MUESTREO Y ESTIMACIÓN ESTADÍSTICA

Muestreo. Métodos de muestreo

- Se llama **población** al conjunto de individuos que posee cierta característica.
- Una **muestra** es una parte de esa población.
- **Muestreo** es el proceso utilizado para la selección de los individuos de una muestra. Si el muestreo se hace bien es posible inferir las características o propiedades de la población, estudiándolas –mediante **encuesta**– en la muestra.

- Los métodos de muestreo más eficaces son los aleatorios. Entre ellos, los más utilizados son:

– **Muestreo aleatorio simple**. En este muestreo cada individuo debe tener la misma probabilidad de ser elegido; además, la selección de un individuo no debe afectar a la probabilidad de que sea seleccionado el siguiente. Por tanto, es un muestreo con reposición.
– **Muestreo sistemático**. Con este método, la muestra se elige siguiendo una secuencia preestablecida. Por ejemplo, encuestar a los individuos que ocupen la posición 7, 17, 27, ... y así sucesivamente, de una lista; o llamar por teléfono al primer individuo de cada página de la guía telefónica. Este muestreo requiere ordenar a los individuos de la población y elegir el criterio a seguir.
– **Muestreo proporcional o estratificado**. En este caso, hay que dividir la población en estratos de acuerdo con algún criterio; por ejemplo, por nivel de estudios, por sexo, por edades, etc. Determinado los estratos, en cada uno de ellos se encuestará a un número de individuos proporcional a la población correspondiente.

Distribución de la media muestral

Si en una población de media μ y desviación típica σ se toman muestras de tamaño *n*, la media de las medias muestrales se distribuye según una normal $N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right)$.

Aclaración: En una población se pueden tomar muchas muestras aleatorias de tamaño *n*. Si las llamamos M1, M2, M3, ..., cada una de ellas tendrá una media $\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$, ..., respectivamente (esas medias, $\bar{x}_i$, serán generalmente distintas). Los elementos de la distribución de medias muestrales son precisamente $\{\bar{x}_1, \bar{x}_2, \bar{x}_3, ...\}$. Pues bien, la media de esa distribución coincide con la media de la población: $\bar{X} = \dfrac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + ...}{\text{número de muestras posibles}} = \mu$;

mientras que la desviación típica de esas medias es $\sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}}$.

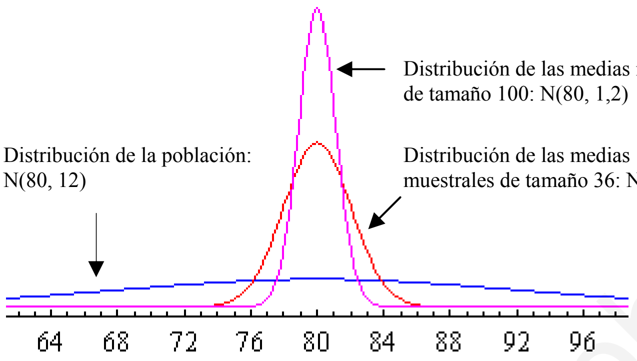
Esto es así para cualquier tipo de población de partida, sea normal o no, siempre y cuando el tamaño muestral $n > 30$.

- Este resultado permite aplicar a la distribución de medias muestrales todos los recursos de la distribución normal. Por ejemplo, cuantificar los resultados con ayuda de la tabla normal estándar: la $N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right)$ se tipifica o transforma en N(0, 1) haciendo el cambio $Z = \dfrac{\bar{x}_i - \mu}{\sigma / \sqrt{n}}$.

Ejemplos:

1. En una población N(80, 12), la media de las medias muestrales de tamaño 36 se distribuye según la normal $N\left(80,\ \dfrac{12}{\sqrt{36}}\right) = N(80,\ 2)$.

Asimismo, la media de las medias muestrales de tamaño 100 se distribuye según la normal $N\left(80,\ \dfrac{12}{\sqrt{100}}\right) = N(80,\ 1{,}2)$.



Como puede verse en la figura, si el tamaño muestral aumenta la campana se estrecha. Esto genera mayor precisión en las estimaciones.

