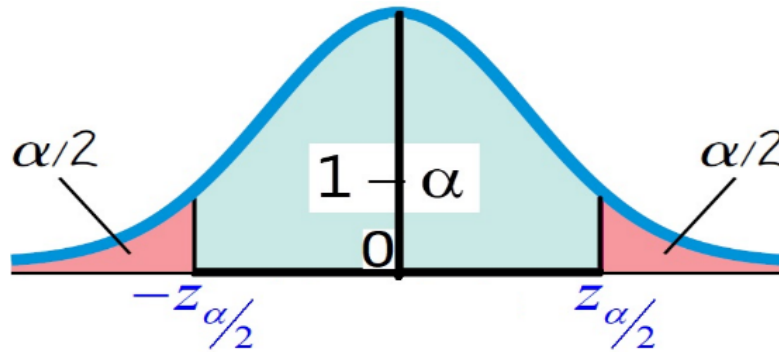


Intervalo de confianza para la media poblacional

Sea X es una v.a. tal que $X \rightarrow N(\mu, \sigma)$ de la que desconocemos la media poblacional μ .

Un intervalo de confianza para μ es un intervalo I en el que $p(\mu \in I) = n_c$.

El valor n_c se llama nivel de confianza y $n_c = 1 - \alpha$ (A α se le llama "nivel de significación")



$$z_{\alpha/2} \text{ cumple } p(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha = n_c \Rightarrow p(Z < z_{\alpha/2}) = 1 - \alpha + \frac{\alpha}{2} = \frac{2 - \alpha}{2} = \frac{1 + 1 - \alpha}{2}$$

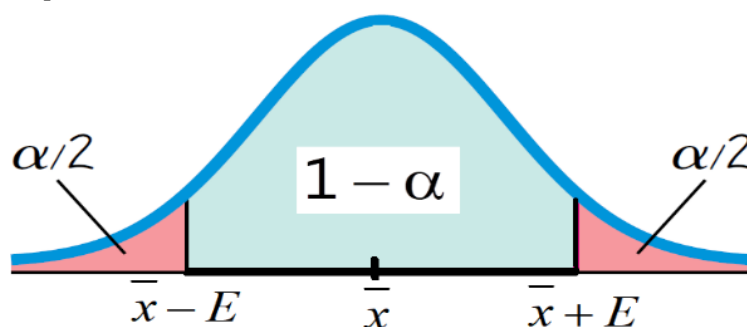
$$\text{Luego, } p(Z < z_{\alpha/2}) = \frac{1 + n_c}{2}$$

Para estimar cuánto vale μ , tomamos una muestra de tamaño n cuya media es $\bar{x}$

Se puede demostrar que el intervalo de confianza para la media es

$$I = (\bar{x} - E, \bar{x} + E) \quad \text{siendo} \quad E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

E se llama error de estimación o error máximo admisible y es el mayor error que podemos cometer cuando tomamos $\bar{x}$ como aproximación de μ

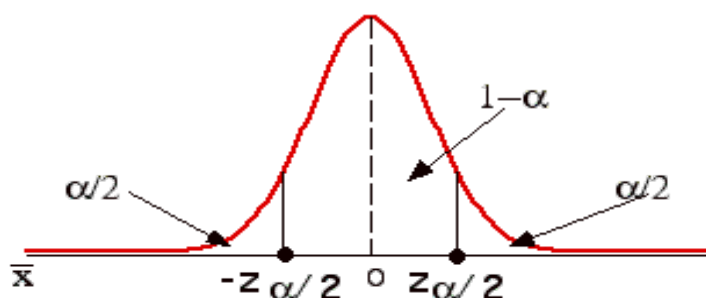


$$n_c = p(\bar{x} - E < \mu < \bar{x} + E)$$

$$\bar{x} - E < \mu < \bar{x} + E \xrightarrow{(-1)} -\bar{x} + E > -\mu > -\bar{x} - E \Leftrightarrow -\bar{x} - E < -\mu < -\bar{x} + E$$

Según el teorema central del límite, $\bar{x} \rightarrow N(\mu, \frac{\sigma}{\sqrt{n}})$, luego $Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \rightarrow N(0,1)$

