13 INFERENCIA ESTADÍSTICA: ESTIMACIÓN DE UNA PROPORCIÓN

Página 298

¿Cuántas caras cabe esperar?

- El intervalo característico correspondiente a una probabilidad del 95% (consideramos "casas raros" al 5% de los casos extremos) es:

$$50 \pm 1,96 \cdot 5 = (40,2;\ 59,8)$$

Esto significa que en el 95% de los casos en que tiremos 100 monedas, el número de caras que obtendremos será mayor que 40 y menor que 60. Cualquier otro resultado será un "caso raro".

Página 299

Un saco de alubias

a) $p = \dfrac{500}{10\,000} = 0,05$

b) $\mu = 600 \cdot 0,05 = 30$; $\sigma = \sqrt{600 \cdot 0,05 \cdot 0,95} = \sqrt{28,5} \simeq 5,34$

c) El intervalo característico correspondiente a una probabilidad del 99% es:

$$30 \pm 2,575 \cdot 5,34 = (16,25;\ 43,75)$$

d) En el 99% de los casos en que saquemos 600 judías de esa saco, el número de judías negras será mayor que 16 y menor que 44. Cualquier otro resultado será un "caso raro" (llamando "casos raros" a ese 1% de casos extremos).

Peces en un pantano

La muestra tiene 514 peces, de los cuales hay 37 marcados. La proporción de peces marcados en la muestra es: $pr = \dfrac{37}{514} = 0,072$. El valor de la proporción de peces marcados en el pantano es $pr = \dfrac{349}{N}$, donde N es el número total de peces.

Aunque este problema se resolverá de forma completa (mediante un intervalo de confianza) al terminar la unidad, podemos suponer que la proporción de peces marcados en la muestra y en el pantano será "aproximadamente" la misma; es decir:

$$\frac{37}{514} \simeq \frac{349}{N} \quad \to N \simeq 4\,848{,}27 \quad \to N \simeq 4848 \text{ peces}$$

(Al considerar una probabilidad determinada, daremos un intervalo de confianza, obteniendo un resultado más preciso que este).

Página 301

1 La variable x es binomial, con $n = 1\,200$ y $p = 0{,}008$.

a) Calcula la probabilidad de que x sea mayor que 10.

b) Halla el intervalo característico para una probabilidad del 95%.

Como $np = 9{,}6 > 5$ y $nq > 5$, podemos aproximar mediante una normal de media $\mu = np = 9{,}6$ y de desviación típica $\sigma = \sqrt{npq} = \sqrt{1\,200 \cdot 0{,}008 \cdot 0{,}992} = 3{,}09$.

Es decir:

$$x \text{ es } B(1\,200;\ 0{,}008) \quad \to \quad x' \text{ es } N(9{,}6;\ 3{,}09) \quad \to \quad z \text{ es } N(0,\ 1)$$

a) $P[x > 10] = P[x' \geq 10{,}5] = P\left[z \geq \dfrac{10{,}5 - 9{,}6}{3{,}09}\right] = P[z \geq 0{,}29] =$

$= 1 - P[z < 0{,}29] = 1 - 0{,}6141 = 0{,}3859$

b) Para una probabilidad del 95%, $z_{\alpha/2} = 1{,}96$.

El intervalo característico será:

$(9{,}6 - 1{,}96 \cdot 3{,}09;\ 9{,}6 + 1{,}96 \cdot 3{,}09)$; es decir: $(3{,}54;\ 15{,}66)$

2 Si tenemos un dado correcto y lo lanzamos 50 veces:

a) ¿Cuál es la probabilidad de que "el 1" salga más de 10 veces?

b) ¿Cuál es la probabilidad de que salga "múltiplo de 3" al menos 20 veces?

a) Llamamos x = "nº de veces que sale el 1"; así, x es $B\left(50;\ \dfrac{1}{6}\right)$.

