \frac{1{,}96 \cdot 5{,}3}{0{,}5} = 20{,}776 \quad \rightarrow \quad n > 431{,}64$$

La muestra ha de ser de, al menos, 432 soldados.

Página 296

3. **Sabemos que la desviación típica de los pesos de los pollos adultos es 300 g. Queremos estimar el peso medio de los pollos adultos de una granja con un error menor que 100 g y para ello tomamos una muestra de 50 individuos. ¿Con qué nivel de confianza podremos realizar la estimación?**

Despejamos $z_{\alpha/2}$ en la fórmula del error:

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \rightarrow 100 = z_{\alpha/2} \cdot \frac{300}{\sqrt{50}} \rightarrow z_{\alpha/2} = \frac{100 \cdot \sqrt{50}}{300} \rightarrow z_{\alpha/2} = 2{,}36$$

Hallamos el nivel de confianza:

$$P[z < z_{\alpha/2}] = P[z < 2{,}36] = 0{,}9909$$

$$\frac{\alpha}{2} = P[z \geq 2{,}36] = 1 - 0{,}9909 = 0{,}0091$$

$$\alpha = 2 \cdot 0{,}091 = 0{,}0182 \rightarrow 1 - \alpha = 0{,}9818$$

El nivel de confianza es del 98,18%.

Página 298

1. Se ha lanzado un dado 400 veces y se ha obtenido 72 veces el valor 4. Estimar el valor de la probabilidad $P[4]$ con un nivel de confianza del 90%.

Para un nivel de confianza del 90%, tenemos que $z_{\alpha/2} = 1{,}645$. La proporción de cuatros obtenida en la muestra es:

$$pr = \frac{72}{400} = 0{,}18$$

El intervalo de confianza para estimar $P[4]$ será:

$$\left(0{,}18 - 1{,}645 \cdot \sqrt{\frac{0{,}18 \cdot 0{,}82}{400}};\ 0{,}18 + 1{,}645 \cdot \sqrt{\frac{0{,}18 \cdot 0{,}82}{400}}\right), \text{ es decir:}$$

$$(0{,}148;\ 0{,}212)$$

Es decir, con un nivel de confianza del 90%, la probabilidad de obtener 4 está entre 0,148 y 0,212.

2. ¿Cuántas veces hemos de lanzar un dado, que suponemos levemente incorrecto, para estimar la probabilidad de "6" con un error menor que 0,002 y un nivel de confianza del 95%?

Para un nivel de confianza del 95%, tenemos que $z_{\alpha/2} = 1{,}96$. Como desconocemos el valor de pr, tomaremos $pr = \frac{1}{6} \approx 0{,}17$ (suponemos el dado levemente incorrecto).

El error máximo admisible es:

$$E = z_{\alpha/2} \cdot \sqrt{\frac{pr(1 - pr)}{n}} \rightarrow 0{,}002 = 1{,}96 \cdot \sqrt{\frac{0{,}17 \cdot 0{,}83}{n}} \rightarrow$$

$$\rightarrow n = 135\,512{,}44$$

Deberemos lanzarlo, al menos, 135 513 veces.

Página 301

1. Repite, paso a paso, el CASO 1 para un nivel de significación α = 0,01.

1º Enunciación:

H_0: p = 0,167 H_1: $p \neq$ 0,167

2º Zona de aceptación:

Las proporciones muestrales se distribuirían:

$$N\left(p,\ \sqrt{\frac{pq}{n}}\right) = N\left(0{,}167,\ \sqrt{\frac{0{,}167 \cdot 0{,}833}{100}}\right) = N(0{,}167,\ 0{,}037)$$

Nivel de significación: $\alpha = 0{,}01 \ \to\ z_{\alpha/2} = 2{,}575$

Zona de aceptación: $(0{,}167 \pm 2{,}575 \cdot 0{,}037) = (0{,}072;\ 0{,}262)$

3º Verificación:

Se extrae la muestra y se calcula el valor del parámetro:

$$pr = \frac{25}{100} = 0{,}25$$

4º Decisión:

0,25 sí está en la zona de aceptación. Se acepta la hipótesis nula. Consideramos el dado correcto.

2. Repite, paso a paso, el CASO 2 para un nivel de significación α = 0,10.

1º Enunciación:

H_0: μ = 102 H_1: $\mu \neq$ 102

2º Zona de aceptación:

Las medias muestrales se distribuirían:

$$N\left(102,\ \frac{11}{\sqrt{400}}\right) = N(102;\ 0{,}55)$$

Nivel de significación: $\alpha = 0{,}10 \ \to\ z_{\alpha/2} = 1{,}645$

Zona de aceptación:

$$(102 \pm 1{,}645 \cdot 0{,}55) = (101{,}09;\ 102{,}90)$$

3º Verificación:

Se extrae la muestra y se calcula el valor del parámetro: $\bar{x} = 101$

4º Decisión:

101 no está en la zona de aceptación. Se rechaza la hipótesis nula.

Los conocimientos de los soldados no son los mismos que hace cinco años.

Página 302

1. a) En una población para la cual es $\sigma = 29$, contrasta la hipótesis de que $\mu = 347$, con un nivel de significación del 1%, mediante una muestra de 200 individuos en la que se obtiene $\bar{x} = 352$.

b) Repite el contraste para $\alpha = 10\%$.

a) **1er paso: Hipótesis:**

H_0: $\mu = 347$; H_1: $\mu \neq 347$

2º paso: Zona de aceptación:

Para un nivel de significación del 1%, $\alpha = 0,01$, tenemos que $z_{\alpha/2} = 2,575$. La zona de aceptación sería el intervalo:

$$\left(347 - 2,575 \cdot \frac{29}{\sqrt{200}}; 347 + 2,575 \cdot \frac{29}{\sqrt{200}}\right); \text{ es decir: } (341,72; 352,28)$$

3er paso: Verificación:

La media muestral obtenida es $\bar{x} = 352$.

4º paso: Decisión:

Como 352 está en la zona de aceptación, aceptamos la hipótesis nula. Es decir, aceptamos que $\mu = 347$.

b) **1er paso: Hipótesis:**

H_0: $\mu = 347$; H_1: $\mu \neq 347$

2º paso: Zona de aceptación:

Para un nivel de significación del 10% $(\alpha = 0,10)$, tenemos que $z_{\alpha/2} = 1,645$. La zona de aceptación sería el intervalo:

$$\left(347 - 1,645 \cdot \frac{29}{\sqrt{200}}; 347 + 1,645 \cdot \frac{29}{\sqrt{200}}\right); \text{ es decir: } (343,63; 350,37)$$

3er paso: Verificación:

La media muestral obtenida es $\bar{x} = 352$.

4º paso: Decisión:

Como 352 no está en la zona de aceptación, rechazamos la hipótesis nula, es decir, aceptamos que $\mu \neq 352$.

Página 303

2. En una población para la cual es $\sigma = 29$, contrasta la hipótesis de que $\mu \leq 347$ con un nivel de significación del 1%, mediante una muestra de 200 individuos en la que se obtiene $\bar{x} = 352$.

1er paso: Hipótesis:

H_0: $\mu \le 347$; H_1: $\mu > 347$

2º paso: Zona de aceptación:

Para $\alpha = 0{,}01$, $z_\alpha = 2{,}33$.

La zona de aceptación es el intervalo:

$$\left(-\infty;\ 347 + 2{,}33 \cdot \frac{29}{\sqrt{200}}\right) = (-\infty;\ 351{,}78)$$

3er paso: Verificación:

La media muestral obtenida es $\bar{x} = 352$.

4º paso: Decisión:

Como 352 no está en la zona de aceptación, rechazamos la hipótesis nula, es decir, aceptamos que $\mu > 347$.

Página 305

1. Respecto a un cierto dado, A opina que $P[6] = 0{,}15$, B opina que $P[6] \le 0{,}15$ y C opina que $P[6] \ge 0{,}15$. Contrasta las tres hipótesis con un nivel de significación de 0,10, sabiendo que se arrojó el dado 1 000 veces y se obtuvo 183 veces el "6".