Tipificando (sumamos $\bar{x}$ y dividimos entre $\sigma/\sqrt{n}$): $\frac{-E}{\sigma/\sqrt{n}} < \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} < \frac{E}{\sigma/\sqrt{n}} \Leftrightarrow \frac{-E}{\sigma/\sqrt{n}} < Z < \frac{E}{\sigma/\sqrt{n}}$



$$n_c = 1 - \alpha \quad (\alpha \text{ se llama nivel de significación})$$

$$\text{Como } n_c = p(-\bar{x} - E < \mu < \bar{x} + E) = p\left(\frac{-E}{\sigma/\sqrt{n}} < Z < \frac{E}{\sigma/\sqrt{n}}\right) = p(-z_{\alpha/2} < Z < z_{\alpha/2})$$

$$z_{\alpha/2} = \frac{E}{\sigma/\sqrt{n}}, \text{ despejando } \boxed{E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}}$$

$$\text{Por otra parte, } z_{\alpha/2} \text{ cumple } n_c = p(-z_{\alpha/2} < Z < z_{\alpha/2}) = 2\phi(z_{\alpha/2}) - 1_c \rightarrow \boxed{\phi(z_{\alpha/2}) = \frac{1 + n_c}{2}}$$

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

Un fabricante de electrodomésticos sabe que la duración media de éstos sigue una normal de media 100 meses y desviación 12 meses. Determinar el mínimo tamaño muestral que garantiza, con una probabilidad de 0,98, que la duración media de los electrodomésticos en dicha muestra se encuentra entre 90 y 110 meses.

El consumo de un cierto producto sigue una distribución Normal de varianza 300. A partir de una muestra de tamaño 25 se ha obtenido una media muestral igual a 180. Halla un intervalo de confianza al 95% para la media del consumo.

El peso de los niños varones a las 10 semanas de vida se distribuye según una normal con desviación típica 87g. ¿Cuántos datos son suficientes para estimar, con una confianza del 95%, el peso medio de esa población con un error no superior a 15 g?

La talla de los individuos de una población sigue una normal de media desconocida y desviación 8 cm. Se han determinado las tallas de 25 individuos, encontrándose una media muestral de 168 cm. Obtener un intervalo de confianza al 95% de la media de la población.

Los estudiantes de las Universidades de cierto país dedican al estudio un número de horas semanales que sigue una distribución Normal de media desconocida y desviación típica 7 horas. Calcula cuántos se deberán seleccionar al azar para determinar la media con una precisión inferior a una hora y con una confianza del 90%.

La edad, en años, de las personas de una población se distribuye Normalmente con media desconocida y varianza 4. ¿Cuál es el tamaño muestral a partir del cual la edad media de la población difiere de la media muestral en menos de 1 año con una confianza del 99%?

Los estudiantes de Bachillerato de una cierta población duermen un número de horas diarias que sigue una Normal de media desconocida y desviación 3. A partir de una muestra de 30 alumnos se ha obtenido una media muestral de 7 horas. Halla un intervalo de confianza al 96% para la media de horas de sueño.

Las ventas mensuales de una fábrica siguen una Normal con desviación típica 90000 euros. En un estudio de los últimos 9 meses se ha encontrado un intervalo de confianza para la media de ventas, cuyos extremos son 466300 y 583900 euros.

¿Cuál ha sido la media de ventas en estos 9 meses?

¿Cuál es el nivel de confianza de ese intervalo?

En una muestra de 225 individuos se ha obtenido una media de 16.5 años. La desviación típica de la población de la que procede la muestra es de 0.7 años. Calcule:

Intervalo al 95% para la media de la población.

¿Qué error se comete en la estimación?

El peso de los adultos de una determinada especie de peces sigue una ley Normal de desviación típica 112 g. ¿Cuál es el tamaño mínimo de la muestra de peces que debería tomarse para obtener, con una confianza del 95%, la media de la población con un error menor de 20 g?

Intervalo de confianza para la proporción poblacional

Sea X es una v.a. tal que $X \rightarrow N(\mu, \sigma)$ de la que desconocemos la proporción, p , de individuos de la población que tiene una cierta característica. Un intervalo de confianza para p es un intervalo I en el que la probabilidad de que p pertenezca a I es n_c . El valor n_c se llama nivel de confianza

y $n_c = 1 - \alpha$ (α se le llama “nivel de significación”)

Razonando de forma análoga que en los intervalos para la media:
$$p(Z < z_{\alpha/2}) = \frac{1 + n_c}{2}$$

Para estimar cuánto vale p , tomamos una muestra de tamaño n cuya proporción de individuos que cumple dicha característica es $\bar{p}$