Como $np > 5$ y $nq > 5$, podemos aproximar mediante una normal de media $\mu = 50 \cdot \dfrac{1}{6} = 8{,}33$ y de desviación típica $\sigma = \sqrt{50 \cdot \dfrac{1}{6} \cdot \dfrac{5}{6}} = 2{,}64$; es decir:

$$x \text{ es } B\left(50;\ \frac{1}{6}\right) \quad \to \quad x' \text{ es } N(8{,}33;\ 2{,}64) \quad \to \quad z \text{ es } N(0,\ 1)$$

$P[x > 10] = P[x' \geq 10{,}5] = P\left[z \geq \dfrac{10{,}5 - 8{,}33}{2{,}64}\right] = P[z \geq 0{,}82] = 1 - P[z < 0{,}82] =$

$= 1 - 0{,}7939 = 0{,}2061$

b) Llamamos x = "nº de veces que sale múltiplo de 3". La probabilidad de obtener un múltiplo de 3 en una tirada es $p = \dfrac{2}{6} = \dfrac{1}{3}$. Así, x es $B\left(50;\ \dfrac{1}{3}\right)$.

Como $np > 5$ y $nq > 5$, podemos aproximar mediante una normal de media $\mu = 50 \cdot \frac{1}{3} = 16,67$ y de desviación típica $\sigma = \sqrt{50 \cdot \frac{1}{3} \cdot \frac{2}{3}} = 3,33$; es decir:

$$x \text{ es } B\left(50; \frac{1}{3}\right) \to x' \text{ es } N(16,67; 3,33) \to z \text{ es } N(0, 1)$$

$$P[x \geq 20] = P[x' \geq 19,5] = P\left[z \geq \frac{19,5 - 16,67}{3,33}\right] = P[z \geq 0,85] = 1 - P[z < 0,85] =$$

$$= 1 - 0,8023 = 0,1977$$

Página 303

1 Como sabemos, en un dado correcto la proporción de veces que sale el 5 es $1/6 = 0,1\widehat{6}$. Halla los intervalos característicos correspondientes al 90%, 95% y 99% para la “proporción de cincos”, en tandas de 100 lanzamientos de un dado correcto.

Las proporciones de cincos en tandas de 100 lanzamientos siguen una distribución normal de media $p = \frac{1}{6} = 0,17$ y de desviación típica $\sqrt{\frac{pq}{n}} = \sqrt{\frac{(1/6) \cdot (5/6)}{100}} =$ $= 0,037$; es decir:

$$pr \text{ es } N(0,17; 0,037)$$

Hallamos los intervalos característicos:

- Para el 90%: $(0,17 \pm 1,645 \cdot 0,037) = (0,109; 0,231)$
- Para el 95%: $(0,17 \pm 1,96 \cdot 0,037) = (0,097; 0,243)$
- Para el 99%: $(0,17 \pm 2,575 \cdot 0,037) = (0,075; 0,265)$

Página 305

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.

Páginas 308

Distribución de las proporciones muestrales. Intervalos característicos

1 **Averigua cómo se distribuyen las proporciones muestrales, pr, para las poblaciones y las muestras que se describen a continuación:**

	a)	b)	c)	d)	e)	f)
PROPORCIÓN, p, EN LA POBLACIÓN	0,5	0,6	0,8	0,1	0,05	0,15
TAMAÑO, n, DE LA MUESTRA	10	20	30	50	100	100

Recordemos que, si $np \geq 5$ y $nq \geq 5$, entonces, las proporciones muestrales siguen una distribución $N\left(p, \sqrt{\frac{pq}{n}}\right)$.

Aplicamos este resultado a cada uno de los casos propuestos. Comprobamos que en todo ellos se tiene que $np \geq 5$ y $nq \geq 5$.

a) $N\left(0{,}5; \sqrt{\frac{0{,}5 \cdot 0{,}5}{10}}\right)$; es decir, $N(0{,}5; 0{,}158)$

b) $N\left(0{,}6; \sqrt{\frac{0{,}6 \cdot 0{,}4}{20}}\right)$; es decir, $N(0{,}6; 0{,}110)$

c) $N\left(0{,}8; \sqrt{\frac{0{,}8 \cdot 0{,}2}{30}}\right)$; es decir, $N(0{,}8; 0{,}073)$

d) $N\left(0{,}1; \sqrt{\frac{0{,}1 \cdot 0{,}9}{50}}\right)$; es decir, $N(0{,}1; 0{,}042)$

e) $N\left(0{,}05; \sqrt{\frac{0{,}05 \cdot 0{,}95}{100}}\right)$; es decir, $N(0{,}05; 0{,}0218)$

f) $N\left(0{,}15; \sqrt{\frac{0{,}15 \cdot 0{,}85}{100}}\right)$; es decir, $N(0{,}15; 0{,}036)$