1er paso: Hipótesis:

	PARA A	PARA B	PARA C
HIPÓTESIS NULA	H_0: $p = 0{,}15$	H_0: $p \le 0{,}15$	H_0: $p \ge 0{,}15$
HIPÓTESIS ALTERNATIVA	H_1: $p \ne 0{,}15$	H_1: $p > 0{,}15$	H_1: $p < 0{,}15$

2º paso: Zona de aceptación:

A $\rightarrow$ $\alpha = 0{,}10$ $\rightarrow$ $z_{\alpha/2} = 1{,}645$

Intervalo: $\left(0{,}15 \pm 1{,}645 \cdot \sqrt{\dfrac{0{,}15 \cdot 0{,}85}{1\,000}}\right) = (0{,}131;\ 0{,}169)$

B $\rightarrow$ $\alpha = 0{,}10$ $\rightarrow$ $z_\alpha = 1{,}28$

Intervalo: $\left(-\infty;\ 0{,}15 + 1{,}28 \cdot \sqrt{\dfrac{0{,}15 \cdot 0{,}85}{1\,000}}\right) = (-\infty;\ 0{,}164)$

C $\rightarrow$ $\alpha = 0{,}10$ $\rightarrow$ $z_\alpha = 1{,}28$

Intervalo: $\left(0{,}15 - 1{,}28 \cdot \sqrt{\dfrac{0{,}15 \cdot 0{,}85}{1\,000}};\ +\infty\right) = (0{,}136;\ +\infty)$

$3^{\underline{er}}$ paso: Verificación:

Hemos obtenido una proporción muestral de:

$$pr = \frac{183}{1\,000} = 0{,}183$$

$4^{\underline{o}}$ paso: Decisión:

A $\rightarrow$ Rechazamos H_0 (es decir, aceptamos que $p \neq 0{,}15$).

B $\rightarrow$ Rechazamos H_0 (es decir, aceptamos que $p > 0{,}15$).

C $\rightarrow$ Aceptamos H_0 (es decir, aceptamos que $p \geq 0{,}15$).

Página 313

EJERCICIOS Y PROBLEMAS PROPUESTOS

PARA PRACTICAR

Intervalos de confianza. Error. Tamaño muestral. Nivel de confianza

1 S **Para estimar la estatura media de los jóvenes entre 15 y 25 años de una localidad, se ha medido a 40 de estos jóvenes, obteniéndose los siguientes resultados:**

ESTATURA (cm)	[148, 153)	[153, 158)	[158, 163)
Nº JÓVENES	2	4	11

ESTATURA (cm)	[163, 168)	[168, 173)	[173, 178)
Nº JÓVENES	14	5	4

Estima, con un nivel de confianza del 99%, el valor de la estatura media de los jóvenes entre 15 y 25 años de dicha localidad.

Hallamos $\bar{x}$ y s para la muestra obtenida:

ESTATURA (cm)	[148, 153)	[153, 158)	[158, 163)	[163, 168)	[168, 173)	[173, 178)
MARCA DE CLASE (x_i)	150,5	155,5	160,5	165,5	170,5	175,5
FRECUENCIA (f_i)	2	4	11	14	5	4

$$\bar{x} = 164 \text{ y } s = 6{,}24$$

Para un nivel de confianza del 99%, se tiene:

$$1 - \alpha = 0{,}99 \rightarrow z_{\alpha/2} = 2{,}575$$

Así, el intervalo de confianza para estimar μ al 99% es:

$$\left(164 - 2{,}575 \cdot \frac{6{,}24}{\sqrt{40}};\ 164 + 2{,}575 \cdot \frac{6{,}24}{\sqrt{40}}\right); \text{ es decir: } (161{,}46;\ 166{,}54)$$

2 **En una muestra de 50 jóvenes encontramos que la dedicación media diaria de ocio es de 400 minutos y su desviación típica de 63 minutos. Calcula el intervalo de confianza de la media de la población al 95% de nivel de confianza.**

Para el 95% $\rightarrow$ $1 - \alpha = 0,95$ $\rightarrow$ $z_{\alpha/2} = 1,96$

El intervalo de confianza para μ al 95% es:

$$\left(400 - 1,96 \cdot \frac{63}{\sqrt{50}};\ 400 + 1,96 \cdot \frac{63}{\sqrt{50}}\right); \text{ es decir: } (382,54;\ 417,46)$$

3 **La desviación típica de una variable estadística es $\sigma = 5$. Para estimar la media de dicha variable, extraemos una muestra aleatoria de tamaño $n = 100$ y obtenemos $\bar{x} = 28$. Obtén un intervalo de confianza del 95% para estimar la media de la población, μ.**

Para el 95% $\rightarrow$ $1 - \alpha = 0,95$ $\rightarrow$ $z_{\alpha/2} = 1,96$

El intervalo de confianza para μ al 95% es:

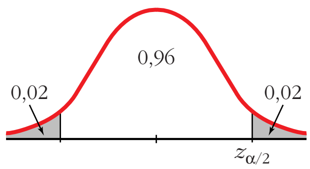
$$\left(28 - 1,96 \cdot \frac{5}{\sqrt{100}};\ 28 + 1,96 \cdot \frac{5}{\sqrt{100}}\right); \text{ es decir: } (27,02;\ 28,98)$$

4 **Los estudiantes de Bachillerato de una cierta comunidad autónoma duermen un número de horas diarias que se distribuye según una ley normal de media μ desconocida y de desviación típica 3. A partir de una muestra de tamaño 30 se ha obtenido una media muestral igual a 7 horas.**

Halla un intervalo de confianza al 96% para la media de horas de sueño, μ.

Hallamos $z_{\alpha/2}$ para el 96% $\rightarrow$ $1 - \alpha = 0,96$:

$P[z \le z_{\alpha/2}] = 0,98$ $\rightarrow$ $z_{\alpha/2} = 2,055$



El intervalo de confianza para μ será:

$$\left(7 - 2,055 \cdot \frac{3}{\sqrt{30}};\ 7 + 2,055 \cdot \frac{3}{\sqrt{30}}\right); \text{ es decir: } (5,87;\ 8,13)$$

5
S **Se desea estudiar el gasto semanal de fotocopias, en céntimos de euro, de los estudiantes de bachillerato de cierta comunidad. Para ello, se ha elegido una muestra aleatoria de 9 de estos estudiantes, resultando los valores siguientes para estos gastos:**

100; 150; 90; 70; 75; 105; 200; 120; 80

Se supone que la variable objeto del estudio sigue una distribución normal de media desconocida y desviación típica igual a 12. Determina un intervalo de confianza al 95% para la media del gasto semanal en fotocopias por estudiante.

Para el 95% $\rightarrow$ $1 - \alpha = 0,95$ $\rightarrow$ $z_{\alpha/2} = 1,96$

Hallamos la media de la muestra:

$$\overline{x} = \frac{100 + 150 + 90 + \ldots + 80}{9} = \frac{990}{9} = 110 \text{ céntimos de euro.}$$

El intervalo de confianza para μ será:

$$\left(110 - 1{,}96 \cdot \frac{12}{\sqrt{9}};\ 110 + 1{,}96 \cdot \frac{12}{\sqrt{9}}\right); \text{ es decir: } (102{,}16;\ 117{,}84)$$

6 **S** **Se ha tomado una muestra aleatoria de 100 individuos a los que se ha medido el nivel de glucosa en sangre, obteniéndose una media muestral de 110 mg/cm^3. Se sabe que la desviación típica de la población es de 20 mg/cm^3.**

a) Obtén un intervalo de confianza, al 90%, para el nivel de glucosa en sangre en la población.

b) ¿Qué error máximo se comete con la estimación anterior?

a) Para el 90% $\rightarrow$ $1 - \alpha = 0{,}90$ $\rightarrow$ $z_{\alpha/2} = 1{,}645$

El intervalo de confianza para μ al 90% es:

$$\left(110 - 1{,}645 \cdot \frac{20}{\sqrt{100}};\ 110 + 1{,}645 \cdot \frac{20}{\sqrt{100}}\right); \text{ es decir: } (106{,}71;\ 113{,}29)$$

b) El error máximo es la mitad de la longitud del intervalo de confianza, es decir:

$$E = 1{,}645 \cdot \frac{20}{\sqrt{100}} = 1{,}645 \cdot 2 = 3{,}29$$

7 **S** **La media de las estaturas de una muestra aleatoria de 400 personas de una ciudad es 1,75 m. Se sabe que la estatura de las personas de esa ciudad es una variable aleatoria que sigue una distribución normal con varianza σ^2 = 0,16 m^2.**

a) Construye un intervalo, de un 95% de confianza, para la media de las estaturas de la población.