2. Si se desea saber cuál es la probabilidad de que:
a) una muestra de tamaño 36 tenga una media superior a 83 o,
b) una muestra de tamaño 100 tenga una media comprendida entre 79 y 81,
hay que proceder como sigue:

a) La normal $N\left(80,\ \dfrac{12}{\sqrt{36}}\right) = N(80,\ 2)$ se tipifica haciendo el cambio $Z = \dfrac{\bar{x}_i - 80}{2}$.

Luego:

$$P(\bar{x} > 83) = P\left(Z > \frac{83-80}{2}\right) = P(Z > 1{,}5) = 1 - P(Z < 1{,}5) = 1 - 0{,}9332 = 0{,}0668$$

b) La normal $N\left(80,\ \dfrac{12}{\sqrt{100}}\right) = N(80,\ 1{,}2)$ se tipifica haciendo el cambio $Z = \dfrac{\bar{x}_i - 80}{1{,}2}$.

Luego:

$$P(79 < \bar{x} < 81) = P\left(\frac{79-80}{1{,}2} < Z < \frac{81-80}{1{,}2}\right) = P(-0{,}83 < Z < 0{,}83) =$$

$$= P(Z < 0{,}83) - P(Z < -0{,}83) = 0{,}7967 - (1 - 0{,}7967) = 0{,}5934$$

Distribución de la proporción muestral

La distribución de la proporción muestral de las muestras de tamaño n, obtenidas en una población con proporción p se ajusta a la normal $N\left(p,\ \sqrt{\dfrac{pq}{n}}\right)$, siendo $q = 1 - p$.

Ejemplos:
3. Supongamos que entre las personas de 35 años, el 80 % sabe conducir; esto es, la proporción de conductores es $p = 0,80$, y $q = 0,20$. En este caso, la proporción de las muestras de tamaño $n = 25$ se distribuye según la normal $N\left(0,8,\ \sqrt{\dfrac{0,8\cdot 0,2}{25}}\right) = N(0,8,\ 0,08)$.

Asimismo, la proporción de las muestras de tamaño 64 se distribuye según la normal $N\left(0,8,\ \sqrt{\dfrac{0,8\cdot 0,2}{64}}\right) = N(0,8,\ 0,05)$.

4. Si se desea saber cuál es la probabilidad de que:
a) en una muestra de tamaño 25 haya menos de un 70 % de conductores o,
b) en una muestra de tamaño 64 haya entre 50 y 60 conductores,
hay que proceder como sigue.

(a) Para una normal N(0,80, 0,08) −muestras de tamaño 25−

$$P(\hat{p} \le 0,70) = P\left(Z \le \frac{0,70 - 0,80}{0,08}\right) = P(Z \le -1,25) = 1 - 0,8944 = 0,1056$$

(b) En este caso, la proporción de la muestra se distribuye según la normal N(0,8, 0,05). Para que haya entre 50 y 60 conductores de un total de 64, la proporción debe estar entre $\dfrac{50}{64} \approx 0,78$ y $\dfrac{60}{64} \approx 0,94$, luego:

$$P(0,78 \le \hat{p} \le 0,94) = P\left(\frac{0,78 - 0,8}{0,05} \le Z \le \frac{0,94 - 0,80}{0,05}\right) = P(-0,4 \le Z \le 2,8) =$$

$$= P(Z \le 2,8) - P(Z \le -0,4) = 0,9974 - (1 - 0,6554) = 0,6528$$

Estimación estadística. Intervalo de confianza

La estimación estadística pretende deducir algún parámetro de la población partiendo del estudio de muestras. Aquí nos limitaremos a estimar la media o la proporción de una población a partir de una muestra.
La estimación se hace siempre en términos probabilísticos: dando un intervalo en el que esperamos que estará el parámetro (μ o p) poblacional e indicando la confianza de nuestra estimación. Ese intervalo se llama de **intervalo de confianza**. A la probabilidad de que tal estimación sea cierta se la llama **nivel de confianza**.