2 **Halla los intervalos característicos para las proporciones muestrales del ejercicio anterior, correspondientes a las probabilidades que, en cada caso, se indican:**

a) 90% **b) 95%** **c) 99%** **d) 95%** **e) 99%** **f) 80%**

a) $z_{\alpha/2} = 1{,}645$

Intervalo $(0{,}5 - 1{,}645 \cdot 0{,}158;\ 0{,}5 + 1{,}645 \cdot 0{,}158)$; es decir: $(0{,}24;\ 0{,}76)$

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

Intervalo $(0{,}6 - 1{,}96 \cdot 0{,}110;\ 0{,}6 + 1{,}96 \cdot 0{,}110)$; es decir: $(0{,}38;\ 0{,}82)$

c) $z_{\alpha/2} = 2{,}575$

Intervalo $(0{,}8 - 2{,}575 \cdot 0{,}073;\ 0{,}8 + 2{,}575 \cdot 0{,}073)$; es decir: $(0{,}61;\ 0{,}99)$

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

Intervalo $(0{,}1 - 1{,}96 \cdot 0{,}042;\ 0{,}1 + 1{,}96 \cdot 0{,}042)$; es decir: $(0{,}018;\ 0{,}182)$

e) $z_{\alpha/2} = 2{,}575$

Intervalo $(0{,}05 - 2{,}575 \cdot 0{,}0218;\ 0{,}05 + 2{,}575 \cdot 0{,}0218)$; es decir: $(-0{,}006;\ 0{,}106)$

f) $z_{\alpha/2} = 1{,}28$

Intervalo $(0{,}15 - 1{,}28 \cdot 0{,}036;\ 0{,}15 + 1{,}28 \cdot 0{,}036)$; es decir: $(0{,}104;\ 0{,}196)$

3 **Cuatro de cada diez habitantes de una determinada población lee habitualmente el periódico Z.**

Halla el intervalo característico para la proporción de habitantes de esa población que leen el periódico Z, en muestras de tamaño 49, correspondiente al 95%.

p = proporción de lectores del periódico Z $= \dfrac{4}{10} = 0{,}4$.

El intervalo característico para la proporción de lectores, pr, en muestras de tamaño n es de la forma:

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

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

El intervalo será:

$$\left(0{,}4 - 1{,}96 \cdot \sqrt{\frac{0{,}4 \cdot 0{,}6}{49}};\ 0{,}4 + 1{,}96 \cdot \sqrt{\frac{0{,}4 \cdot 0{,}6}{49}}\right);\ \text{es decir: } (0{,}26;\ 0{,}54)$$

4 **En un saco mezclamos judías blancas y judías pintas en la relación de 14 blancas por cada pinta.**

Extraemos un puñado de 100 judías.

a) ¿Cuál es la probabilidad de que la proporción de judías pintas esté comprendida entre 0,05 y 0,1?

b) Halla un intervalo en el cual se encuentre el 99% de las proporciones de las muestras de tamaño 100.

a) La proporción de judías pintas es $p = \frac{1}{15}$. Si extraemos un puñado de 100 judías, tenemos una binomial $B\left(100, \frac{1}{15}\right)$.

Una proporción entre 0,05 y 0,1 significa que haya entre $100 \cdot 0{,}05 = 5$ y $100 \cdot 0{,}1 = 10$ judías pintas.

Por tanto, si x es $B\left(100, \frac{1}{15}\right)$, tenemos que calcular $P[5 < x < 10]$.

Como $100 \cdot \frac{1}{15} > 5$ y $100 \cdot \frac{14}{15} > 5$, podemos aproximar la binomial mediante una normal de media $\mu = 100 \cdot \frac{1}{15} = 6{,}67$ y desviación típica $\sigma = \sqrt{100 \cdot \frac{1}{15} \cdot \frac{14}{15}} = 2{,}49$.