b) ¿Cuál sería el mínimo tamaño muestral necesario para que pueda decirse que la verdadera media de las estaturas está a menos de 2 cm de la media muestral, con una confianza del 90%?

a) $n = 400$; $\overline{x} = 1{,}75$ m; $\sigma^2 = 0{,}16$ $\rightarrow$ $\sigma = \sqrt{0{,}16} = 0{,}4$

Para el 95% $\rightarrow$ $1 - \alpha = 0{,}95$ $\rightarrow$ $z_{\alpha/2} = 1{,}96$

El intervalo de confianza es:

$$\left(1{,}75 - 1{,}96 \cdot \frac{0{,}4}{\sqrt{400}};\ 1{,}75 + 1{,}96 \cdot \frac{0{,}4}{\sqrt{400}}\right); \text{ es decir: } (1{,}7108;\ 1{,}7892)$$

b) 90% de confianza $\rightarrow$ $1 - \alpha = 0{,}90$ $\rightarrow$ $z_{\alpha/2} = 1{,}645$

El error máximo admisible es: $E = z_{\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}}$

Buscamos n para que $E < 0{,}02$ m:

$$1{,}645 \cdot \frac{0{,}4}{\sqrt{n}} < 0{,}02 \quad \to \quad \frac{0{,}658}{\sqrt{n}} < 0{,}02 \quad \to \quad \frac{\sqrt{n}}{0{,}658} > \frac{1}{0{,}02}$$

$$\sqrt{n} > \frac{0{,}658}{0{,}02} \quad \to \quad \sqrt{n} > 32{,}9 \quad \to \quad n > 1\,082{,}41$$

Debemos tomar una muestra de, al menos, 1 083 personas.

8 **S** **Las medidas de los diámetros de una muestra al azar de 200 cojinetes de bolas, hechos por una determinada máquina dieron una media de 2 cm y una desviación típica de 0,1 cm. Halla los intervalos de confianza del:**

a) 68,26% **b) 95,44%** **c) 99,73%**

para el diámetro medio de todos los cojinetes.

a) Para el 68,26% $\to$ $1 - \alpha = 0{,}6826$ $\to$ $z_{\alpha/2} = 1$

El intervalo de confianza para μ es:

$$\left(2 - 1 \cdot \frac{0{,}1}{\sqrt{200}};\ 2 + 1 \cdot \frac{0{,}1}{\sqrt{200}}\right); \text{ es decir: } (1{,}993;\ 2{,}007)$$

b) Para el 95,44% $\to$ $1 - \alpha = 0{,}9544$ $\to$ $z_{\alpha/2} = 2$

El intervalo de confianza es:

$$\left(2 - 2 \cdot \frac{0{,}1}{\sqrt{200}};\ 2 + 2 \cdot \frac{0{,}1}{\sqrt{200}}\right); \text{ es decir: } (1{,}986;\ 2{,}014)$$

c) Para el 99,73% $\to$ $1 - \alpha = 0{,}9973$ $\to$ $z_{\alpha/2} = 3$

El intervalo de confianza es:

$$\left(2 - 3 \cdot \frac{0{,}1}{\sqrt{200}};\ 2 + 3 \cdot \frac{0{,}1}{\sqrt{200}}\right); \text{ es decir: } (1{,}979;\ 2{,}021)$$

9 **La duración de las bombillas fabricadas por una empresa sigue una distribución normal de media desconocida y desviación típica 50 horas. Para estimar la duración se experimenta con una muestra de tamaño *n*.**

Calcular el valor de *n* para que, con un nivel de confianza del 95%, se consiga un error en la estimación inferior a 5 horas.

Para el 95% $\to$ $1 - \alpha = 0{,}95$ $\to$ $z_{\alpha/2} = 1{,}96$

El error máximo admisible es:

$E = z_{\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}}$. Buscamos n para que $E < 5$ horas.

Como $z_{\alpha/2} = 1{,}96$ y $\sigma = 50$, queda:

$$1{,}96 \cdot \frac{50}{\sqrt{n}} < 5 \quad \to \quad \frac{98}{\sqrt{n}} < 5 \quad \to \quad \sqrt{n} > \frac{98}{5} = 19{,}6 \quad \to \quad n > 384{,}16$$

Debemos tomar una muestra de, al menos, 385 bombillas.

Página 314

10 S **Al medir el tiempo de reacción, un psicólogo estima que la desviación típica del mismo es de 0,5 segundos. ¿Cuál será el número de medidas que deberá hacer para que sea del 95% la confianza de que el error de su estimación no excederá de 0,05 segundos?**

Para el 95% de confianza, $1 - \alpha = 0{,}95 \rightarrow z_{\alpha/2} = 1{,}96$

El error máximo admisible es:

$E = z_{\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}}$. Buscamos n para que $E \le 0{,}05$:

$$1{,}96 \cdot \frac{0{,}5}{\sqrt{n}} \le 0{,}05 \rightarrow \frac{0{,}98}{\sqrt{n}} \le 0{,}05 \rightarrow \sqrt{n} \ge \frac{0{,}98}{0{,}05} = 19{,}6 \rightarrow n \ge 384{,}16$$

Deberá hacer, al menos, 385 medidas.

11 **Al medir el diámetro de los cojinetes producidos por una empresa, se estima que la desviación típica de dicho diámetro es de 0,05 cm. Se han hecho 121 mediciones.**

¿Se puede afirmar, con el 99% de confianza, que el error en la estimación de la media no excederá a 0,01 cm?

Para el 99% $\rightarrow 1 - \alpha = 0{,}99 \rightarrow z_{\alpha/2} = 2{,}575$

El error máximo admisible es:

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 2{,}575 \cdot \frac{0{,}05}{\sqrt{121}} = 0{,}0117 > 0{,}01 \text{ cm}$$

Por tanto, no podemos afirmar que el error en la estimación no excederá a 0,01 cm.

12 S **Las ventas mensuales de una tienda de electrodomésticos se distribuyen según una ley normal, con desviación típica 900 €. En un estudio estadístico de las ventas realizadas en los últimos nueve meses, se ha encontrado un intervalo de confianza para la media mensual de las ventas, cuyos extremos son 4 663 € y 5 839 €.**

a) ¿Cuál ha sido la media de las ventas en estos nueve meses?

b) ¿Cuál es el nivel de confianza para este intervalo?

a) La media de las ventas es el punto medio del intervalo; es decir:

$$\bar{x} = \frac{4663 + 5839}{2} = 5251 \text{ €}$$

b) El estudio se ha realizado en los últimos 9 meses, es decir, se ha considerado una muestra de tamaño $n = 9$.

El error máximo admisible es la mitad de la longitud del intervalo, es decir:

$$E = \frac{5839 - 4663}{2} = 588$$

Así, sabemos que: $n = 9$; $\sigma = 900$; $E = 588$, y como:

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \to 588 = z_{\alpha/2} \cdot \frac{900}{\sqrt{9}} \to 588 = z_{\alpha/2} \cdot \frac{900}{3} \to$$

$$\to 588 = z_{\alpha/2} \cdot 300 \to z_{\alpha/2} = \frac{588}{300} = 1{,}96 \to$$

$$\to z_{\alpha/2} = 1{,}96$$

que corresponde a un nivel de confianza del 95%.

13 **Se realizó una encuesta a 350 familias preguntando si poseían ordenador en casa, encontrándose que 75 de ellas lo poseían. Estima la proporción real de las familias que disponen de ordenador con un nivel de confianza del 95%.**

La proporción de familias con ordenador en la muestra es $pr = \frac{75}{350} = \frac{3}{14}$

Para el 95% de confianza, $1 - \alpha = 0{,}95 \to z_{\alpha/2} = 1{,}96$

El intervalo de confianza para p es:

$$\left(\frac{3}{14} - 1{,}96 \cdot \sqrt{\frac{(3/14)(1 - 3/14)}{350}}; \frac{3}{14} + 1{,}96 \cdot \sqrt{\frac{(3/14)(1 - 3/14)}{350}}\right);$$ es decir:

$(0{,}17; 0{,}26)$

14 **Tomada al azar una muestra de 500 personas en cierta comunidad autónoma, se encontró que 220 leían algún periódico habitualmente. Calcula, con un nivel de confianza del 95%, el intervalo en el que se encontrará la verdadera proporción de lectores de periódicos y explica el proceso seguido para dicho cálculo.**

La proporción de lectores del periódico en la muestra es $pr = \frac{220}{500} = 0{,}44$.