Intervalo de confianza para la media

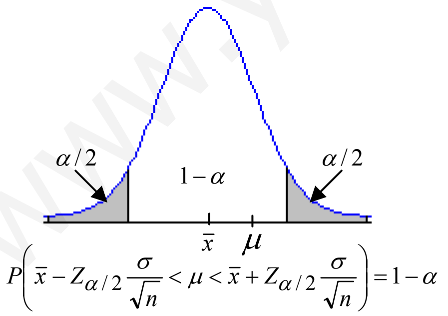
El intervalo de confianza de la media poblacional que se obtiene a partir de una muestra de media $\bar{x}$, es:

$$\left(\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

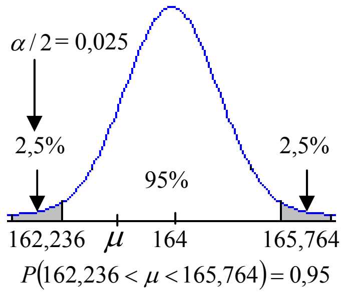
siendo σ la desviación típica poblacional, n el tamaño muestral y $Z_{\alpha/2}$ el valor correspondiente en la tabla normal para un nivel de confianza de $1 - \alpha$.

El nivel de confianza $1 - \alpha$, da la probabilidad que hay de que la media poblacional, μ, pertenezca a ese intervalo; es frecuente darla en porcentajes. Son equivalentes las expresiones "nivel de confianza $1 - \alpha$" y "**significación α**"; así, suele hablarse de obtener un intervalo con un nivel de confianza del 95 % o para una significación $\alpha = 0{,}05$. (Este valor de α indica la probabilidad que hay de errar en la estimación.)

Notas: 1. La estimación se considera válida si la población de partida es normal o si $n \geq 30$.
2. Si se desconoce la desviación típica poblacional, esta puede sustituirse por la "cuasidesviación" típica de la muestra, cuya fórmula es $s_{n-1} = \sqrt{\dfrac{\sum(x_i - \bar{x})^2}{n-1}}$. (En las calculadoras suele indicarse con la tecla σ_{n-1}).



Intervalo de confianza $1 - \alpha$



Intervalo de confianza al 95%

- Niveles de confianza usuales:

$1 - \alpha$ (%)	0,6826 (68,26%)	0,90 (90%)	0,95 (95%)	0,9544 (95,44%)	0,99 (99%)
α	0,3174	0,10	0,05	0,0456	0,01
$\alpha/2$	0,1587	0,05	0,025	0,0228	0,005
$Z_{\alpha/2}$	1	1,645	1,96	2	2,575

Ejemplos:

5. Se sabe que la desviación típica de la estatura de las chicas de 18 años es 9 cm. Si se toma una muestra aleatoria de tamaño 100, en la que se obtiene una estatura media de $\bar{x} = 164$, el intervalo de confianza para un nivel del 95% de la media de la población será:

$$\left(164 - 1{,}96\cdot\frac{9}{\sqrt{100}},\ 164 + 1{,}96\cdot\frac{9}{\sqrt{100}}\right) = (162{,}236,\ 165{,}764)$$

Esto significa que la media de estaturas para toda la población, μ, estará entre 162,236 cm y 165,764 cm, con una probabilidad del 0,95: $P(162{,}236 < \mu < 165{,}764) = 0{,}95$.
Puede suceder que la verdadera media sea mayor que 165,764, pero la probabilidad de eso es escasa: $P(\mu > 165{,}764) = 0{,}025$ ($\rightarrow$ $\alpha/2 = 0{,}025$)
También puede suceder que la verdadera media sea menor que 162,236, pero con una probabilidad igualmente pequeña: $P(\mu < 162{,}236) = 0{,}025$ ($\rightarrow$ $\alpha/2 = 0{,}025$).