Así, si x es $B\left(100, \frac{1}{15}\right) \rightarrow x'$ es $N(6{,}67;\ 2{,}49) \rightarrow z$ es $N(0, 1)$. Calculamos:

$$P[5 < x < 10] = P[5{,}5 \le x' \le 9{,}5] = P\left[\frac{5{,}5 - 6{,}67}{2{,}49} \le z \le \frac{9{,}5 - 6{,}67}{2{,}49}\right] =$$

$$= P[-0{,}47 \le z \le 1{,}14] = P[z \le 1{,}14] - P[z \le -0{,}47] =$$

$$= P[z \le 1{,}14] - P[z \ge 0{,}47] = P[z \le 1{,}14] - (1 - P[z \le 0{,}47]) =$$

$$= 0{,}8729 - (1 - 0{,}6808) = 0{,}5537$$

b) Si consideramos muestras de tamaño 100, el intervalo característico para la proporción muestral es de la forma:

$$\left(p - z_{\alpha/2} \cdot \sqrt{\frac{pq}{100}},\ p + z_{\alpha/2} \cdot \sqrt{\frac{pq}{100}}\right)$$

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

Así, el intervalo será:

$$\left(\frac{1}{15} - 2{,}575 \cdot \sqrt{\frac{(1/15) \cdot (14/15)}{100}},\ \frac{1}{15} + 2{,}575 \cdot \sqrt{\frac{(1/15) \cdot (14/15)}{100}}\right);$$

es decir: $(0{,}0024;\ 0{,}1309)$

5 **En una localidad de 6 000 habitantes, la proporción de menores de 16 años es de 1 500/6 000.**

a) ¿Cuál es la distribución de la proporción de menores de 16 años en muestras de 50 habitantes de dicha población?

b) Halla la probabilidad de que, en una muestra de 50 habitantes, haya entre 15 y 20 menores de 16 años.

a) La proporción, pr, de menores de 16 años en muestras de tamaño $n = 50$ sigue una distribución normal de media $p = \dfrac{1500}{6000} = 0,25$ y de desviación típica:

$$\sqrt{\frac{pq}{n}} = \sqrt{\frac{0,25 \cdot 0,75}{50}} = 0,061, \text{ es decir, } pr \text{ es } N(0,26;\ 0,061).$$

b) El número de menores de 16 años en una muestra de 50 es una binomial $B(50;\ 0,25)$. Como $np > 5$ y $nq > 5$, podemos aproximar mediante una normal de media $\mu = 50 \cdot 0,25 = 12,5$ y de desviación típica:

$$\sigma = \sqrt{npq} = \sqrt{50 \cdot 0,25 \cdot 0,75} = 3,06$$

Así, si x es $B(50;\ 0,25)$ $\rightarrow$ x' es $N(12,5;\ 3,06)$ $\rightarrow$ z es $N(0,\ 1)$, entonces:

$$P[15 < x < 20] = P[15,5 < x' < 19,5] = P\left[\frac{15,5 - 12,5}{3,06} < z < \frac{19,5 - 12,5}{3,06}\right] =$$

$$= P[0,98 < z < 2,29] = P[z < 2,29] - P[z < 0,98] =$$

$$= 0,9890 - 0,8365 = 0,1525$$

6 El 42% de los habitantes de un municipio es contrario a la gestión del alcalde y el resto son partidarios de este. Si se toma una muestra de 64 individuos, ¿cuál es la probabilidad de que ganen los que se oponen al alcalde?

En muestras de 64, el número de personas que se oponen al alcalde, x, sigue una distribución binomial $B(64,\ 0,42)$. Tenemos que calcular $P[x > 32]$.

Como $np > 5$ y $nq > 5$, podemos aproximar mediante una normal de media $\mu = n \cdot p = 64 \cdot 0,42 = 26,88$ y de desviación típica $\sqrt{npq} = \sqrt{64 \cdot 0,42 \cdot 0,58} = 3,95$. Así, si: x es $B(64,\ 0,42)$ $\rightarrow$ x' es $N(26,88;\ 3,95)$ $\rightarrow$ z es $N(0,\ 1)$, entonces:

$$P[x > 32] = P[x' \geq 32,5] = P\left[z \geq \frac{32,5 - 26,88}{3,95}\right] = P[z \geq 1,42] =$$

$$= 1 - P[z < 1,42] = 1 - 0,9222 = 0,0778$$

7 La probabilidad de que un bebé sea varón es 0,515. Si han nacido 184 bebés, ¿cuál es la probabilidad de que haya 100 varones o más?