Para un 95% de confianza, $1 - \alpha = 0{,}95 \to z_{\alpha/2} = 1{,}96$

El intervalo de confianza para p es:

$$\left(0{,}44 - 1{,}96 \cdot \sqrt{\frac{0{,}44 \cdot 0{,}56}{500}}; 0{,}44 + 1{,}96 \cdot \sqrt{\frac{0{,}44 \cdot 0{,}56}{500}}\right);$$ es decir:

$(0{,}396; 0{,}484)$

15 **¿De qué tamaño conviene tomar la muestra de una línea de producción para tener una confianza del 95% de que la proporción estimada no difiere de la verdadera en más de un 4%? Se sabe, por estudios previos, que la proporción de objetos defectuosos es del orden del 0,05.**

Para el 95% de confianza, $1 - \alpha = 0{,}95 \to z_{\alpha/2} = 1{,}96$

El error máximo admisible es:

$E = z_{\alpha/2} \cdot \sqrt{\frac{pr(1 - pr)}{n}}$. Buscamos n para que $E \le 0{,}04$ (no más de un 4%):

$$1{,}96 \cdot \sqrt{\frac{0{,}05(1-0{,}05)}{n}} \le 0{,}04 \quad \rightarrow \quad n \ge 114{,}05$$

El tamaño mínimo de la muestra ha de ser $n = 115$.

16 **S** **Se desea estimar la proporción, *p*, de individuos daltónicos de una población a través del porcentaje observado en una muestra aleatoria de individuos, de tamaño *n*.**

a) Si el porcentaje de individuos daltónicos en la muestra es igual al 30%, calcula el valor de *n* para que, con un nivel de confianza de 0,95, el error cometido en la estimación sea inferior al 3,1%.

b) Si el tamaño de la muestra es de 64 individuos, y el porcentaje de individuos daltónicos en la muestra es del 35%, determina, usando un nivel de significación del 1%, el correspondiente intervalo de confianza para la proporción de daltónicos de la población.

a) Para un nivel de confianza del 95%, $1 - \alpha = 0{,}95 \quad \rightarrow \quad z_{\alpha/2} = 1{,}96$

El error máximo admisible es:

$E = z_{\alpha/2} \cdot \sqrt{\frac{pr(1-pr)}{n}}$. Buscamos n para que $E < 0{,}031$ (inferior al 3,1%):

$$1{,}96 \cdot \sqrt{\frac{0{,}3 \cdot 0{,}7}{n}} < 0{,}031 \quad \rightarrow \quad n > 839{,}48$$

La muestra ha de ser, como mínimo, de 840 individuos.

b) Para un nivel de significación del 1%, tenemos que:

$$\alpha = 0{,}01 \quad \rightarrow \quad 1 - \alpha = 0{,}99 \quad \rightarrow \quad z_{\alpha/2} = 2{,}575$$

El intervalo de confianza para p será:

$$\left(0{,}35 - 2{,}575 \cdot \sqrt{\frac{0{,}35 \cdot 0{,}65}{64}};\ 0{,}35 + 2{,}575 \cdot \sqrt{\frac{0{,}35 \cdot 0{,}65}{64}}\right);\ \text{es decir:}$$

(0,196; 0,504)

Contrastes de hipótesis

17 **S** **Un fabricante de lámparas eléctricas está ensayando un nuevo método de producción que se considerará aceptable si las lámparas obtenidas por este método dan lugar a una población normal de duración media 2 400 horas, con una desviación típica igual a 300.**

Se toma una muestra de 100 lámparas producidas por este método y esta muestra da una duración media de 2 320 horas. ¿Se puede aceptar la hipótesis de validez del nuevo proceso de fabricación con un riesgo igual o menor al 5%?

1$^{\text{er}}$ paso: Hipótesis: Tenemos que contrastar:

$$H_0: \mu = 2\,400 \quad \text{frente a} \quad H_1: \mu \ne 2\,400$$

2º paso: Zona de aceptación:

$$\left(\mu_0 - z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}},\ \mu_0 + z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}}\right)$$

Conocemos los siguientes datos:

$\mu_0 = 2\,400;\ \ \sigma_0 = 300;\ \ n = 100$

$\alpha = 0,05 \ \ \to \ \ z_{\alpha/2} = 1,96$

Por tanto, la zona de aceptación será:

$\left(2\,400 - 1,96 \cdot \frac{300}{\sqrt{100}};\ 2\,400 + 1,96 \cdot \frac{300}{\sqrt{100}}\right)$; es decir, el intervalo:

$(2\,341,2;\ 2\,458,8)$

3^er^ paso: Verificación:

Hemos obtenido una media muestral de $\bar{x} = 2\,320$.

4º paso: Decisión:

Como $\bar{x} = 2\,320$ no cae dentro de la zona de aceptación, rechazamos H_0, es decir, no podemos aceptar la validez del nuevo proceso de fabricación.

18 **Un laboratorio afirma que un calmante quita la jaqueca en 14 minutos en los casos corrientes. Con el fin de comprobar esta información, se eligen al azar 30 pacientes con jaqueca y se toma como variable en el experimento el tiempo que transcurre entre la administración del calmante y el momento en que desaparece la jaqueca.**

Los resultados obtenidos en esta muestra fueron media 17 minutos y desviación típica 7 minutos. ¿Podemos admitir como cierta la afirmación del laboratorio a un nivel de confianza del 95%?

1^er^ paso: Hipótesis: Tenemos que contrastar:

H_0: $\mu = 14$ frente a H_1: $\mu \neq 14$

2º paso: Zona de aceptación:

$$\left(\mu_0 - z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}},\ \mu_0 + z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}}\right)$$

Conocemos los siguientes datos:

$\mu_0 = 14;\ \ \sigma_0 = 7;\ \ n = 30$

$1 - \alpha = 0,95 \ \ \to \ \ z_{\alpha/2} = 1,96$

Por tanto, la zona de aceptación será:

$\left(14 - 1,96 \cdot \frac{7}{\sqrt{30}};\ 14 + 1,96 \cdot \frac{7}{\sqrt{30}}\right)$; es decir, el intervalo $(11,495;\ 16,505)$

3$^{\underline{er}}$ paso: Verificación:

Hemos obtenido una media muestral de $\bar{x} = 17$ minutos.

4º paso: Decisión:

Como $\bar{x} = 17$ no cae dentro de la zona de aceptación, rechazamos H_0, es decir, no podemos aceptar que el calmante quite la jaqueca en 14 minutos.

19 S **Se sabe que la renta anual de los individuos de una localidad sigue una distribución normal de media desconocida y desviación típica 2 400 €. Se ha observado la renta anual de 16 individuos de esa localidad escogidos al azar, y se ha obtenido un valor medio de 16 000 €. Contrasta, a un nivel de significación del 5%, si la media de la distribución es 14 500 €. Para ello, responde:**

a) ¿Cuáles son la hipótesis nula y la alternativa del contraste?

b) Determina la forma de la región crítica.

c) ¿Se acepta la hipótesis nula, con el nivel de confianza indicado?

a) Tenemos que contrastar la hipótesis nula:

$$H_0: \mu = 14\,500$$

frente a la hipótesis alternativa:

$$H_1: \mu \neq 14\,500$$

b) La zona de aceptación es el intervalo:

$$\left(\mu_0 - z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}},\ \mu_0 + z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}}\right)$$

Como $\mu_0 = 14\,500$; $\sigma = 2\,400$; $n = 16$

$\alpha = 0{,}05 \rightarrow z_{\alpha/2} = 1{,}96$; tenemos que la zona de aceptación es el intervalo:

$$\left(14\,500 - 1{,}96 \cdot \frac{2\,400}{\sqrt{16}};\ 14\,500 + 1{,}96 \cdot \frac{2\,400}{\sqrt{16}}\right); \text{ es decir, } (13\,324;\ 15\,676)$$

c) Como hemos obtenido una media muestral de $\bar{x} = 16\,000$ €, que no está en el intervalo de aceptación, rechazamos H_0, es decir, no podemos aceptar que la media sea de 14 500 €.

