Estadística II
Ejercicios Tema 1
Curso 2009/10 – Soluciones

Estimación puntual

1. Una muestra aleatoria simple de precios de una cierta fruta en 9 tiendas ha dado los valores siguientes

2.5 2.1 1.8 2.1 2.3 2.2 2.2 2.0 1.9

Se pide que calcules:

a) La media, la varianza y la desviación típica de la muestra.

b) ¿Cuáles de los estimadores empleados en la pregunta anterior son insesgados (para los valores correspondientes de la población)?

c) Emplea un procedimiento de estimación insesgado para estimar la varianza de la media muestral.

d) Emplea un estimador insesgado para estimar la proporción de tiendas en las que el precio de venta de la fruta es inferior a 2.1 euros.

Solución.

a) Aplicamos las fórmulas habituales para obtener

$$\begin{aligned}
\bar{x} &= \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{19{,}1}{9} = 2{,}122 \\
s^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2 = \frac{40{,}89}{9} - 2{,}122^2 = 0{,}0395 \\
s &= \sqrt{s^2} = 0{,}1988
\end{aligned}$$

b) La media muestral es un estimador insesgado de la media de la población, pero el estimador insesgado de la varianza de la población es

$$\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \frac{1}{n-1}\sum_{i=1}^{n} x_i^2 - \frac{n}{n-1}\bar{x}^2 = \frac{40{,}89}{8} - \frac{9}{8}2{,}122^2 = 0{,}0444$$

De este estimador obtenemos el estimador de la desviación típica.

c) La varianza de la media muestral para una muestra aleatoria simple vale σ^2/n. Una estimación insesgada de este valor a partir de la muestra dada será

$$\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\frac{1}{n} = \frac{0{,}0444}{9} = 0{,}004938$$

d) Un estimador insesgado para una proporción en la población es la proporción en la muestra. En este caso, contando el número de tiendas con precios inferiores a 2.1 euros (3) tenemos

$$\hat{p} = \frac{3}{9} = 0{,}333$$

2. Se ha realizado un estudio de mercado con 64 personas, a las que se ha preguntado si estarían dispuestas a comprar dos productos. 36 han contestado afirmativamente para el primer producto, 32 lo han hecho para el segundo y 26 lo han hecho para ambos.

Se pide que:

a) Emplees un procedimiento de estimación insesgado para estimar el porcentaje de los consumidores que estarían dispuestos a comprar ambos productos.

b) Emplees un procedimiento de estimación insesgado para estimar el porcentaje de los consumidores que estarían dispuestos a comprar el primer producto pero no el segundo.

c) Indica la varianza del estimador correspondiente a esta segunda pregunta y estima su valor.

Solución.

a) Para estimar proporciones en la población de manera insesgada podemos emplear como estimadores las proporciones muestrales correspondientes. En nuestro caso el valor pedido es

$$\hat{p}_1 = \frac{26}{64} = 0{,}406$$

b) Para calcular esta segunda estimación empleamos el mismo procedimiento que antes, pero necesitamos el número de personas en la muestra que cumplen la condición. En este caso de los 36 clientes dispuestos a comprar el primer producto, $36 - 26 = 10$ no están dispuestos a comprar el segundo, luego la proporción deseada es

$$\hat{p}_2 = \frac{10}{64} = 0{,}156$$

c) El estimador se puede escribir como la media de variables Bernoulli asociadas a la condición fijada. Su varianza por tanto vendrá dada por la varianza de cada variable $p_2(1-p_2)$ dividida por n,

$$\mathrm{Var}[\hat{p}_2] = \frac{p_2(1-p_2)}{n}$$

Una estimación de esta varianza será $0{,}156(1-0{,}156)/64 = 0{,}00206$.

3. Si X_1, X_2, X_3 es una muestra aleatoria simple de una población con media μ y desviación típica σ, se define a partir de esta muestra dos estimadores para la media de la población:

$$\hat{\mu}_1 = \frac{X_1 + 2X_2 + 2X_3}{5}, \quad \hat{\mu}_2 = \frac{X_1 + 3X_2 + X_3}{5}$$

Se pide que:

a) Verifiques que ambos estimadores son insesgados.

b) Determines cuál de los dos estimadores es más eficiente y calcules su eficiencia relativa.

c) Encuentres otro estimador insesgado que sea más eficiente que cualquiera de ellos.

Solución.

a) Calculamos el valor esperado de ambos estimadores, teniendo en cuenta que $E[X_i] = \mu$,

$$\begin{aligned} E[\hat{\mu}_1] &= E\left[\frac{X_1 + 2X_2 + 2X_3}{5}\right] = \frac{E[X_1] + 2E[X_2] + 2E[X_3]}{5} = \frac{5\mu}{5} = \mu \\ E[\hat{\mu}_2] &= E\left[\frac{X_1 + 3X_2 + X_3}{5}\right] = \frac{E[X_1] + 3E[X_2] + E[X_3]}{5} = \frac{5\mu}{5} = \mu \end{aligned}$$

luego ambos estimadores son insesgados.

b) Calculamos las eficiencias de los estimadores a partir de sus varianzas, teniendo en cuenta que las variables X_i son independientes y que $\mathrm{Var}[X_i] = \sigma^2$,

$$\begin{aligned} \mathrm{Var}[\hat{\mu}_1] &= \mathrm{Var}\left[\frac{X_1 + 2X_2 + 2X_3}{5}\right] = \frac{\mathrm{Var}[X_1] + 4\mathrm{Var}[X_2] + 4\mathrm{Var}[X_3]}{25} = \frac{9\sigma^2}{25} = 0{,}36\sigma^2 \\ \mathrm{Var}[\hat{\mu}_2] &= \mathrm{Var}\left[\frac{X_1 + 3X_2 + X_3}{5}\right] = \frac{\mathrm{Var}[X_1] + 9\mathrm{Var}[X_2] + \mathrm{Var}[X_3]}{25} = \frac{11\sigma^2}{25} = 0{,}44\sigma^2 \end{aligned}$$

El primer estimador es más eficiente (tiene una varianza menor), y la eficiencia relativa es el cociente de varianzas, en este caso $9/11$.

c) Podemos ver que la eficiencia aumenta (la varianza disminuye) si los pesos son más parecidos. El candidato más razonable (óptimo en algunos casos particulares) es la media muestral, que tiene pesos iguales para todas las observaciones. Su eficiencia es

$$\mathrm{Var}[\hat{\mu}] = \mathrm{Var}\left[\frac{X_1 + X_2 + X_3}{3}\right] = \frac{\mathrm{Var}[X_1] + \mathrm{Var}[X_2] + \mathrm{Var}[X_3]}{9} = \frac{3\sigma^2}{9} = 0{,}33\sigma^2$$

mejor que los valores anteriores.