Halla el intervalo característico correspondiente al 95% para la proporción de varones en muestras de 184 bebés.

- El número de varones entre 184 bebés, x, sigue una distribución binomial $B(184;\ 0,515)$. Tenemos que calcular $P[x \geq 100]$. Como $np > 5$ y $nq > 5$, podemos aproximar mediante una normal de media $\mu = np = 184 \cdot 0,515 = 94,76$ y de desviación típica $\sqrt{npq} = \sqrt{184 \cdot 0,515 \cdot 0,485} = 6,78$. Así, si:

  x es $B(184;\ 0,515)$ $\rightarrow$ x' es $N(94,76;\ 6,78)$ $\rightarrow$ z es $N(0,\ 1)$, entonces:

$$P[x \geq 100] = P[x' \geq 99{,}5] = P\left[z \geq \frac{99{,}5 - 94{,}76}{6{,}78}\right] = P[z \geq 0{,}70] =$$

$$= 1 - P[z < 0{,}70] = 1 - 0{,}7580 = 0{,}2420$$

- El intervalo característico para la proporción muestral es de la forma:

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

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

Así, el intervalo será:

$$\left(0{,}515 - 1{,}96 \cdot \sqrt{\frac{0{,}515 \cdot 0{,}485}{184}},\ 0{,}515 + 1{,}96 \cdot \sqrt{\frac{0{,}515 \cdot 0{,}485}{184}}\right);$$

es decir: $(0{,}4428; 0{,}5872)$

8 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 $pr = \dfrac{75}{350} = \dfrac{3}{14}$

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

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:

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

9 Se selecciona aleatoriamente una muestra de 600 personas en una ciudad y se les pregunta si consideran que el tráfico en la misma es aceptablemente fluido. Responden afirmativamente 250 personas.

¿Cuál es el intervalo de confianza de la proporción de ciudadanos de esa ciudad que consideran aceptable la fluidez del tráfico, con un nivel de confianza del 90%?

La proporción muestral es $pr = \dfrac{250}{600} = \dfrac{5}{12} \rightarrow 1 - pr = \dfrac{7}{12}$

Para un nivel de confianza del 90%, sabemos que $z_{\alpha/2} = 1{,}645$.

El intervalo de confianza para la proporción de ciudadanos que consideran aceptable la fluidez del tráfico es:

$$\left(pr - z_{\alpha/2} \cdot \sqrt{\frac{pr\,(1 - pr)}{n}}\ ;\ pr + z_{\alpha/2} \cdot \sqrt{\frac{pr\,(1 - pr)}{n}}\right)$$

En este caso queda:

$$\left(\frac{5}{12} - 1{,}645 \cdot \sqrt{\frac{\frac{5}{12} \cdot \frac{7}{12}}{600}}\ ;\ \frac{5}{12} + 1{,}645 \cdot \sqrt{\frac{\frac{5}{12} \cdot \frac{7}{12}}{600}} \right);\ \text{es decir: } (0{,}3836;\ 0{,}4498)$$

10 **Sabemos que al lanzar al suelo 100 chinchetas, en el 95% de los casos, la proporción de ellas que quedan con la punta hacia arriba está en el intervalo (0,1216; 0,2784). Calcula la probabilidad p de que una de esas chinchetas caiga con la punta hacia arriba y comprueba que la amplitud del intervalo dado es correcta.**

- p es el centro del intervalo, es decir:

$$p = \frac{0{,}2784 + 0{,}1216}{2} = 0{,}2 = p$$

- Veamos que la amplitud del intervalo dado es correcta:

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

  El intervalo característico es:

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

  En este caso ($p = 0{,}2$; $q = 0{,}8$; $n = 100$; $z_{\alpha/2} = 1{,}96$), queda:

$$\left(0{,}2 - 1{,}96 \cdot \sqrt{\frac{0{,}2 \cdot 0{,}8}{100}};\ 0{,}2 + 1{,}96 \cdot \sqrt{\frac{0{,}2 \cdot 0{,}8}{100}} \right);\ \text{es decir:}$$

  (0,1216; 0,2784), como queríamos probar.