Página 315

20 S **Se sabe por experiencia que el tiempo obtenido por los participantes olímpicos de la prueba de 100 metros, en la modalidad de decathlon, es una variable aleatoria que sigue una distribución normal con media 12 segundos y desviación típica 1,5 segundos. Para contrastar, con un nivel de significación del 5%, si no ha variado el tiempo medio en la última Olimpiada, se extrajo una muestra aleatoria de 10 participantes y se anotó el tiempo obtenido por cada uno, con los resultados siguientes, en segundos:**

13 12 11 10 11 11 9 10 12 11

a) ¿Cuáles son la hipótesis nula y la alternativa del contraste?

b) Determina la región crítica.

c) Realiza el contraste.

a) Tenemos que contrastar la hipótesis nula:

H_0: $\mu = 12$

frente a la hipótesis alternativa:

H_1: $\mu \neq 12$

b) La zona de aceptación es el intervalo:

$$\left(\mu_0 - z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}},\ \mu_0 + z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}}\right)$$

Como $\mu_0 = 12$; $\sigma = 1,5$; $n = 10$;

$\alpha = 0,05 \rightarrow z_{\alpha/2} = 1,96$; tenemos que la zona de aceptación es el intervalo:

$$\left(12 - 1,96 \cdot \frac{1,5}{\sqrt{10}};\ 12 + 1,96 \cdot \frac{1,5}{\sqrt{10}}\right); \text{ es decir, } (11,07;\ 12,93)$$

c) Calculamos la media de la muestra:

$$\bar{x} = \frac{13 + 12 + 11 + \ldots + 11}{10} = \frac{110}{10} = 11$$

Como no está dentro del intervalo de aceptación, rechazamos H_0, es decir, no podemos aceptar que la media siga siendo la misma.

21 **Sabemos que la vida media de las lavadoras de una determinada marca sigue una distribución normal con una desviación típica de 0,7 años. Un fabricante de dicha marca afirma que sus lavadoras tienen una vida media de 10 años. Para comprobar dicha afirmación se obtuvo una muestra de 50 lavadoras, cuya duración media fue de $\bar{x} = 9,2$ años.**

a) Formula la hipótesis nula y la hipótesis alternativa del contraste.

b) Determina la zona de aceptación y realiza el contraste (con un nivel de confianza del 95%).

c) Explica, en el contexto del problema, en qué consisten cada uno de los errores del tipo I y II.

a) Tenemos que contrastar la hipótesis nula:

H_0: $\mu = 10$

frente a la hipótesis alternativa:

H_1: $\mu \neq 10$

b) La zona de aceptación es el intervalo:

$$\left(\mu_0 - z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}},\ \mu_0 + z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}}\right)$$

Como $\mu_0 = 10$; $\sigma = 0{,}7$; $n = 50$

$\alpha = 0{,}05 \rightarrow z_{\alpha/2} = 1{,}96$; tenemos que la zona de aceptación es el intervalo:

$$\left(10 - 1{,}96 \cdot \frac{0{,}7}{\sqrt{50}};\ 10 + 1{,}96 \cdot \frac{0{,}7}{\sqrt{50}}\right);\ \text{es decir, } (9{,}81;\ 10{,}19)$$

Como hemos obtenido una media muestral de $\bar{x} = 9{,}2$, que no pertenece a la zona de aceptación, rechazamos H_0, es decir, no podemos aceptar que la duración media sea de 10 años.

c) El error de tipo I consiste en rechazar H_0 siendo verdadera. En este caso, consistiría en rechazar que la media es de 10 años, siendo verdad.

El error de tipo II consiste en aceptar H_0 siendo falsa. En este caso, consistiría en aceptar que la media es de 10 años, siendo falso.

22 S **Se ha comprobado que el tiempo de espera (en minutos) hasta ser atendido, en cierto servicio de urgencias, sigue un modelo normal de probabilidad. A partir de una muestra de 100 personas que fueron atendidas en dicho servicio, se ha calculado un tiempo medio de espera de 14,25 minutos y una desviación típica de 2,5 minutos.**

a) ¿Podríamos afirmar, con un nivel de significación del 5% ($\alpha = 0{,}05$), que el tiempo medio de espera, en ese servicio de urgencias, no es de 15 minutos?

b) ¿Qué podríamos concluir si el nivel de significación hubiese sido del 0,1% ($\alpha = 0{,}001$)?

c) ¿Existe contradicción en ambas situaciones?

Justifica las respuestas.

a) **1^er^ paso: Hipótesis:** Tenemos que contrastar:

H_0: $\mu = 15$ frente a H_1: $\mu \neq 15$

2º paso: Zona de aceptación:

$$\left(\mu_0 - z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}},\ \mu_0 + z_{\alpha/2} \cdot \frac{\sigma_0}{\sqrt{n}}\right)$$

Como $\mu_0 = 15$; $\sigma = 2{,}5$; $n = 100$;

$\alpha = 0{,}05 \rightarrow z_{\alpha/2} = 1{,}96$; tenemos que la zona de aceptación es:

$$\left(15 - 1{,}96 \cdot \frac{2{,}5}{\sqrt{100}};\ 15 + 1{,}96 \cdot \frac{2{,}5}{\sqrt{100}}\right);\ \text{es decir, el intervalo } (14{,}51;\ 15{,}49)$$

3^er^ paso: Verificación:

Hemos obtenido una media muestral de $\bar{x} = 14{,}25$.

4º paso: Decisión:

Como la media muestral está fuera de la zona de aceptación, rechazamos H_0, es decir, no podemos aceptar que el tiempo medio sea de 15 minutos.

b) Si $\alpha = 0{,}001$, entonces $z_{\alpha/2} = 3{,}27$, y la zona de aceptación sería:

$$\left(15 - 3{,}27 \cdot \frac{2{,}5}{\sqrt{100}};\ 15 + 3{,}27 \cdot \frac{2{,}5}{\sqrt{100}}\right);$$ es decir, el intervalo (14,18; 15,82)

Por tanto, como $\bar{x} = 14{,}25$ sí está en el intervalo de aceptación, no podríamos rechazar H_0, es decir, aceptaríamos que el tiempo medio es de 15 minutos.

c) No existe contradicción. En el apartado b) el riesgo que estamos asumiendo es muy pequeño, mucho menor que en el caso a), por tanto, el intervalo es más amplio.

23 S **La duración de las bombillas de 100 vatios que fabrica una empresa sigue una distribución normal con una desviación típica de 120 horas. Su vida media está garantizada durante un mínimo de 800 horas.**

Se escoge al azar una muestra de 50 bombillas de un lote y, después de comprobarlas, se obtiene una vida media de 750 horas. Con un nivel de significación de 0,01, ¿habría que rechazar el lote por no cumplir la garantía?

1.er paso: Hipótesis: Queremos contrastar:

H_0: $\mu \geq 800$ frente a H_1: $\mu < 800$

2.º paso: Zona de aceptación:

$$\left(\mu_0 - z_\alpha \cdot \frac{\sigma_0}{\sqrt{n}};\ +\infty\right)$$

Para $\alpha = 0{,}01 \rightarrow z_\alpha = 2{,}33$. Como $\mu_0 = 800$; $\sigma_0 = 120$ y $n = 50$, la zona de aceptación será:

$$\left(800 - 2{,}33 \cdot \frac{120}{\sqrt{50}};\ +\infty\right);$$ es decir, el intervalo $(760{,}46;\ +\infty)$

3.er paso: Verificación:

Hemos obtenido una media muestral de $\bar{x} = 750$ horas.

4.º paso: Decisión:

Como la media muestral no está dentro de la zona de aceptación, rechazamos H_0, es decir, habría que rechazar el lote por no cumplir la garantía.

24 **Una empresa asegura que unas determinadas pastillas de jabón duran más de 11 días. Para comprobarlo, se realiza una encuesta en 100 casos. Estas son las respuestas:**

DURACIÓN (días)	5 a 9	10 a 14	15 a 19	20 a 24
RESPUESTAS	24	46	19	11

¿Se puede dar como válida la afirmación de la empresa, para un nivel de significación de $\alpha = 0{,}05$?