Intervalos de confianza para la media - Población normal y varianza conocida

4. Uno de los productos de una empresa es harina en paquetes de 500 gr. Los pesos de estos paquetes en realidad siguen una distribución normal con desviación típica conocida, igual a 8 gr. En un determinado día se ha seleccionado una muestra aleatoria simple de 64 envases, y se ha medido que su peso medio es de 498 gr.

   a) Calcula un intervalo de confianza al 99 % para el peso medio de los paquetes envasados en el día.
   b) ¿Cuánto valdría el intervalo de confianza anterior si el nivel de confianza se fijase al 95 %?
   c) Supongamos que el valor real de la desviación típica para la producción del día fuese de 12 gr. ¿Cuál sería el intervalo de confianza al 99 % correspondiente a este valor?
   d) Si aumentasemos el tamaño de la muestra con un nivel de confianza del 99 % y el valor promedio del peso en la muestra se mantuviese en 498 gr. (con desviación típica igual a 12 gr.) ¿Para que tamaño de muestra quedaría fuera del intervalo el peso oficial de 500 gr.?

Solución.

a) En este caso, al disponer de una muestra aleatoria simple de datos normales y conocer la desviación típica de la población, podemos utilizar el resultado

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

En la distribución normal estándar el valor que deja a su derecha una probabilidad del 0.5 % (la mitad del complementario del nivel de confianza, $\alpha/2$) es $z_{\alpha/2} = 2{,}576$. Obtenemos este valor por ejemplo buscando en las tablas el valor correspondiente a $\alpha/2 = 0{,}005$.
El intervalo de confianza lo construimos como los valores que cumplen

$$-z_{\alpha/2} \leq \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}$$

y sustituyendo valores teniendo en cuenta que $\bar{x} = 498$, $\sigma = 8$ y $n = 64$,

$$-2{,}576 \leq \frac{498 - \mu}{8/\sqrt{64}} \leq 2{,}576 \Leftrightarrow 495{,}424 \leq \mu \leq 500{,}576$$

el intervalo pedido.

b) Si cambia el nivel de confianza, en los cálculos anteriores cambia el valor de $z_{\alpha/2}$ que pasa a ser ahora el valor correspondiente a $\alpha/2 = 0{,}025$, y $z_{\alpha/2} = 1{,}96$. Con este valor el intervalo queda

$$-1{,}96 \leq \frac{498 - \mu}{8/\sqrt{64}} \leq 1{,}96 \Leftrightarrow 496{,}04 \leq \mu \leq 499{,}96$$

un poco menor que el anterior.

c) Cambiamos ahora el valor de σ a 12. El nuevo valor del intervalo será

$$-2{,}576 \leq \frac{498 - \mu}{12/\sqrt{64}} \leq 2{,}576 \Leftrightarrow 494{,}136 \leq \mu \leq 501{,}864$$

y el intervalo aumenta proporcionalmente al valor de σ.

d) Para responder a la pregunta debemos despejar el valor de n para el que se tiene que un extremo del intervalo coincide con el valor 500. Tenemos que

$$500 = 498 + 2{,}576\frac{12}{\sqrt{n}} \Rightarrow n = \left(\frac{2{,}576 \times 12}{500 - 498}\right)^2 = 239$$

Para valores de n mayores que este, el valor 500 no está en el intevalo de confianza.

5. En el proceso de selección de candidatos para un puesto en cierta empresa se pide a dichos candidatos que realicen un test de aptitud. Los resultados de dicho test se supone que siguen una distribución normal con desviación típica conocida, igual a 1.15 puntos. Se ha seleccionado una muestra aleatoria simple de 16 resultados de dicho test y se ha obtenido un valor medio muestral de 4.26 puntos.

   Se pide que:

   a) Calcules un intervalo de confianza al 90 % para la calificación media de todos los candidatos.

   b) Te han dado un intervalo de confianza calculado por otra persona con los datos anteriores, desde 3.86 a 4.66 puntos. ¿Qué nivel de confianza corresponde a dicho intervalo?

   Solución.

   a) Tenemos una muestra aleatoria simple de una población normal con desviación típica conocida. Podemos utilizar el resultado

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

Para el caso que nos indican, empezamos calculando el valor del cuantil de interés de la distribución normal estándar, $z_{\alpha/2}$. En este caso $\alpha = 0{,}1$ y de las tablas de la normal estándar el valor que corresponde a $\alpha/2 = 0{,}05$ es $z_{0{,}05} = 1{,}645$. El intervalo pedido tiene la forma

$$-z_{\alpha/2} \leq \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2} \tag{1}$$

y sustituyendo valores teniendo en cuenta que $\bar{x} = 4{,}26$, $\sigma = 1{,}15$ y $n = 16$,

$$-1{,}645 \leq \frac{4.26 - \mu}{1{,}15/\sqrt{16}} \leq 1{,}645 \Leftrightarrow 3{,}787 \leq \mu \leq 4{,}733$$

el intervalo pedido.

   b) Para el intervalo dado tenemos que su longitud es $4{,}66 - 3{,}86 = 0{,}8$ (y su centro es $(3{,}86 + 4{,}66)/2 = 4{,}26$, la media muestral). La expresión del intervalo, despejando en (1), es

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

y su longitud es $2z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$. Igualando y despejando el valor de $z_{\alpha/2}$ tenemos

$$0{,}8 = 2z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 2z_{\alpha/2}\frac{1{,}15}{\sqrt{16}} \Rightarrow z_{\alpha/2} = 1{,}391$$

y de las tablas de la normal este valor corresponde a $\alpha/2 = 0{,}0821$ luego el nivel de confianza para este intervalo es $1 - \alpha = 0{,}836$, esto es, el 83.6 %.

6. Las ventas semanales de automóviles (medidas en euros) en los concesionarios de una cadena de venta de coches se supone que siguen una distribución normal con desviación típica conocida igual a 21.000 euros. Se ha recogido una muestra de 9 valores de ventas semanales, medidas en miles de euros:

   123 145 88 150 115 128 97 104 125

   Se pide que calcules:

   a) Un intervalo de confianza para la media de las ventas mensuales en todos los concesionarios al 95 %.

b) El valor de dicho intervalo si se cambia el nivel de confianza al 99 %.

c) ¿Cuál sería el intervalo de confianza al 95 % si empleamos solo los 6 primeros valores de la muestra?