6. El peso medio de una muestra elegida al azar de 196 perdices de una determinada variedad es de 320 g, y la desviación típica de 35 g. Calcula el intervalo de confianza aproximado para la media poblacional para un nivel de confianza del 90 %

Para $\bar{x} = 320$, $\sigma = 35$, $n = 196$ y, para el 90% de confianza, $Z_{\alpha/2} = 1{,}645$, el intervalo es,

$$\left(320 - 1{,}645\cdot\frac{35}{\sqrt{196}},\ 320 + 1{,}645\cdot\frac{35}{\sqrt{196}}\right) = (320 - 4{,}1,\ 320 + 4{,}1) = (315{,}9,\ 324{,}1)$$

- **Error admitido**

Al dar el intervalo de confianza para la media afirmamos, con una confianza de $1 - \alpha$, que

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

Esto significa que en nuestra estimación, admitimos un error máximo de $E = Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$.

Como puede observarse, este error depende σ, α y n: aumenta al hacerlo la desviación típica y la confianza; disminuye cuando aumenta el tamaño muestral.

Ejemplo:

7. Para los datos del ejemplo anterior: $\sigma = 9$ cm, $1 - \alpha = 0{,}95$ y $n = 100$, el error máximo es:

$$E = 1{,}96\cdot\frac{9}{\sqrt{100}} = 1{,}764 \text{ cm}$$

Por eso suele afirmarse que la media poblacional será $\mu = 164 \pm 1{,}764$. (Recuérdese que la media muestral era $\bar{x} = 164$.)
Si se varía el tamaño muestral y la significación (nivel de confianza), por ejemplo:

Para $n = 81$ y $\alpha = 0{,}1$ se tendrá: $\rightarrow$ $E = 1{,}645\cdot\dfrac{9}{\sqrt{81}} = 1{,}654$ cm.

Para $n = 36$ y $\alpha = 0{,}01$ $\rightarrow$ $E = 2{,}575\cdot\dfrac{9}{\sqrt{36}} = 3{,}8625$ cm.

- **Tamaño muestral**

Si fijamos el error máximo (= *emáx*) deseado, podemos deducir el mínimo tamaño muestral necesario para tal propósito, pues basta con despejar. Así:

$$E = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < emáx \;\Rightarrow\; Z_{\alpha/2}\frac{\sigma}{(emáx)} < \sqrt{n} \;\Rightarrow\; n > \left(Z_{\alpha/2}\frac{\sigma}{(emáx)}\right)^2$$

Ejemplo:

8. Si se quiere obtener una estimación de la media de estaturas de una población, que se distribuye con una desviación típica de σ = 9 cm, con un error máximo de 2 cm y con un nivel de confianza del 95%, ¿cuál será el tamaño muestral necesario para determinarla?

Se desea que $E = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < 2$. Como $Z_{0,025} = 1,96$ y $\sigma = 9 \Rightarrow n > \left(1,96\cdot\frac{9}{2}\right)^2 = 77,8$

Por tanto, habrá que encuestar a 78 personas de esa población, al menos.

Intervalo de confianza para la proporción

El intervalo de confianza de la proporción, con un nivel de confianza de $1-\alpha$, es:

$$\left(\hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}},\; \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right)$$

siendo $\hat{p}$ la proporción de la muestra, $\hat{q} = 1-\hat{p}$, n el tamaño muestral y $Z_{\alpha/2}$ el valor correspondiente en la tabla normal para una significación α.

Notas: 1. La estimación se considera válida si np y nq son mayores que 5, lo cual se consigue cuando n es suficientemente grande

2. Si se conociese p por estudios anteriores, se emplea la desviación típica $\sqrt{\frac{pq}{n}}$ en vez de $\sqrt{\frac{\hat{p}\hat{q}}{n}}$, siendo el intervalo de confianza: $\left(\hat{p} - Z_{\alpha/2}\sqrt{\frac{pq}{n}},\; \hat{p} + Z_{\alpha/2}\sqrt{\frac{pq}{n}}\right)$.