Calculamos la media muestral y la desviación típica:

DURACIÓN (días)	5 a 9	10 a 14	15 a 19	20 a 24
x_i	7	12	17	22
f_i	24	46	19	11

$$\bar{x} = \frac{\Sigma f_i x_i}{n} = \frac{1\,285}{100} = 12{,}85 \text{ días}; \quad s = 4{,}59$$

1$^{\underline{er}}$ paso: Hipótesis: Queremos contrastar:

$$H_0: \mu \le 11 \text{ frente a } H_1: \mu > 11$$

2º paso: Zona de aceptación:

$$\left(-\infty;\ \mu_0 - z_\alpha \cdot \frac{\sigma_0}{\sqrt{n}}\right)$$

Para $\alpha = 0{,}05 \rightarrow z_\alpha = 1{,}645$. Como $\mu_0 = 11$; $\sigma_0 = 4{,}59$ y $n = 100$, la zona de aceptación es:

$$\left(-\infty;\ 11 + 1{,}645 \cdot \frac{4{,}59}{\sqrt{100}}\right); \text{ es decir, el intervalo } (-\infty;\ 11{,}76)$$

3$^{\underline{er}}$ paso: Verificación:

La media muestral obtenida es $\bar{x} = 12{,}85$ días.

4º paso: Decisión:

Como $\bar{x} = 12{,}85$ está fuera de la zona de aceptación, rechazamos H_0, es decir, aceptamos que las pastillas de jabón duran más de 11 días.

25 S **Una encuesta, realizada a 64 empleados de una fábrica, concluyó que el tiempo medio de duración de un empleo en la misma era de 6,5 años, con una desviación típica de 4.**

¿Sirve esta información para aceptar, con un nivel de significación del 5%, que el tiempo medio de empleo en esa fábrica es menor o igual que 6? Justifica adecuadamente la respuesta.

1$^{\underline{er}}$ paso: Hipótesis: Tenemos que contrastar:

$$H_0: \mu \le 6 \text{ frente a } H_1: \mu > 6$$

2º paso: Zona de aceptación:

$$\left(-\infty;\ \mu_0 - z_\alpha \cdot \frac{\sigma_0}{\sqrt{n}}\right)$$

Para un nivel de significación del 5%, tenemos que $z_\alpha = 1{,}645$. Por tanto, el intervalo es:

$$\left(-\infty;\ 6 + 1{,}645 \cdot \frac{4}{\sqrt{64}}\right); \text{ es decir, } (-\infty;\ 6{,}8225)$$

3$^{\text{er}}$ paso: Verificación:

La media muestral obtenida es $\bar{x} = 6,5$ años.

4º paso: Decisión:

Como la media muestral pertenece al intervalo de aceptación, no podemos rechazar H_0, es decir, aceptamos que el tiempo medio es menor o igual que 6 años.

Página 316

26 S **En función de la información disponible, la dirección de un centro de secundaria ha establecido que la media de horas semanales dedicadas por el alumnado de ese centro al estudio es superior a 15, con una desviación típica igual a 1 hora. Durante el presente curso, el Departamento de Matemáticas quiere demostrar que esta media ha disminuido. Para ello, elige una muestra aleatoria de 150 alumnos, obteniendo una media muestral de 12,7 horas.**

a) ¿Puede afirmarse, con un nivel de confianza del 90%, que ha disminuido el tiempo medio dedicado al estudio por los alumnos y alumnas del centro?

b) Responde a la pregunta a) con un nivel de significación del 1%.

a) **1$^{\text{er}}$ paso: Hipótesis:** Queremos contrastar:

H_0: $\mu \geq 15$ frente a H_1: $\mu < 15$

2º paso: Zona de aceptación:

$$\left(\mu_0 - z_\alpha \cdot \frac{\sigma_0}{\sqrt{n}}, +\infty\right)$$

Para el 90% de confianza, $1 - \alpha = 0,90 \rightarrow z_\alpha = 1,28$

Como $\mu_0 = 15$; $\sigma_0 = 1$; $n = 150$, la zona de aceptación es:

$$\left(15 - 1,28 \cdot \frac{1}{\sqrt{150}}; \ +\infty\right); \text{ es decir, el intervalo } (14,895; +\infty)$$

3$^{\text{er}}$ paso: Verificación:

Hemos obtenido una media muestral de $\bar{x} = 12,7$ horas.

4º paso: Decisión:

Como la media muestral está fuera del intervalo de aceptación, rechazamos H_0, es decir, aceptamos que la media ha disminuido.

b) Si consideramos un nivel de significación del 1%, entonces $\alpha = 0,01 \rightarrow z_\alpha = 2,33$. La zona de aceptación sería el intervalo:

$$\left(15 - 2,33 \cdot \frac{1}{\sqrt{150}}; \ +\infty\right); \text{ es decir } (14,810; +\infty)$$

La media muestral sigue estando fuera de la zona de aceptación. Por tanto, tomaríamos la misma decisión: rechazaríamos H_0, es decir, aceptaríamos que la media ha disminuido.

27 **S** **Según la normativa sobre contaminación atmosférica, los motores de los automóviles no deben emitir más de 5 ppm (partes por millón) de CO_2. Dentro de sus procesos de control de calidad, un fabricante ha medido la emisión de CO_2 en una muestra de 36 motores, obteniendo una media de 5,5 ppm y una desviación típica de 0,6 ppm.**

Contrasta, con un nivel de significación igual a 0,05, la hipótesis de que los motores de este fabricante cumplen en media la normativa sobre contaminación.

1^er^ paso: Hipótesis: Tenemos que contrastar:

$$H_0: \mu \le 5 \quad \text{frente a} \quad H_1: \mu > 5$$

2º paso: Zona de aceptación:

$$\left(-\infty; \ \mu_0 + z_\alpha \cdot \frac{\sigma_0}{\sqrt{n}}\right)$$

Para un nivel de significación $\alpha = 0{,}05 \ \to \ z_\alpha = 1{,}645$. Como $\mu_0 = 5$; $\sigma_0 = 0{,}6$ y $n = 36$, tenemos que la zona de aceptación es:

$$\left(-\infty; \ 5 + 1{,}645 \cdot \frac{0{,}6}{\sqrt{36}}\right); \ \text{ es decir, el intervalo } (-\infty; 5{,}16)$$

3^er^ paso: Verificación:

La media muestral obtenida es $\bar{x} = 5{,}5$ ppm.

4º paso: Decisión:

Como la media muestral está fuera de la zona de aceptación, rechazamos H_0, es decir, no podemos aceptar que los motores de este fabricante cumplan la normativa sobre contaminación.

28 **S** **La Concejalía de Juventud de un Ayuntamiento maneja el dato de que la edad a la que los hijos se independizan de sus padres es una variable normal con media 29 años y desviación típica 3 años. Aunque la desviación típica no plantea dudas, sí se sospecha que la media ha descendido, sobre todo por la política de ayuda al empleo que ha llevado a cabo el Ayuntamiento. Así, de un estudio reciente sobre 100 jóvenes que se acaban de independizar, se ha obtenido una media de 28,1 años de edad.**

a) Con un nivel de significación del 1%, ¿puede defenderse que la edad media no ha disminuido, frente a que sí lo ha hecho como parecen indicar los datos? Plantea el contraste o test de hipótesis y resuélvelo.

b) Explica, en el contexto del problema, en qué consisten cada uno de los errores del tipo I y II.

a) **1^er^ paso: Hipótesis:** Tenemos que contrastar:

$$H_0: \mu \ge 29 \quad \text{frente a} \quad H_1: \mu < 29$$

2º paso: Zona de aceptación:

$$\left(\mu_0 - z_\alpha \cdot \frac{\sigma_0}{\sqrt{n}},\ +\infty\right)$$

Para un nivel de significación de $\alpha = 0{,}01$, tenemos que $z_\alpha = 2{,}33$. Así, el intervalo es:

$$\left(29 - 2{,}33 \cdot \frac{3}{\sqrt{100}};\ +\infty\right);\ \text{ es decir, } (28{,}301;\ +\infty)$$

3er paso: Verificación:

La media muestral obtenida es $\bar{x} = 28{,}1$ años.

4º paso: Decisión:

Como la media muestral está fuera del intervalo de aceptación, rechazamos H_0, es decir, aceptamos que la media de edad ha disminuido.

b) • El error de tipo I consiste en rechazar H_0 siendo verdadera. En el contexto de este problema sería aceptar que la media ha disminuido, siendo falso.

• El error de tipo II consiste en aceptar H_0 siendo falsa. En este problema sería aceptar que la media no ha disminuido, siendo falso.