11
S **De 120 alumnos, la proporción de que tengan dos o más hermanos es de 48/120. Indica los parámetros de la distribución a la que se ajustarían las muestras de tamaño 30.**

En muestras de tamaño $n = 30$, la proporción muestral, pr, seguiría una distribución normal de media:

$$\mu = np = 30 \cdot \frac{48}{120} = 30 \cdot 0{,}4 = 12$$

y de desviación típica:

$$\sigma = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0{,}4 \cdot 0{,}6}{30}} = 0{,}089$$

Es decir, pr es $N(12;\ 0{,}089)$.

12 **¿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, $1 - \alpha = 0{,}95 \quad \rightarrow \quad z_{\alpha/2} = 1{,}96$

El error máximo admisible es:

$$E = z_{\alpha/2} \cdot \sqrt{\frac{pr(1 - pr)}{n}}$$. Buscamos n para que $E \le 0{,}04$ (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 $n = 115$.

13 **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:

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

$(0{,}196;\ 0{,}504)$

14 **En una muestra de 100 rótulos publicitarios se observa que aparecen 6 defectuosos.**

a) Estima la proporción real de rótulos defectuosos, con un nivel de confianza del 99%.

b) ¿Cuál es el error máximo cometido al hacer la estimación anterior?

c) ¿De qué tamaño tendríamos que coger la muestra, con un nivel de confianza del 99%, para obtener un error inferior a 0,05?

a) La proporción muestral es $pr = \frac{6}{100} = 0,06 \rightarrow 1 - pr = 0,94$

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

El intervalo de confianza para estimar la proporción real de rótulos defectuosos es:

$$\left(pr - z_{\alpha/2} \cdot \sqrt{\frac{pr\,(1 - pr)}{n}} \,;\, pr + z_{\alpha/2} \cdot \sqrt{\frac{pr\,(1 - pr)}{n}} \right)$$

En este caso queda:

$$\left(0,06 - 2,575 \cdot \sqrt{\frac{0,06 \cdot 0,94}{100}} \,;\, 0,06 + 2,575 \cdot \sqrt{\frac{0,06 \cdot 0,94}{100}} \right) \text{es decir: } (0;\, 0,12)$$

b) $E = z_{\alpha/2} \cdot \sqrt{\frac{pr\,(1 - pr)}{n}} = 2,575 \cdot \sqrt{\frac{0,06 \cdot 0,94}{100}} \simeq 0,06$

c) En la expresión del error, sabemos que:

$E = 0,05$

$z_{\alpha/2} = 2,575$ (para un nivel de confianza del 99%)

$pr = 0,06; \quad 1 - pr = 0,94$

Por tanto:

$$E = z_{\alpha/2} \cdot \sqrt{\frac{pr\,(1 - pr)}{n}} \rightarrow 0,05 = 2,575 \cdot \sqrt{\frac{0,06 \cdot 0,94}{n}} \rightarrow n \simeq 149,58$$

Habrá que tomar una muestra de, al menos, 150 rótulos.

15 Tomada al azar una muestra de 60 estudiantes de una universidad, se encontró que un tercio hablaban el idioma inglés.

a) Halla, con un nivel de confianza del 90%, un intervalo para estimar la proporción de estudiantes que hablan el idioma inglés entre los estudiantes de esa universidad.

b) A la vista del resultado anterior se pretende repetir la experiencia para conseguir una cota de error de 0,01 con el mismo nivel de confianza del 90%. ¿Cuántos individuos ha de tener la muestra?