Solución.

a) Tenemos una muestra aleatoria simple de una población normal con desviación típica conocida. Podemos utilizar el resultado

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Para el caso que nos indican, empezamos calculando el valor del cuantil de interés, $z_{\alpha/2}$. En este caso $\alpha = 0{,}05$ y de las tablas de la normal estándar el valor que corresponde a $\alpha/2 = 0{,}025$ es $z_{0,025} = 1{,}96$. El intervalo pedido tiene la forma

$$-z_{\alpha/2} \leq \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}$$

y sustituyendo valores teniendo en cuenta que $\bar{x}$, la media muestral, vale en este caso $1075/9 = 119{,}44$, $\sigma = 21$ y $n = 9$,

$$-1{,}96 \leq \frac{119{,}44-\mu}{21/\sqrt{9}} \leq 1{,}96 \Leftrightarrow 105{,}72 \leq \mu \leq 133{,}16$$

el intervalo pedido.

b) Si cambia el nivel de confianza, cambiamos el valor de $z_{\alpha/2}$ a $z_{0,005} = 2{,}576$ y obtenemos

$$-2{,}576 \leq \frac{119{,}44-\mu}{21/\sqrt{9}} \leq 2{,}576 \Leftrightarrow 101{,}412 \leq \mu \leq 137{,}476$$

c) Si empleamos los 6 primeros valores, tenemos $\bar{x} = 749/6 = 124{,}83$ y $n = 6$, con lo que el intervalo será

$$-1{,}96 \leq \frac{124{,}83-\mu}{21/\sqrt{6}} \leq 1{,}96 \Leftrightarrow 106{,}80 \leq \mu \leq 142{,}87$$

Intervalos de confianza para la media - Muestras grandes

7. A una muestra aleatoria simple de 225 expertos en marketing se les solicitó su opinión sobre el nivel de uso de determinadas técnicas cuantitativas en sus empresas (en una escala de 0 a 5). Los resultados obtenidos dieron una calificación media en la muestra de 3.35 y una desviación típica de 0.72.

   a) Calcula un intervalo de confianza para el valor medio del nivel de uso de dichas técnicas en la población al 99 %

   b) ¿A qué nivel de confianza correspondería un intervalo de 3.25 a 3.45?

   c) Si se mantuviesen los valores de la media y la desviación típica muestral, encuentra los tamaños de muestra para los que el intervalo de confianza al 99 % tendría un tamaño inferior a 0.20.

Solución.

a) En este caso disponemos de una muestra aleatoria simple de una población que puede no seguir una distribución normal, pero el tamaño de la muestra es elevado por lo que podemos aplicar el teorema central del límite, y tener aproximadamente

$$\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim N(0,1)$$

Para el caso que nos indican, empezamos calculando el valor del cuantil de interés de la distribución normal estándar, $z_{\alpha/2}$. En este caso $\alpha = 0{,}01$ y de las tablas de la normal estándar el valor que corresponde a $\alpha/2 = 0{,}005$ es $z_{0,005} = 2{,}576$. El intervalo pedido tiene la forma

$$-z_{\alpha/2} \leq \frac{\bar{x}-\mu}{s/\sqrt{n}} \leq z_{\alpha/2}$$

y sustituyendo valores teniendo en cuenta que $\bar{x} = 3{,}35$, $s = 0{,}72$ (la desviación típica muestral) y $n = 225$,

$$-2{,}576 \leq \frac{3{,}35 - \mu}{0{,}72/\sqrt{225}} \leq 2{,}576 \Leftrightarrow 3{,}226 \leq \mu \leq 3{,}474$$

el intervalo pedido.

b) La longitud del intervalo dado es $3{,}45 - 3{,}25 = 0{,}2$ y su centro es $(3{,}25 + 3{,}45)/3 = 3{,}35$, la media muestral. La longitud teórica del intervalo de confianza

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

es $2z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$. Igualando y despejando el valor de $z_{\alpha/2}$ tenemos

$$0{,}2 = 2z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 2z_{\alpha/2}\frac{0{,}72}{\sqrt{225}} \Rightarrow z_{\alpha/2} = 2{,}083$$

y de las tablas de la normal este valor corresponde a $\alpha/2 = 0{,}0186$, luego el nivel de confianza para este intervalo es $1 - \alpha = 1 - 2 \times 0{,}0186 = 0{,}963$, esto es, 96.3 %.

c) El tamaño del intervalo es $2z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$. Manteniendo los datos que nos indican y despejando el valor de n tenemos

$$0{,}2 = 2z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 2 \times 2{,}576\frac{0{,}72}{\sqrt{n}} \Rightarrow n = \left(\frac{2 \times 2{,}576 \times 0{,}72}{0{,}2}\right)^2 = 344$$

Para tamaños muestrales mayores o iguales que 344 el tamaño del intervalo es inferior a 0.2.

8. En una muestra aleatoria simple de 900 consumidores se determinó un consumo medio mensual de 325 euros en transporte, con una desviación típica muestral de 112 euros.

   a) Calcula un intervalo de confianza al 90 % para la media del consumo mensual en transporte en la población.

   b) Calcula el intervalo de confianza al 90 % para la media de la población si la desviación típica muestral fuera de 140 euros y el tamaño de muestra fuera de 400 consumidores, manteniendose el consumo medio mensual.