29 **Un dentista afirma que el 40% de los niños de 10 años presentan indicios de caries dental. Tomada una muestra de 100 niños, se observó que 30 presentaban indicios de caries.**

Utilizando la aproximación normal, comprueba, a un nivel de significación del 5%, si el resultado proporciona evidencia que permita rechazar la afirmación del dentista.

1er paso: Hipótesis: Tenemos que contrastar:

H_0: $p = 0{,}4$ frente a H_1: $p \neq 0{,}4$

2º paso: Zona de aceptación:

$$\left(p_0 - z_{\alpha/2} \cdot \sqrt{\frac{p_0 q_0}{n}};\ p_0 + z_{\alpha/2} \cdot \sqrt{\frac{p_0 q_0}{n}}\right)$$

Para un nivel de significación $\alpha = 0{,}05$ tenemos que $z_{\alpha/2} = 1{,}96$. El intervalo será:

$$\left(0{,}4 - 1{,}96 \cdot \sqrt{\frac{0{,}4 \cdot 0{,}6}{100}};\ 0{,}4 + 1{,}96 \cdot \sqrt{\frac{0{,}4 \cdot 0{,}6}{100}}\right);\ \text{ es decir, } (0{,}304;\ 0{,}496)$$

3er paso: Verificación:

La proporción obtenida en la muestra es $pr = \dfrac{30}{100} = 0{,}3$.

4º paso: Decisión:

Como la proporción muestral queda fuera de la zona de aceptación, rechazamos H_0, es decir, rechazamos la afirmación del dentista.

30 **Una empresa de productos farmacéuticos afirma en su publicidad que uno de sus medicamentos reduce considerablemente los síntomas de la alergia primaveral en el 90% de la población.**

Una asociación de consumidores ha experimentado dicho fármaco en una muestra de 200 socios de la misma, obteniendo el resultado indicado en la publicidad en 170 personas.

Determina si la asociación de consumidores puede considerar que la afirmación de la empresa es estadísticamente correcta al nivel de significación de 0,05.

1er paso: Hipótesis: Tenemos que contrastar:

$H_0: p = 0{,}9$ frente a $H_1: p \neq 0{,}9$

2º paso: Zona de aceptación:

$$\left(p_0 - z_{\alpha/2} \cdot \sqrt{\frac{p_0 q_0}{n}};\ p_0 + z_{\alpha/2} \cdot \sqrt{\frac{p_0 q_0}{n}}\right)$$

Para un nivel de significación $\alpha = 0{,}05$, tenemos que $z_{\alpha/2} = 1{,}96$. El intervalo será:

$$\left(0{,}9 - 1{,}96 \cdot \sqrt{\frac{0{,}9 \cdot 0{,}1}{200}};\ 0{,}9 + 1{,}96 \cdot \sqrt{\frac{0{,}9 \cdot 0{,}1}{200}}\right); \text{ es decir, } (0{,}858;\ 0{,}942)$$

3er paso: Verificación:

La proporción obtenida en la muestra es $pr = \frac{170}{200} = 0{,}85$.

4º paso: Decisión:

Como la proporción muestral está fuera del intervalo de aceptación, rechazamos H_0, es decir, no podemos considerar válida la afirmación de la empresa.

31 S **Se afirma que, en una determinada ciudad, al menos el 30% de las familias poseen ordenador. Se toma una muestra aleatoria de 200 familias de la ciudad y resulta que 50 poseen ordenador.**

A un nivel de significación de 0,05, ¿hay suficiente evidencia para refutar la afirmación?

1er paso: Hipótesis: Queremos contrastar:

$H_0: p \geq 0{,}3$ frente a $H_1: p < 0{,}3$

2º paso: Zona de aceptación:

$$\left(p_0 - z_{\alpha} \cdot \sqrt{\frac{p_0 q_0}{n}};\ +\infty\right)$$

Para un nivel de significación $\alpha = 0{,}05$ tenemos que $z_\alpha = 1{,}645$. Por tanto, el intervalo será:

$$\left(0{,}3 - 1{,}645 \cdot \sqrt{\frac{0{,}3 \cdot 0{,}7}{200}};\ +\infty\right); \text{ es decir, } (0{,}247;\ +\infty)$$

3^er^ paso: Verificación:

La proporción obtenida en la muestra es $pr = \frac{50}{200} = 0{,}25$.

4º paso: Decisión:

Como la proporción muestral está dentro del intervalo de aceptación, no podemos rechazar H_0, es decir, aceptamos que, al menos, el 30% de las familias poseen ordenador.

32 **S** **Un investigador, utilizando información de anteriores comicios, sostiene que, en una determinada zona, el nivel de abstención en las próximas elecciones es del 40% como mínimo.**

Se elige una muestra aleatoria de 200 individuos para los que se concluye que 75 estarían dispuestos a votar.

Determina, con un nivel de significación del 1%, si se puede admitir como cierta la afirmación del investigador.

La proporción de abstenciones en la muestra es:

$$pr = \frac{120}{200} = 0{,}625 \ (62{,}5\%)$$

1^er^ paso: Hipótesis: Tenemos que contrastar:

$$H_0: p \le 0{,}4 \ \text{ frente a } \ H_1: p > 0{,}4$$

2º paso: Zona de aceptación:

$$\left(-\infty;\ p_0 + z_\alpha \cdot \sqrt{\frac{p_0 q_0}{n}}\right)$$

Para un nivel de significación del 1%, tenemos que $\alpha = 0{,}01 \ \rightarrow \ z_\alpha = 2{,}33$. El intervalo será:

$$\left(-\infty;\ 0{,}4 + 2{,}33 \cdot \sqrt{\frac{0{,}4 \cdot 0{,}6}{200}}\right); \ \text{ es decir, } (-\infty;\ 0{,}48)$$

3^er^ paso: Verificación:

La proporción de abstenciones en la muestra es $pr = 0{,}625$.

4º paso: Decisión:

Como la proporción muestral está fuera de la zona de aceptación, rechazamos H_0, es decir, aceptamos que la proporción de abstenciones superará el 40%.

Página 317

33 **S** **El 42% de los escolares de un cierto país suelen perder al menos un día de clase a causa de gripes y catarros. Sin embargo, un estudio sobre 1 000 escolares revela que en el último curso hubo 450 en tales circunstancias.**

Las autoridades sanitarias defienden que el porcentaje del 42% para toda la población de escolares se ha mantenido.

a) Contrasta, con un nivel de significación del 5%, la hipótesis defendida por las autoridades sanitarias, frente a que el porcentaje ha aumentado como parecen indicar los datos, explicando claramente a qué conclusión se llega.

b) ¿Cómo se llama la probabilidad de concluir erróneamente que el tanto por ciento se ha mantenido?

a) **1^{er} paso: Hipótesis:** Queremos contrastar:

H_0: $p \leq 0{,}42$ frente a H_1: $p > 0{,}42$

2º paso: Zona de aceptación:

$$\left(-\infty;\ p_0 + z_\alpha \cdot \sqrt{\frac{p_0 q_0}{n}}\right)$$

Para un nivel de significación del 5%, tenemos que $z_\alpha = 1{,}645$. Por tanto, el intervalo será:

$$\left(-\infty;\ 0{,}42 + 1{,}645 \cdot \sqrt{\frac{0{,}42 \cdot 0{,}58}{1\,000}}\right);\ \text{ es decir, } (-\infty;\ 0{,}446)$$

3^{er} paso: Verificación:

La proporción obtenida en la muestra es $pr = \frac{450}{1\,000} = 0{,}45$.

4º paso: Decisión:

Como la proporción muestral está fuera de la zona de aceptación, rechazamos H_0, es decir, aceptamos que la proporción ha aumentado.

b) La probabilidad de concluir erróneamente que el tanto por ciento se ha mantenido, es decir, de aceptar H_0, siendo falsa, es la probabilidad de cometer un error de tipo II.