3. En estudios profesionales (encuestas) suele asignarse inicialmente $p = 0,5$ y $q = 0,5$.

Ejemplo:

9. Para conocer el grado de satisfacción de los usuarios en relación con el funcionamiento (bueno o malo) de los autobuses urbanos, se pregunta a 400 usuarios. De ellos, 184 dicen que el funcionamiento es bueno. Con esto, la proporción real de los usuarios satisfechos del servicio de autobuses, para un nivel de confianza del 99% (significación 0,01) pertenecerá al intervalo

$$\left(0,46 - 2,575\sqrt{\frac{0,46\cdot 0,54}{400}},\; 0,46 + 2,575\sqrt{\frac{0,46\cdot 0,54}{400}}\right) = (0,396,\ 0,524)$$

Nótese que: $\hat{p} = 184/400 = 0,46$, $\hat{q} = 1-\hat{p} = 0,54$, $n = 400$. y $Z_{\alpha/2} = Z_{0,005} = 2,575$.

Esto significa que el porcentaje de usuarios que está satisfecho del funcionamiento de los autobuses urbanos estará entre el 39,6 % y el 52,4 %. Una proporción real fuera de ese intervalo puede ser posible, pero la probabilidad de que realmente lo sea es de 0,01; por tanto, se asume un riesgo del 1 %.

- **Error admitido**

El error máximo admitido es: $E = Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$, que puede controlarse variando la significación α y el tamaño muestral *n*.

- **Tamaño muestral**

Si fijamos el error máximo (= *emáx*) deseado, podemos deducir el mínimo tamaño muestral *n*, pues:

$$E = Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} < emáx \quad \Rightarrow \quad n > \left(Z_{\alpha/2}\right)^2 \frac{\hat{p}\hat{q}}{(emáx)^2}$$

Ejemplos:

11. En el caso del ejemplo anterior: $\hat{p} =$ 0,46, 1 − α = 0,99 ($Z_{\alpha/2} = Z_{0,005}$ = 2,575) y *n* = 400, el error máximo es:

$$2{,}575\sqrt{\frac{0{,}46 \cdot 0{,}54}{400}} = 0{,}064$$

Luego la proporción de la población será *p* = 0,46 ± 0,064.

Si se varía el tamaño muestral y el nivel de confianza, por ejemplo:

a) Para *n* = 900 y 1 − α = 0,9544, se tendrá: → $E = 2\cdot\sqrt{\frac{0{,}46 \cdot 0{,}54}{900}}$ = 0,033 (± 3,3 %).

b) Para *n* = 100 y α = 0,05 → $E = 1{,}96\cdot\sqrt{\frac{0{,}46 \cdot 0{,}54}{100}}$ = 0,098 (± 9,8 %)

12. Si se quiere obtener una estimación de la proporción con un error máximo del 3 %, con un nivel de confianza del 98 %, el tamaño muestral necesario será:

$$n > (2{,}33)^2 \frac{0{,}46 \cdot 0{,}54}{0{,}03^2} = 1498{,}4$$

El 3 % ⇒ *E* = 0,03; 1 − α = 0,98 ⇒ α = 0,02, α/2 = 0,01 y $Z_{0,01}$ = 2,33

Por tanto, habrá que encuestar, al menos, a 1499 personas.

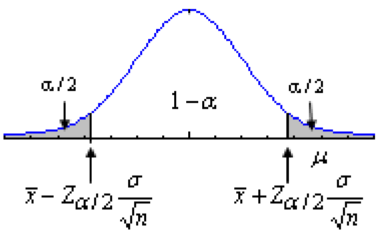
Problemas

1. El número de días de permanencia de los enfermos en un hospital sigue una ley Normal de media μ días y desviación típica 3 días.
a) Determina un intervalo de confianza para estimar μ, a un nivel del 97%, con una muestra aleatoria de 100 enfermos cuya media es 8,1 días.
b) ¿Qué tamaño mínimo debe tener una muestra aleatoria para poder estimar μ con un error máximo de 1 día y un nivel de confianza del 92%

Solución:
a) El intervalo de confianza de la media poblacional, para las muestras de tamaño muestral n de media $\bar{x}$ y desviación típica σ es:

$$\left(\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

siendo $Z_{\alpha/2}$ el valor correspondiente en la tabla normal para una confianza de $1 - \alpha$.