Solución.

a) Disponemos de una muestra aleatoria simple de una población cuya distribución desconocemos, pero el tamaño de la muestra es elevado, por lo que podemos aplicar el teorema central del límite, y aproximadamente

$$\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim N(0,1)$$

Para el caso que nos indican, empezamos calculando el valor del cuantil de interés de la distribución normal estándar, $z_{\alpha/2}$. En este caso $\alpha = 0{,}1$ y de las tablas de la normal estándar el valor que corresponde a $\alpha/2 = 0{,}05$ es $z_{0{,}05} = 1{,}645$. El intervalo pedido tiene la forma

$$-z_{\alpha/2} \leq \frac{\bar{x} - \mu}{s/\sqrt{n}} \leq z_{\alpha/2}$$

y sustituyendo valores teniendo en cuenta que $\bar{x} = 325$, $s = 112$ (la desviación típica muestral) y $n = 900$,

$$-1{,}645 \leq \frac{325 - \mu}{112/\sqrt{900}} \leq 1{,}645 \Leftrightarrow 318{,}86 \leq \mu \leq 331{,}14$$

b) Si cambian los valores de s y n a $s = 140$ y $n = 400$, obtenemos

$$-1{,}645 \leq \frac{325 - \mu}{140/\sqrt{400}} \leq 1{,}645 \Leftrightarrow 313{,}485 \leq \mu \leq 336{,}515$$

9. Se ha obtenido una muestra aleatoria simple de 100 datos de consumo diario de energía eléctrica, obteniéndose los valores siguientes:

$$\sum_{i=1}^{100} x_i = 1{,}723 \qquad \sum_{i=1}^{100} x_i^2 = 45{,}21$$

   a) Calcula la media y la desviación típica muestral. Estima también por un procedimiento insesgado la varianza de la población.

   b) Calcula el intervalo de confianza al 95 % para la media de la población.

Solución.

   a) De los datos que nos ofrecen calculamos

$$\begin{aligned} \bar{x} &= \frac{1}{n}\sum_i x_i = \frac{1{,}723}{100} = 0{,}01723 \\ s_m^2 &= \frac{1}{n}\sum_i (x_i-\bar{x})^2 = \frac{1}{n}\sum_i x_i^2 - \bar{x}^2 = \frac{45{,}21}{100} - 0{,}01723^2 = 0{,}4518 \\ s_m &= \sqrt{s_m^2} = 0{,}6722 \end{aligned}$$

   Para estimar de manera insesgada la varianza debemos emplear la cuasivarianza muestral,

$$s^2 = \frac{1}{n-1}\sum_i (x_i-\bar{x})^2 = \frac{1}{n-1}\sum_i x_i^2 - \frac{n}{n-1}\bar{x}^2 = \frac{45{,}21}{99} - \frac{100}{99}0{,}01723^2 = 0{,}4564$$

   b) Disponemos de una muestra aleatoria simple de una población cuya distribución desconocemos, pero el tamaño de la muestra es elevado, por lo que podemos aplicar el teorema central del límite, y aproximadamente

$$\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim N(0,1)$$

   Para el caso que nos indican, empezamos calculando el valor del cuantil de interés de la distribución normal estándar, $z_{\alpha/2}$. En este caso $\alpha = 0{,}05$ y de las tablas de la normal estándar el valor que corresponde a $\alpha/2 = 0{,}025$ es $z_{0{,}025} = 1{,}96$. El intervalo pedido tiene la forma

$$-z_{\alpha/2} \le \frac{\bar{x}-\mu}{s/\sqrt{n}} \le z_{\alpha/2}$$

   y sustituyendo valores teniendo en cuenta que $\bar{x} = 0{,}01723$, $s^2 = 0{,}4564$ (empleamos el estimador insesgado de la varianza de la población), $s = 0{,}6755$ y $n = 100$,

$$-1{,}96 \le \frac{0{,}01723-\mu}{0{,}6755/\sqrt{100}} \le 1{,}96 \Leftrightarrow -0{,}115 \le \mu \le 0{,}150$$

Intervalos de confianza para proporciones - Muestras grandes

10. Se ha preguntado a una muestra aleatoria de 90 pequeñas empresas si consideran que las mejoras en la calidad de productos y procesos son las iniciativas más importantes para la mejora de su competitividad. De ellas, 35 contestaron afirmativamente.

   a) Calcula un intervalo de confianza del 99 % para la proporción en la población que considera estas iniciativas las más importantes.

   b) ¿Cuál sería el valor del intervalo para un nivel de confianza del 95 %?

Solución.

   a) Para construir el intervalo de confianza pedido asociaremos a las respuestas de las empresas una variable aleatoria Bernoulli que tome el valor 1 si la respuesta es afirmativa, con probabilidad p. La proporción que considera importantes las iniciativas mencionadas se puede considerar como la media de estas variables Bernoulli. La varianza de estas variables es $p(1-p)$. Como el tamaño muestral es razonablemente elevado podemos aplicar el teorema central del límite para tener aproximadamente

$$\frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} \sim N(0,1)$$

donde p es la proporción en la población, $\hat{p}$ denota la proporción en la muestra y hemos aproximado la desviación típica en la población mediante su equivalente muestral $\hat{p}(1-\hat{p})$.

Para la distribución normal estándar tenemos que el cuantil de interés es el correspondiente a $\alpha/2$ con $\alpha = 0{,}01$ en este caso. Dicho valor será en este caso $z_{0,005} = 2{,}576$.

El intervalo pedido tiene la forma

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

y sustituyendo valores teniendo en cuenta que $\hat{p} = 35/90 = 0{,}389$ y $n = 90$,

$$-2{,}576 \leq \frac{0{,}389-p}{\sqrt{0{,}389(1-0{,}389)/90}} \leq 2{,}576 \Leftrightarrow 0{,}257 \leq \mu \leq 0{,}521$$

b) Si el nivel de confianza cambia al 95 %, ahora $\alpha = 0{,}05$ y el cuantil de interés será el correspondiente a $\alpha/200{,}025$; de las tablas de la normal estándar $z_{0,025} = 1{,}96$. Cambiando este valor en la fórmula anterior tenemos

$$-1{,}96 \leq \frac{0{,}389-p}{\sqrt{0{,}389(1-0{,}389)/90}} \leq 1{,}96 \Leftrightarrow 0{,}288 \leq \mu \leq 0{,}490$$

11. Un ayuntamiento está pensando mejorar las infraestructuras deportivas de la ciudad, y para financiar esta mejora contempla un puequeño aumento temporal del impuesto de bienes inmuebles. Ha realizado una encuesta entre 150 vecinos, de los cuales un 44 % se han mostrado a favor de esta medida.

    a) Calcula un intervalo de confianza del 95 % para la proporción de los vecinos del ayuntamiento a favor de la medida.

    b) Si se mantuviese la proporción favorable de vecinos y el nivel de confianza, ¿qué tamaño de muestra mínimo sería necesario para dejar fuera del intervalo al valor del 50 %?

Solución.

a) Asociamos a las respuestas de los vecinos una variable aleatoria Bernoulli que tome el valor 1 si la respuesta es favorable, con probabilidad p. La proporción que considera importantes las iniciativas mencionadas se puede considerar como la media de estas variables Bernoulli, cuya varianza es $p(1-p)$. Como el tamaño muestral es razonablemente elevado podemos aplicar el teorema central del límite para tener aproximadamente

$$\frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} \sim N(0,1)$$

donde p es la proporción en la población, $\hat{p}$ denota la proporción en la muestra y hemos aproximado la desviación típica en la población mediante su equivalente muestral $\hat{p}(1-\hat{p})$.