La proporción muestral es $pr = \frac{1}{3} \rightarrow 1 - pr = \frac{2}{3}$

Para un nivel de confianza del 90%, sabemos que $z_{\alpha/2} = 1,645$.

a) El intervalo de confianza para estimar la proporción en la población es:

$$\left(pr - z_{\alpha/2} \cdot \sqrt{\frac{pr\,(1-pr)}{n}}\ ;\ pr + z_{\alpha/2} \cdot \sqrt{\frac{pr\,(1-pr)}{n}}\right)$$

En este caso queda:

$$\left(\frac{1}{3} - 1{,}645 \cdot \sqrt{\frac{\frac{1}{3} \cdot \frac{2}{3}}{60}}\ ;\ \frac{1}{3} + 1{,}645 \cdot \sqrt{\frac{\frac{1}{3} \cdot \frac{2}{3}}{60}}\right);\ \text{es decir: } (0{,}2332;\ 0{,}4334)$$

b) En la expresión del error, sabemos que:

$E = 0{,}01$

$z_{\alpha/2} = 1{,}645$ (para un nivel de confianza del 90%)

$pr = \frac{1}{3};\quad 1 - pr = \frac{2}{3}$

Por tanto:

$$E = z_{\alpha/2} \cdot \sqrt{\frac{pr\,(1-pr)}{n}} \rightarrow 0{,}01 = 1{,}645 \cdot \sqrt{\frac{\frac{1}{3} \cdot \frac{2}{3}}{60}} \rightarrow n \simeq 6\,013{,}4$$

Habrá que tomar una muestra de, al menos, 6 014 individuos.

16 **Para estimar la proporción de habitantes de una determinada ciudad que poseen ordenador personal, se quiere utilizar una muestra aleatoria de tamaño *n*. Calcula el valor mínimo de *n* para garantizar que, con un nivel de confianza del 95%, el error en la estimación no sea superior al 2%.**

☛ ***Como se desconoce la proporción, se tiene que tomar el caso más desfavorable, que será 0,5.***

En la expresión del error, sabemos que:

$E = 0{,}02$ (error no superior al 2%)

$z_{\alpha/2} = 1{,}96$ (nivel de confianza del 95%)

$pr = 0{,}5;\quad 1 - pr = 0{,}5$ (al desconocer la proporción, debemos tomar el caso más desfavorable, que es 0,5).

Por tanto:

$$E = z_{\alpha/2} \cdot \sqrt{\frac{pr\,(1-pr)}{n}} \rightarrow 0{,}02 = 1{,}96 \cdot \sqrt{\frac{0{,}5 \cdot 0{,}5}{60}} \rightarrow n \simeq 2401$$

Habrá que tomar una muestra de, al menos, 2401 individuos.

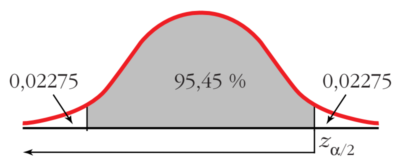
17 **En una encuesta realizada a 800 personas elegidas al azar del censo electoral, 240 declaran su intención de votar al partido *A*.**

a) Estima, con un nivel de confianza del 95,45%, entre qué valores se encuentra la intención de voto al susodicho partido en todo el censo.

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

La proporción muestral es $pr = \dfrac{240}{800} = 0,3 \;\rightarrow\; 1 - pr = 0,7$

a) Para un nivel de confianza del 95,45%, hallamos $z_{\alpha/2}$:



$$0,02275 + 0,9545 = 0,97725 \;\rightarrow\; P[z \le z_{\alpha/2}] = 0,97725 \;\rightarrow\; z_{\alpha/2} = z_{\alpha/2} = 2$$

El intervalo de confianza para estimar la proporción en la población es:

$$\left(pr - z_{\alpha/2} \cdot \sqrt{\frac{pr\,(1 - pr)}{n}}\,;\; pr + z_{\alpha/2} \cdot \sqrt{\frac{pr\,(1 - pr)}{n}}\right);\text{ en este caso queda:}$$

$$\left(0,3 - 2 \cdot \sqrt{\frac{0,3 \cdot 0,7}{800}}\,;\; 0,3 + 2 \cdot \sqrt{\frac{0,3 \cdot 0,7}{800}}\right);\text{ es decir: } (0,2676;\ 0,3324)$$

La proporción de votantes del partido A en la población se encuentra, con un nivel de confianza del 95,45%, entre el 26,76% y el 33,24%.

b) Si aumenta el nivel de confianza, mayor es la amplitud del intervalo; es decir, cuanto más seguros queramos estar de nuestra estimación, mayor será el error máximo admisible.

Si disminuye el nivel de confianza, disminuye la amplitud del intervalo.