Para $\bar{x} = 8{,}1$, $\sigma = 3$, $n = 100$ y, para el 97 % de confianza, $Z_{\alpha/2} = 2{,}17$, se tiene:

$$\left(8{,}1 - 2{,}17\cdot\frac{3}{\sqrt{100}},\ 8{,}1 + 2{,}17\cdot\frac{3}{\sqrt{100}}\right) = (8{,}1 - 0{,}651,\ 8{,}1 + 0{,}651) = (7{,}449,\ 8{,}751)$$

b) El error admitido, E, viene dado por $E = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$.

Para una confianza del 92%, $Z_{\alpha/2} = 1{,}75$. Como $\sigma = 3$ y se desea que $E \le 1$, debe cumplirse que:

$$1{,}75\cdot\frac{3}{\sqrt{n}} \le 1 \ \Rightarrow\ \sqrt{n} \ge 5{,}25 \ \Rightarrow\ n \ge 27{,}56 \to n \ge 28$$

El tamaño muestral mínimo debe ser de 28 enfermos.

2. Los gastos semanales en los hogares españoles siguen una distribución normal con desviación típica 30 euros. A partir de una muestra aleatoria de tamaño 25 se ha obtenido una media muestral de 175 euros.
a) Halla el intervalo de confianza del 95% para la media del gasto semanal.
b) ¿Qué tamaño mínimo debe tener la muestra que permita estimar la media con un nivel de confianza del 99%, con un error que sea la décima parte del obtenido en el apartado anterior?

Solución:
a) Para $\bar{x} = 175$, $\sigma = 30$, $n = 25$ y, para el 95 % de confianza, $Z_{\alpha/2} = 1{,}96$, el intervalo pedido es,

$$\left(175 - 1{,}96\cdot\frac{30}{\sqrt{25}},\ 175 + 1{,}96\cdot\frac{30}{\sqrt{25}}\right) = (175 - 11{,}76,\ 175 + 11{,}76) = (163{,}24,\ 186{,}76)$$

b) El error admitido E, viene dado por $E = Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$.

Para un nivel de confianza del 99%, $Z_{\alpha/2}$ = 2,575, con σ = 30 y E < 1,176 (la décima parte de 11,76) se tendrá:

$$2{,}575\cdot\frac{30}{\sqrt{n}} < 1{,}176 \;\Rightarrow\; \sqrt{n} > 65{,}69 \;\Rightarrow\; n > 4315$$

3. Una encuesta, realizada sobre una muestra de los jóvenes de una ciudad, para determinar el gasto mensual medio (expresado en euros) en teléfono móvil, concluyó con el intervalo de confianza al 95%: (10,794, 13,206).
a) ¿Cuál es el gasto mensual medio muestral?
b) ¿Cuál es el correspondiente intervalo de confianza al 99%?
c) Si, aproximando con cuatro cifras decimales, la desviación típica del gasto mensual es de 7,9989 euros, ¿cuál es el tamaño de la muestra encuestada?

Solución:
a) El intervalo de confianza $\left(\bar{x} - Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}},\ \bar{x} + Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right)$ es simétrico respecto de la media, lo que significa que la media es el punto medio de dicho intervalo. Por tanto:

$$\bar{x} = \frac{10{,}794 + 13{,}206}{2} = 12$$

El gasto medio mensual en teléfono móvil es de 12 €.

b) Como para una confianza del 95%, $Z_{\alpha/2}$ = 1,96, se tiene que:

$$\left(12 - 1{,}96\frac{\sigma}{\sqrt{n}},\ 12 + 1{,}96\frac{\sigma}{\sqrt{n}}\right) = (10{,}974,\ \ 13{,}206) \Rightarrow$$

$$\Rightarrow\ 1{,}96\frac{\sigma}{\sqrt{n}} = 1{,}206 \Rightarrow \frac{\sigma}{\sqrt{n}} = 0{,}6153.$$