Para la distribución normal estándar tenemos que el cuantil de interés es el correspondiente a $\alpha/2$ con $\alpha = 0{,}05$ en este caso. De las tablas de la normal estándar para $\alpha/2 = 0{,}005$ dicho valor será $z_{0,025} = 1{,}96$.

El intervalo pedido tiene la forma

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

y sustituyendo valores teniendo en cuenta que $\hat{p} = 0{,}44$ y $n = 150$,

$$-1{,}96 \leq \frac{0{,}44-p}{\sqrt{0{,}44(1-0{,}44)/150}} \leq 1{,}96 \Leftrightarrow 0{,}361 \leq \mu \leq 0{,}519$$

b) Obtenemos el valor de n deseado igualando el extremo superior del intervalo

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

al valor 0,5, y despejando en la ecuación el valor de n.

$$0{,}5 = \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0{,}44 + 1{,}96\sqrt{\frac{0{,}44(1-0{,}44)}{n}}$$

$$\Rightarrow \quad n = \left(\frac{1{,}96\sqrt{0{,}44(1-0{,}44)}}{0{,}5-0{,}44}\right)^2 = 263$$

Hace falta por tanto una muestra de al menos 263 vecinos.

12. Se ha realizado una encuesta entre 500 votantes para preguntar por su intención de voto en relación con un cierto partido político. 183 de los encuestados han mostrado su intención de votar a este partido en las próximas elecciones.

a) Calcula un intervalo de confianza al 99 % para la intención de voto por dicho partido en la población.

b) ¿Qué tamaño de muestra mínimo sería necesario para que el tamaño del intervalo fuese de un 1 %, suponiendo que la proporción de votantes favorables y el nivel de confianza se mantuviesen constantes?

c) Si para los datos iniciales te diesen un intervalo de confianza del 30.3 % al 42.9 %, ¿a qué nivel de confianza correspondería?

Solución.

a) Asociamos a la intención de voto de los votantes una variable aleatoria Bernoulli que tome el valor 1 si se tiene la intención de votar al partido, con probabilidad p. La proporción de la población con la intención de votar al partido se puede considerar como la media de estas variables Bernoulli, cuya varianza es $p(1-p)$. Como el tamaño muestral es razonablemente elevado podemos aplicar el teorema central del límite para tener aproximadamente

$$\frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} \sim N(0,1)$$

donde p es la proporción en la población, $\hat{p}$ denota la proporción en la muestra y hemos aproximado la desviación típica en la población mediante su equivalente muestral $\hat{p}(1-\hat{p})$.
Para la distribución normal estándar tenemos que el cuantil de interés es el correspondiente a $\alpha/2$ con $\alpha = 0{,}01$ en este caso. Dicho valor será en este caso $z_{0{,}005} = 2{,}576$.
El intervalo pedido tiene la forma

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

y sustituyendo valores teniendo en cuenta que $\hat{p} = 183/500 = 0{,}366$ y $n = 500$,

$$-2{,}576 \leq \frac{0{,}366-p}{\sqrt{0{,}366(1-0{,}366)/500}} \leq 2{,}576 \Leftrightarrow 0{,}311 \leq \mu \leq 0{,}421$$

b) Si queremos que el intervalo tenga un tamaño de 0.01 (un 1 %) debemos despejar el valor de n necesario del tamaño del intervalo. Dicho intervalo viene dado por

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

y su tamaño es $2z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}$. Igualando tenemos

$$0{,}01 = 2z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 2 \times 2{,}576\sqrt{\frac{0{,}366(1-0{,}366)}{n}}$$

$$\Rightarrow \quad n = \left(\frac{2 \times 2{,}576\sqrt{0{,}366(1-0{,}366)}}{0{,}01}\right)^2 = 61592$$

luego en este caso haría falta una muestra de tamaño muy elevado.

c) El tamaño del intervalo que nos dan es $0{,}429 - 0{,}303 = 0{,}126$ y su centro es $(0{,}429 + 0{,}303)/2 = 0.366$, la proporción muestral. Despejamos el valor de $z_{\alpha/2}$ de la expresión del tamaño del intervalo.

$$0{,}126 = 2z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 2 \times z_{\alpha/2}\sqrt{\frac{0{,}366(1-0{,}366)}{500}}$$
$$\Rightarrow \quad z_{\alpha/2} = \frac{0{,}126}{2\sqrt{0{,}366(1-0{,}366)/500}} = 2{,}924$$

Este valor de $z_{\alpha/2}$ corresponde en las tablas a $\alpha/2 = 0{,}00173$ y por tanto el nivel de confianza es $1 - \alpha = 1 - 2 \times 0{,}00173 = 0{,}99654$, esto es, el 99.65 %.

Intervalos de confianza para la media - Datos normales y varianza desconocida

13. En una clínica se ofrecen tratamientos para la reducción de peso en pacientes. Se supone que la disminución de peso que se observa tras un tratamiento de dos meses sigue una distribución normal. De una muestra aleatoria simple de 16 pacientes se han obtenido los valores de reducción que se indican a continuación:

12,5	14,3	9,8	15,3	10,5	11,8	9,5	8,4
11,3	8,9	10,6	12,0	14,1	8,8	12,1	9,4

a) Calcula un intervalo de confianza al 95 % para la reducción de peso media en todos los pacientes que siguen el tratamiento.

b) ¿Cuál sería el valor del intervalo para un nivel de confianza del 99 %?