18 **Una reciente encuesta, realizada en un cierto país sobre una muestra aleatoria de 800 personas, arroja el dato de que 300 de ellas son analfabetas. Para estimar la proporción de analfabetos del país hemos obtenido el siguiente intervalo de confianza: (0,3414; 0,4086)**

¿Cuál es el nivel de confianza con el que se ha hecho la estimación?

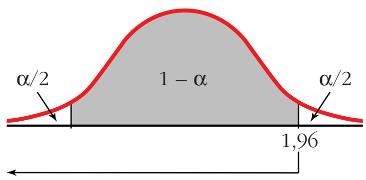
La proporción muestral es $pr = \dfrac{300}{800} = \dfrac{3}{8} \;\rightarrow\; 1 - pr = \dfrac{5}{8}$

El error máximo admisible es la semiamplitud del intervalo de confianza; es decir:

$$E = \frac{0,4086 - 0,3414}{2} = 0,0336$$

Por tanto:

$$E = z_{\alpha/2} \cdot \sqrt{\frac{pr\,(1 - pr)}{n}} \;\rightarrow\; 0{,}0336 = z_{\alpha/2} \cdot \sqrt{\frac{(3/8)\cdot(5/8)}{800}} \;\rightarrow\; z_{\alpha/2} = 1{,}96$$



$P[z \le 1{,}96] = 0{,}9750$

$\dfrac{\alpha}{2} = P[z > 1{,}96] = 1 - 0{,}9750 = 0{,}025$

$\alpha = 0{,}025 \cdot 2 = 0{,}05 \;\rightarrow\; 1 - \alpha = 0{,}95$

El nivel de confianza es del 95%.

19 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 $1 - \alpha = 0{,}95 \;\rightarrow\; z_{\alpha/2} = 1{,}96$, tenemos que:

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

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

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

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

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

20 a) Un fabricante de medicamentos afirma que cierta medicina cura una enfermedad de la sangre en el 80% de los casos. Los inspectores de sanidad utilizan el medicamento en una muestra de 100 pacientes y deciden aceptar dicha afirmación si se curan 75 o más.

Si lo que afirma el fabricante es realmente cierto, ¿cuál es la probabilidad de que los inspectores rechacen dicha afirmación?

b) Si en la muestra se curan 60 individuos, con una confianza del 95%, ¿cuál es el error máximo cometido al estimar que el porcentaje de efectividad del medicamento es del 60%?

a) Si lo que dice el fabricante es realmente cierto, tenemos que:

$p = 0{,}8 \ \rightarrow \ 1 - p = 0{,}2$

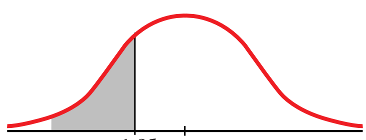
Considerando una muestra de tamaño $n = 100$, las proporciones muestrales, pr, siguen una distribución normal de media $p = 0{,}8$ y de desviación típica

$\sqrt{\dfrac{pq}{n}} = \sqrt{\dfrac{0{,}8 \cdot 0{,}2}{100}} = 0{,}04$; es decir, pr es $N(0{,}8;\ 0{,}04)$.

La probabilidad de que los inspectores rechacen la afirmación es:

$P\left[pr < \dfrac{\mathbf{75}}{\mathbf{100}}\right]$,

Calculamos esta probabilidad:



$P\left[pr < \dfrac{75}{100}\right] = P[pr < 0{,}75] =$

$= P\left[z < \dfrac{0{,}75 - 0{,}8}{0{,}04}\right] = P[z < -1{,}25] =$

$= P[z > 1{,}25] = 1 - P[z \le 1{,}25] = 1 - 0{,}8944 = 0{,}1056$

b) Si la proporción muestral es $pr = \dfrac{60}{100} = 0{,}6 \ \rightarrow \ 1 - pr = 0{,}4$

Para $z_{\alpha/2} = 1{,}96$ (nivel de confianza del 95%),

el error máximo será:

$E = z_{\alpha/2} \cdot \sqrt{\dfrac{pr\,(1 - pr)}{n}} = 1{,}96 \cdot \sqrt{\dfrac{0{,}6 \cdot 0{,}4}{100}} \simeq 0{,}096$

El error máximo cometido es de unas 10 personas.