Por tanto, como para una confianza del 99%, $Z_{\alpha/2}$ = 2,575, el intervalo de confianza correspondiente será:

$$\left(12 - 2{,}575\cdot\frac{\sigma}{\sqrt{n}},\ 12 + 2{,}575\frac{\sigma}{\sqrt{n}}\right) = (12 - 2{,}575\cdot 0{,}6153,\ \ 12 + 2{,}575\cdot 0{,}6153) =$$

$$= (10{,}4156,\ \ 13{,}5844)$$

c) Como se ha dicho, $\dfrac{\sigma}{\sqrt{n}} = 0{,}6153$; luego si σ =7,9989 ⇒ $\dfrac{7{,}9989}{\sqrt{n}} = 0{,}6153$ ⇒

$$\Rightarrow \sqrt{n} = 13 \Rightarrow n = 169.$$

El tamaño de la muestra ha sido de 169 jóvenes.

4. Con un nivel de confianza igual a 0,95, a partir de un estudio muestral, el intervalo de confianza de la proporción de habitantes de una comunidad que tienen ordenador portátil es (0,1804, 0,2196).
a) ¿Cuál es la proporción muestral de habitantes de esa comunidad que tienen ordenador portátil? ¿Cuál es el tamaño de la muestra?
b) ¿Cuál debería ser el tamaño muestral para estimar la citada proporción, con una confianza del 95%, con un error máximo de 0,01?

Solución:
a) En este caso, el intervalo de confianza para la proporción de la población, será:
$\left(\hat{p} - Z_{\alpha/2}\sqrt{\dfrac{0,5\cdot 0,5}{n}},\ \hat{p} + Z_{\alpha/2}\sqrt{\dfrac{0,5\cdot 0,5}{n}}\right)$, siendo $\hat{p}$ la proporción de la muestra.
Como puede observarse, el intervalo es simétrico; por tanto, la proporción muestral es el punto medio del intervalo.
Luego, si el intervalo es (0,1804, 0,2196) se tendrá que $\hat{p} = \dfrac{0,1804 + 0,2196}{2} = 0,20$.
El 20% de la población tiene ordenador portátil.

Como $\hat{p} + Z_{\alpha/2}\sqrt{\dfrac{0,5\cdot 0,5}{n}} = 0,2196$ se tendrá:

$$0,20 + 1,96\sqrt{\frac{0,5\cdot 0,5}{n}} = 0,2196 \Rightarrow \sqrt{\frac{0,5\cdot 0,5}{n}} = 0,01 \Rightarrow n = \frac{0,25}{0,0001} = 2500$$

El tamaño de la muestra ha sido de 2500 personas.

b) El error máximo vale $E = Z_{\alpha/2}\sqrt{\dfrac{0,5\cdot 0,5}{n}} \Leftrightarrow n = \left(Z_{\alpha/2}\right)^2 \dfrac{0,5\cdot 0,5}{E^2}$.
Si se desea que $E \le 0,01$ y que $Z_{\alpha/2} = 1,96$ (correspondiente a un nivel de confianza del 95%), entonces, $n \ge 1,96^2 \dfrac{0,5\cdot 0,5}{0,01^2} = 9604$
El tamaño muestral debería ser de 9604 personas.

5. Para estimar la proporción de las viviendas de una determinada ciudad que tienen aire acondicionado se quiere utilizar una muestra aleatoria de tamaño *n*. Calcule el valor mínimo de *n* para que, con un nivel de confianza del 97%, el error en la estimación sea más pequeño que 0,05. (Como se desconoce la proporción, se debe tomar el caso más desfavorable, que será 0,5).

Solución:
El error admitido es $E = Z_{\alpha/2}\sqrt{\dfrac{pq}{n}} \Rightarrow n = (Z_{\alpha/2})^2 \dfrac{pq}{E^2}$
Para una confianza del 97 %, $Z_{\alpha/2} = 2,17$. Como se desea que $E < 0,05$ se tendrá que

$$0,05 > 2,17\sqrt{\frac{0,5\cdot 0,5}{n}} \Rightarrow n > 2,17^2 \frac{0,5\cdot 0,5}{0,05^2} \Rightarrow n > 3470,89.$$

Luego, el tamaño muestral mínimo debe ser 471.