Solución.

a) Tenemos una muestra aleatoria simple de una población normal con desviación típica desconocida. El resultado sobre distribuciones que podemos emplear es

$$\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t_{n-1}$$

donde t_{n-1} denota una distribución t de Student con $n-1$ grados de libertad.
Para el caso que nos indican, empezamos calculando el valor del cuantil de interés, $t_{n-1,\alpha/2}$. Tenemos $\alpha = 0.05$ y por tanto $\alpha/2 = 0.025$, también $n = 16$. De las tablas de la t de Student con 15 grados de libertad el valor que corresponde a $\alpha/2 = 0{,}025$ es $t_{15,0{,}025} = 2{,}131$. El intervalo pedido tiene la forma

$$-t_{n-1,\alpha/2} \le \frac{\bar{x} - \mu}{s/\sqrt{n}} \le t_{n-1,\alpha/2}$$

Sustituimos valores teniendo en cuenta que las cantidades muestrales de interes son

$$\begin{aligned} \bar{x} &= \frac{1}{n}\sum_i x_i = \frac{179{,}3}{16} = 11{,}206 \\ s^2 &= \frac{1}{n-1}\sum_i (x_i - \bar{x})^2 = \frac{1}{n-1}\sum_i x_i^2 - \frac{n}{n-1}\bar{x}^2 = \frac{2075{,}45}{15} - \frac{16}{15}11{,}206^2 = 4{,}411 \\ s &= \sqrt{s^2} = 2{,}1 \end{aligned}$$

para obtener el intervalo

$$-2{,}131 \le \frac{11{,}206 - \mu}{2{,}1/\sqrt{16}} \le 2{,}131 \Leftrightarrow 10{,}09 \le \mu \le 12{,}33$$

b) Para un nivel de confianza del 99 % tenemos $\alpha = 0{,}01$ y $\alpha/2 = 0{,}005$. De las tablas de la t de Student tenemos $t_{\infty,0{,}005} = 1{,}98$, y sustituyendo este valor en la expresión del intervalo,

$$-2{,}947 \le \frac{11{,}206 - \mu}{2{,}1/\sqrt{16}} \le 2{,}947 \Leftrightarrow 9{,}66 \le \mu \le 12{,}75$$

14. En un tramo de una carretera se han medido las velocidades a las que circulaban 12 conductores, y se han obtenido los valores siguientes

89	82	95	102	84	80
86	79	96	81	98	86

Suponiendo que las velocidades en dicho tramo sigan una distribución normal,

a) Calcula un intervalo de confianza al 99 % para la velocidad media de circulación de todos los conductores en dicho tramo.

b) ¿Cuál sería el valor del intervalo si se calcula a partir de los primeros 8 valores de la muestra?

c) En el caso de la muestra completa, ¿para qué valores del nivel de confianza el intervalo resultante no incluiría el límite de velocidad, que es de 90 Km/h en ese tramo?

Solución.

a) Tenemos una muestra aleatoria simple de una población normal con desviación típica desconocida. El resultado básico a emplear para construir el intervalo de confianza pedido es

$$\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{n-1}$$

donde t_{n-1} denota una distribución t de Student con $n-1$ grados de libertad.
Para el caso que nos indican, empezamos calculando el valor del cuantil de interés, $t_{n-1,\alpha/2}$. Tenemos $\alpha = 0{,}01$ y por tanto $\alpha/2 = 0{,}005$, así como $n = 12$. De las tablas de la t de Student con 11 grados de libertad el valor que corresponde a $\alpha/2 = 0{,}005$ es $t_{11,0{,}005} = 3{,}106$. El intervalo pedido tiene la forma

$$-t_{n-1,\alpha/2} \leq \frac{\bar{x}-\mu}{s/\sqrt{n}} \leq t_{n-1,\alpha/2}$$

Sustituimos valores teniendo en cuenta que las cantidades muestrales de interes son

$$\begin{aligned}
\bar{x} &= \frac{1}{n}\sum_i x_i = \frac{1058}{12} = 88{,}17 \\
s^2 &= \frac{1}{n-1}\sum_i (x_i-\bar{x})^2 = \frac{1}{n-1}\sum_i x_i^2 - \frac{n}{n-1}\bar{x}^2 = \frac{93944}{11} - \frac{12}{11}88{,}17^2 = 60{,}33 \\
s &= \sqrt{s^2} = 7{,}767
\end{aligned}$$

para obtener el intervalo

$$-3{,}106 \leq \frac{88{,}17-\mu}{7{,}767/\sqrt{12}} \leq 3{,}106 \Leftrightarrow 81{,}20 \leq \mu \leq 95{,}13$$

b) Si consideramos los primeros ocho valores, $n = 8$ y de la muestra tenemos

$$\begin{aligned}
\bar{x} &= \frac{1}{n}\sum_i x_i = \frac{697}{8} = 87{,}125 \\
s^2 &= \frac{1}{n-1}\sum_i (x_i-\bar{x})^2 = \frac{1}{n-1}\sum_i x_i^2 - \frac{n}{n-1}\bar{x}^2 = \frac{61167}{7} - \frac{8}{7}87{,}125^2 = 62{,}982 \\
s &= \sqrt{s^2} = 7{,}936
\end{aligned}$$

Sustituyendo estos valores

$$-3{,}106 \leq \frac{87{,}125-\mu}{7{,}936/\sqrt{8}} \leq 3{,}106 \Leftrightarrow 78{,}41 \leq \mu \leq 95{,}84$$

c) Para encontrar el valor pedido igualamos el límite superior del intervalo, $\bar{x} + t_{n-1,\alpha/2}s/\sqrt{n}$, al valor del límite de interés, 90, y despejamos el valor de $t_{n-1,\alpha/2}$. Obtenemos

$$90 = \bar{x} + t_{n-1,\alpha/2}\frac{s}{\sqrt{n}} = 88{,}17 + t_{n-1,\alpha/2}\frac{7{,}767}{12} \Rightarrow t_{n-1,\alpha/2} = \frac{90 - 88{,}17}{7{,}767/12} = 2{,}832$$

Debemos buscar este valor en la tabla de la t de Student con 11 grados de libertad. El valor correcto es (aunque no es fácil obtenerlo de la tabla) $\alpha/2 = 0{,}0082$ y por tanto el nivel de confianza es $1 - \alpha = 1 - 2 \times 0{,}0082 = 0.9836$, esto es, el 98.36 %.

15. En la universidad se hace un seguimiento de los salarios de los antiguos alumnos. Para una muestra de 20 antiguos alumnos de una titulación se ha observado que la media (muestral) de los salarios mensuales un año después de su graduación es de 1700 euros y la desviación típica es de 350 euros. Suponemos que la distribución en la población es normal.

a) Calcula un intervalo de confianza al 95 % para los sueldos medios de los antiguos alumnos de la titulación.

b) ¿Cuál sería el valor del intervalo con un nivel de confianza del 99 %?

c) ¿Para qué tamaño de muestra se tendría un intervalo de tamaño menor de 100 euros, si se mantuviesen los valores indicados de media y desviación típica muestral y un nivel de confianza del 95 %?