CUESTIONES TEÓRICAS

34 **Con una muestra de 500 individuos hemos estimado, con un nivel de confianza del 90%, que la estatura media de los soldados de un cierto reemplazo está entre 174,3 cm y 175,1 cm (problema de la página inicial, pág. 290).**

a) Si la desviación típica de la población era desconocida, averigua la media, $\bar{x}$, y la desviación típica, s, de la muestra.

b) ¿Cuál sería el intervalo si la muestra fuera de tamaño la cuarta parte (500 : 4 = 125) y mantuviéramos el nivel de confianza?

a) • La media muestral es el punto medio del intervalo de confianza:

$$\bar{x} = \frac{174{,}3 + 175{,}1}{2} = 174{,}7 \text{ cm}$$

• La semiamplitud del intervalo es:

$$\frac{175{,}1 - 174{,}3}{2} = 0{,}4, \text{ que coincide con:}$$

$$z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} = 0{,}4 \quad \to \quad s = \frac{0{,}4 \cdot \sqrt{n}}{z_{\alpha/2}}$$

Para un 90% de confianza, tenemos que $z_{\alpha/2} = 1{,}645$. Por tanto, la desviación típica de la muestra es:

$$s = \frac{0{,}4 \cdot \sqrt{500}}{1{,}645} = 5{,}44$$

b) Si $z_{\alpha/2} = 1{,}645$, y mantenemos las condiciones del problema, salvo el tamaño muestral, que es $\frac{n}{4}$, el intervalo tendría el doble de amplitud que el anterior, pues:

$$E = z_{\alpha/2} \cdot \frac{s}{\sqrt{500/4}} = 2 \cdot z_{\alpha/2} \cdot \frac{s}{\sqrt{500}} = 2 \cdot 0{,}4 = 0{,}8$$

Es decir, el intervalo sería:

$(174{,}7 - 0{,}8;\ 174{,}7 + 0{,}8)$; esto es: $(173{,}9;\ 175{,}5)$

35 S **Supongamos que, a partir de una muestra aleatoria de tamaño $n = 25$, se ha calculado el intervalo de confianza para la media de una población normal, obteniéndose una amplitud igual a ± 4. Si el tamaño de la muestra hubiera sido $n = 100$, permaneciendo invariables todos los demás valores que intervienen en el cálculo, ¿cuál habría sido la amplitud del intervalo?**

La semiamplitud del intervalo es igual al error máximo admisible:

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}. \text{ Para } n = 25, \text{ sabemos que } E = 4.$$

Si $n = 100$, tendríamos que:

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{100}} = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{4 \cdot 25}} = \frac{1}{2}\left(z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{25}}\right) = \frac{1}{2} \cdot 4 = 2$$

La amplitud del intervalo sería ± 2.

36 **A partir de una muestra aleatoria, hemos estimado el peso de los toros de una manada mediante el intervalo (443, 528). ¿Cuál es la media de la muestra obtenida?**

La media muestral es el punto medio del intervalo de confianza, es decir:

$$\bar{x} = \frac{443 + 528}{2} = 485{,}5$$

37 **Mediante una muestra de 100 individuos estimamos la estatura de un colectivo de personas con un nivel de confianza del 95%. El error máximo admisible obtenido es $E = 1{,}274$. ¿Cuál es la desviación típica de la muestra obtenida?**

Para un nivel de confianza del 95%, tenemos que $z_{\alpha/2} = 1{,}96$.

El error máximo admisible es:

$E = z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$. Como $E = 1{,}274$, $n = 100$ y $z_{\alpha/2} = 1{,}96$, tenemos que:

$$1{,}274 = 1{,}96 \cdot \frac{s}{\sqrt{100}} \to s = \frac{12{,}74}{1{,}96} = 6{,}5 \to s = 6{,}5$$

38 **A partir de una muestra de tamaño 400 se estima la proporción de individuos que leen el periódico en una gran ciudad. Se obtiene una cota de error de 0,0392 con un nivel de confianza del 95%.**

a) ¿Podríamos, con la misma muestra, mejorar el nivel de confianza en la estimación? ¿Qué le ocurriría a la cota de error?

b) ¿Sabrías calcular la proporción, *pr*, obtenida en la muestra?

a) Aumentando la cota de error mejoraría el nivel de confianza.

b) La cota de error es:

$$E = z_{\alpha/2} \cdot \sqrt{\frac{pr(1-pr)}{n}}$$

Como $E = 0{,}0392$; $n = 400$ y $1 - \alpha = 0{,}95 \to z_{\alpha/2} = 1{,}96$, tenemos que:

$$0{,}0392 = 1{,}96 \cdot \sqrt{\frac{pr(1-pr)}{400}} \to \frac{0{,}0392}{1{,}96} = \sqrt{\frac{pr(1-pr)}{400}} \to$$

$$\to 0{,}02 = \sqrt{\frac{pr(1-pr)}{400}} \to 0{,}0004 = \frac{pr(1-pr)}{400} \to 0{,}16 = pr(1-pr)$$

$$0{,}16 = pr - pr^2 \to pr^2 - pr + 0{,}16 = 0$$

$$pr = \frac{1 \pm \sqrt{1 - 0{,}64}}{2} = \frac{1 \pm \sqrt{0{,}36}}{2} = \frac{1 \pm 0{,}6}{2} \begin{cases} pr = 0{,}8 \\ pr = 0{,}2 \end{cases}$$

Podría ser $pr = 0{,}8$ o bien $pr = 0{,}2$. Con los datos que tenemos, no podemos decidir cuál de estos dos resultados es el válido.

PARA PROFUNDIZAR

39 **En un test de hipótesis para estudiar si el cociente intelectual medio de los estudiantes de una universidad es 113, hemos seleccionado una muestra aleatoria de 180 estudiantes, obteniendo una media de 115. La zona de aceptación obtenida ha sido el intervalo (111,98; 114,02) y sabemos que la desviación típica es $\sigma = 7$. Por tanto, hemos rechazado la hipótesis.**

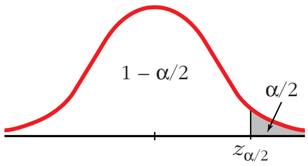
¿Cuál es la probabilidad de habernos equivocado, es decir, de haber rechazado la hipótesis, cuando en realidad era verdadera? ¿Cómo se llama este tipo de error?

El error que consiste en rechazar H_0 cuando esta es verdadera, se llama error de tipo I. La probabilidad de cometerlo es precisamente α, el nivel de significación. Lo calculamos en este caso concreto:

- La semiamplitud del intervalo de aceptación es:

$$z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}. \text{ En este caso concreto es igual a: } \frac{114{,}02 - 111{,}98}{2} = 1{,}02$$

- Sabemos que $\sigma = 7$ y que $n = 180$. Por tanto, podemos despejar $z_{\alpha/2}$:



$$1{,}02 = z_{\alpha/2} \cdot \frac{7}{\sqrt{180}} \rightarrow z_{\alpha/2} = 1{,}95 \rightarrow$$

$$\rightarrow 1 - \frac{\alpha}{2} = 0{,}9744 \rightarrow \alpha = 0{,}0512$$

- La probabilidad de haber cometido un error de tipo I (rechazar H_0 siendo cierta) es $\alpha = 0{,}0512$.

40 **En una determinada provincia, la nota media en matemáticas de los alumnos de 2º de Bachillerato del curso pasado fue de 5,8, con una desviación típica de 2,3 puntos. Con un nivel de significación de 0,05 y suponiendo que la desviación típica sigue siendo la misma, queremos contrastar la hipótesis de que la media no ha variado. Para ello, vamos a extraer una muestra aleatoria de tamaño 100. Así, la zona de aceptación será el intervalo (5,35; 6,25). Si al final la media real fuera de 5 puntos, ¿cuál es la probabilidad de obtener una media muestral que nos lleve a cometer un error de tipo II (es decir, aceptar H_0 siendo falsa)?**

Si la media real fuera $\mu = 5$, las medias muestrales en muestras de tamaño $n = 100$, con $\sigma = 2{,}3$, se distritribuirían según una $N\left(5; \frac{2{,}3}{\sqrt{100}}\right)$; es decir, según una $N(5;\ 0{,}23)$.

Así, la probabilidad de aceptar H_0 siendo falsa (esto es, la probabilidad de cometer un error de tipo II) sería la probabilidad de obtener una media muestral que cayera dentro de la zona de aceptación, es decir:

$$P[5{,}35 < \overline{x} < 6{,}25] = P\left[\frac{5{,}35 - 5}{0{,}23} < z < \frac{6{,}25 - 5}{0{,}23}\right] = P[1{,}52 < z < 5{,}43] =$$

$$= P[z < 5{,}43] - P[z < 1{,}52] = 1 - 0{,}9357 = 0{,}0643$$