Solución.

a) Tenemos una muestra aleatoria simple de una población normal con desviación típica desconocida. El resultado básico a emplear para construir el intervalo de confianza pedido es

$$\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t_{n-1}$$

donde t_{n-1} denota una distribución t de Student con $n-1$ grados de libertad.
Para el caso que nos indican, empezamos calculando el valor del cuantil de interés, $t_{n-1,\alpha/2}$. Tenemos $\alpha = 0{,}05$ y por tanto $\alpha/2 = 0{,}025$, así como $n = 20$. De las tablas de la t de Student con 19 grados de libertad el valor que corresponde a $\alpha/2 = 0{,}025$ es $t_{19,0{,}025} = 2{,}093$. El intervalo pedido tiene la forma

$$-t_{n-1,\alpha/2} \leq \frac{\bar{x} - \mu}{s/\sqrt{n}} \leq t_{n-1,\alpha/2}$$

Sustituimos valores teniendo en cuenta que las cantidades muestrales de interes son $\bar{x} = 1700$ y $s = 350$, para obtener el intervalo

$$-2{,}093 \leq \frac{1700 - \mu}{350/\sqrt{20}} \leq 2{,}093 \Leftrightarrow 1536{,}2 \leq \mu \leq 1863{,}8$$

b) Modificamos el nivel de confianza y tenemos ahora $\alpha = 0{,}01$ y por tanto $\alpha/2 = 0{,}005$, así como $n = 20$. De las tablas de la t de Student con 19 grados de libertad el valor que corresponde a $\alpha/2 = 0{,}005$ es $t_{19,0{,}005} = 2.861$.
Sustituimos este valor para obtener el intervalo

$$-2{,}861 \leq \frac{1700 - \mu}{350/\sqrt{20}} \leq 2{,}861 \Leftrightarrow 1476{,}1 \leq \mu \leq 1923{,}9$$

c) Para encontrar el tamaño de muestra pedido igualamos el tamaño del intervalo de confianza en función de n al valor dado, y despejamos n. El intervalo tiene la forma

$$\bar{x} - t_{n-1,\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}$$

y su tamaño es $2t_{n-1,\alpha/2}s/\sqrt{n}$. Tenemos que tanto $t_{n-1,\alpha/2}$ como $\sqrt{n}$ dependen de n, y es complicado despejar el valor deseado de

$$2t_{n-1,\alpha/2}s/\sqrt{n} = 100.$$

Pero para valores elevados de $n-1$ el valor del cuantil $t_{n-1,\alpha/2}$ apenas varía con n (ver las tablas), y podemos emplear en una primera aproximación el valor correspondiente a $n=\infty$ para hacer el cálculo, $t_{\infty,0{,}025} = 1{,}96$. Igualando y despejando obtenemos

$$100 = 2t_{n-1,\alpha/2}\frac{s}{\sqrt{n}} = 2\times 1{,}96\frac{350}{\sqrt{n}} \Rightarrow n = \left(\frac{2\times 1{,}96\times 350}{100}\right)^2 = 189$$

Para $n=189$ el valor de $t_{n-1,\alpha/2}$ es 1,9726 (del ordenador) y el valor del tamaño del intervalo de confianza es $2t_{n-1,\alpha/2}s/\sqrt{n} = 100{,}44$, muy cerca del valor deseado. El valor correcto se obtiene (por prueba y error) para $n=191$.

Intervalos de confianza para la varianza - Datos normales

16. Una compañía telefónica ha realizado un estudio del tiempo de uso diario del móvil por sus clientes. Para una muestra de 25 clientes se ha obtenido un valor de la desviación típica muestral de 12.5 minutos. Suponemos que el tiempo de uso sigue una distribución normal.

    a) Calcula un intervalo de confianza al 95 % para la varianza del tiempo de uso en la población.

    b) ¿Cuál sería el valor del intervalo para un nivel de confianza del 99 %?

Solución.

a) Tenemos una muestra aleatoria simple de una población normal y nos piden estimar un intervalo de confianza para la varianza de la población. El resultado básico para construir el intervalo de confianza pedido es

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$$

donde χ^2_{n-1} denota una distribución chi-cuadrado con $n-1$ grados de libertad.
Para el caso que nos indican, empezamos calculando el valor de los cuantiles de interés. Como la distribución chi-cuadrado no es simétrica, debemos calcular dos valores, $\chi^2_{n-1,1-\alpha/2}$ y $\chi^2_{n-1,\alpha/2}$. Tenemos $\alpha=0{,}05$ y por tanto $\alpha/2=0{,}025$ y $1-\alpha/2=0{,}975$, así como $n=25$. De las tablas de la chi-cuadrado con 24 grados de libertad obtenemos los valores $\chi^2_{24,0{,}975}=12{,}401$ y $\chi^2_{24,0{,}025}=39{,}364$. El intervalo pedido tiene la forma

$$\chi^2_{n-1,1-\alpha/2} \le \frac{(n-1)s^2}{\sigma^2} \le \chi^2_{n-1,\alpha/2}$$

Sustituimos valores teniendo en cuenta que la cantidad muestral de interes es $s=12{,}5$, para obtener el intervalo

$$12{,}401 \le \frac{24\times 12{,}5^2}{\sigma^2} \le 39{,}364 \Leftrightarrow \frac{24\times 12{,}5^2}{39{,}364} \le \sigma^2 \le \frac{24\times 12{,}5^2}{12{,}401} \Leftrightarrow 95{,}265 \le \sigma^2 \le 302.395$$

b) Modificamos el nivel de confianza y tenemos ahora $\alpha=0{,}01$ y por tanto $\alpha/2=0{,}005$ y $1-\alpha/2=0{,}995$. De las tablas de la chi-cuadrado con 24 grados de libertad obtenemos los valores $\chi^2_{24,0{,}995}=9{,}886$ y $\chi^2_{24,0{,}005}=45{,}559$.
Sustituimos estos valores para obtener el nuevo intervalo

$$9{,}886 \le \frac{24\times 12{,}5^2}{\sigma^2} \le 45{,}559 \Leftrightarrow \frac{24\times 12.5^2}{45{,}559} \le \sigma^2 \le \frac{24\times 12{,}5^2}{9{,}886} \Leftrightarrow 82{,}311 \le \sigma^2 \le 379{,}324$$

17. Se ha consultado a 12 analistas financieros sobre los beneficios por acción que pueda ofrecer una empresa en el próximo ejercicio, y se ha obtenido la muestra siguiente:

2.3	1.6	0.6	1.9	1.2	3.5
1.6	2.0	0.8	0.0	1.1	1.6

Suponemos que los valores indicados forman una muestra aleatoria simple y que estas predicciones siguen una distribución normal.

a) Calcula un intervalo de confianza al 95 % para la desviación típica de la población.

b) ¿Cuál sería el valor del intervalo si se calculase empleando únicamente los primeros 6 valores de la muestra?

Solución.

a) Tenemos una muestra aleatoria simple de una población normal y nos piden estimar un intervalo de confianza para la desviación típica de la población. Lo haremos a partir de un intervalo de confianza para la varianza. El resultado básico para construir el intervalo de confianza pedido es

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$$

donde χ^2_{n-1} denota una distribución chi-cuadrado con $n-1$ grados de libertad.

Para el caso que nos indican, empezamos calculando el valor de los cuantiles de interés. Tenemos $\alpha = 0{,}05$ y por tanto $\alpha/2 = 0{,}025$ y $1-\alpha/2 = 0{,}975$, así como $n = 12$. De las tablas de la chi-cuadrado con 11 grados de libertad obtenemos los valores $\chi^2_{11,0{,}975} = 3{,}816$ y $\chi^2_{11,0{,}025} = 21{,}92$. El intervalo para la varianza tiene la forma

$$\chi^2_{n-1,1-\alpha/2} \leq \frac{(n-1)s^2}{\sigma^2} \leq \chi^2_{n-1,\alpha/2}$$

Sustituimos valores teniendo en cuenta que la cantidad muestral de interes es

$$\begin{aligned}
\bar{x} &= \frac{1}{n}\sum_i x_i = \frac{18{,}2}{12} = 1{,}517 \\
s^2 &= \frac{1}{n-1}\sum_i (x_i-\bar{x})^2 = \frac{1}{n-1}\sum_i x_i^2 - \frac{n}{n-1}\bar{x}^2 = \frac{36{,}48}{11} - \frac{12}{11}1{,}517^2 = 0{,}807
\end{aligned}$$

para obtener el intervalo

$$3{,}816 \leq \frac{11\times 0{,}807}{\sigma^2} \leq 21{,}92 \Leftrightarrow \frac{11\times 0{,}807}{21{,}92} \leq \sigma^2 \leq \frac{11\times 0{,}807}{3{,}816} \Leftrightarrow 0{,}405 \leq \sigma^2 \leq 2{,}326$$

y para la desviación típica, tomando raíces cuadradas,

$$0{,}636 \leq \sigma \leq 1,525$$

b) Si empleamos únicamente los primeros 6 valores, cambian los cuantiles a emplear, de manera que ahora nos basaremos en los valores $\chi^2_{5,0{,}975} = 0{,}831$ y $\chi^2_{5,0{,}025} = 12{,}83$. Cambia también la estimación a emplear para la varianza

$$\begin{aligned}
\bar{x} &= \frac{1}{n}\sum_i x_i = \frac{11{,}1}{6} = 1{,}85 \\
s^2 &= \frac{1}{n-1}\sum_i (x_i-\bar{x})^2 = \frac{1}{n-1}\sum_i x_i^2 - \frac{n}{n-1}\bar{x}^2 = \frac{25{,}51}{5} - \frac{6}{5}1{,}85^2 = 0{,}995
\end{aligned}$$

El intervalo es ahora

$$0{,}831 \leq \frac{5\times 0{,}995}{\sigma^2} \leq 12{,}83 \Leftrightarrow \frac{5\times 0{,}995}{12{,}83} \leq \sigma^2 \leq \frac{5\times 0{,}995}{0{,}831} \Leftrightarrow 0{,}388 \leq \sigma^2 \leq 5{,}987$$

y para la desviación típica, tomando raíces cuadradas,

$$0{,}623 \leq \sigma \leq 2{,}447$$

18. En una empresa se pasa un control de calidad a los productos enviados por un proveedor. Para una de las características que se quiere controlar se selecciona una muestra de 20 productos y se obtiene una desviación típica muestral de 2,3. Suponemos que los valores de la característica siguen una distribución normal.

a) Calcula un intervalo de confianza al 90 % para la varianza de la población.

b) Supongamos que para un tamaño de muestra de 40 productos se tiene la misma desviación típica muestral. ¿Cuál sería ahora el nuevo intervalo de confianza, para el mismo nivel de confianza?

Solución.

a) Tenemos una muestra aleatoria simple de una población normal y nos piden estimar un intervalo de confianza para la varianza. El resultado básico para construir el intervalo de confianza pedido es

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$$

donde χ^2_{n-1} denota una distribución chi-cuadrado con $n-1$ grados de libertad.

Para el caso que nos indican, empezamos calculando el valor de los cuantiles de interés. Tenemos $\alpha = 0{,}1$ y por tanto $\alpha/2 = 0{,}05$ y $1-\alpha/2 = 0{,}95$, así como $n = 20$. De las tablas de la chi-cuadrado con 19 grados de libertad obtenemos los valores $\chi^2_{19,0,95} = 10{,}117$ y $\chi^2_{19,0,05} = 30{,}144$. El intervalo para la varianza tiene la forma

$$\chi^2_{n-1,1-\alpha/2} \leq \frac{(n-1)s^2}{\sigma^2} \leq \chi^2_{n-1,\alpha/2}$$

Sustituimos valores teniendo en cuenta que la cantidad muestral de interes es $s^2 = 2{,}3^2 = 5{,}29$, para obtener el intervalo

$$10{,}117 \leq \frac{19 \times 5{,}29}{\sigma^2} \leq 30{,}144 \Leftrightarrow \frac{19 \times 5{,}29}{30{,}144} \leq \sigma^2 \leq \frac{19 \times 5{,}29}{10{,}117} \Leftrightarrow 3{,}334 \leq \sigma^2 \leq 9{,}935$$

b) Si el tamaño de muestra es ahora $n = 40$, los valores de los cuantiles cambian a $\chi^2_{39,0,95} = 25{,}695$ y $\chi^2_{39,0,05} = 54{,}572$. El intervalo es ahora

$$25{,}695 \leq \frac{39 \times 5{,}29}{\sigma^2} \leq 54{,}572 \Leftrightarrow \frac{39 \times 5{,}29}{54{,}572} \leq \sigma^2 \leq \frac{39 \times 5{,}29}{25{,}695} \Leftrightarrow 3{,}781 \leq \sigma^2 \leq 8{,}029$$