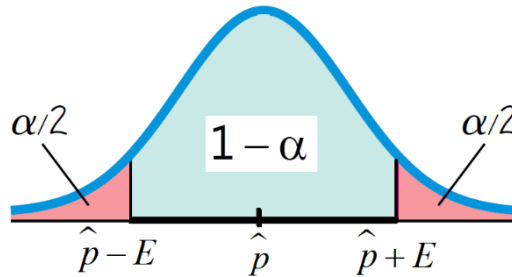
Se puede demostrar que el intervalo de confianza para la proporción es

$$I = (\bar{p} - E, \bar{p} + E)$$

siendo

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

E se llama error de estimación o error máximo admisible y es el mayor error que podemos cometer cuando tomamos $\hat{p}$ como aproximación de p



$$n_c = p(\hat{p} - E < p < \hat{p} + E)$$

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

$$\phi(z_{\alpha/2}) = \frac{1 + n_c}{2}$$

Razonando de manera similar al caso anterior,

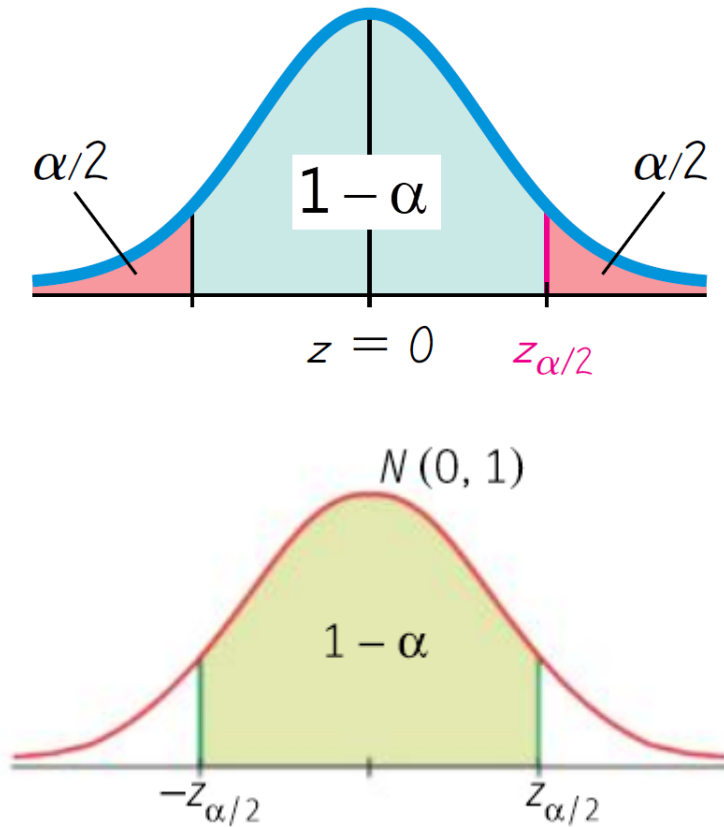
Parámetro a estimar	Intervalo de confianza	Error máximo admisible	Interpretación gráfica
Media	$I = (\bar{x} - E, \bar{x} + E)$ El punto medio del intervalo es $\bar{x}$ y su amplitud es 2E	$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$	
Proporción	$I = (\hat{p} - E, \hat{p} + E)$ El punto medio del intervalo es $\hat{p}$ y la amplitud es 2E	$E = z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	

Sea X es una v.a. continua de media μ y desviación típica σ y sea $\bar{X}$ la v.a. que representa la media de las muestras de tamaño n .

Por el teorema del limite central, si $n \geq 30$ ó $X \rightarrow N(\mu, \sigma)$, la v.a. $\bar{X} \rightarrow N(\mu, \frac{\sigma}{\sqrt{n}}) \Rightarrow Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \rightarrow N(0,1)$

Sea I el intervalo de confianza a nivel de significación α (el nivel de confianza es, entonces, $n_c = 1 - \alpha$)

Sea $z_{\alpha/2}$ el valor que cumple: $p(Z < z_{\alpha/2}) = 1 - \frac{\alpha}{2} \Rightarrow p(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha = n_c$



Según la definición $n_c = p(\mu \in I)$. Por otra parte, $n_c = 1 - \alpha = p(-z_{\alpha/2} < Z < z_{\alpha/2}) \xrightarrow{\cdot \frac{\sigma}{\sqrt{n}}} p(-z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}})$
 $\xrightarrow{-\bar{X}} p(-\bar{X} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < -\mu < -\bar{X} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}) \xrightarrow{(-1)} p(\bar{X} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} > \mu > \bar{X} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}) = p\left(\mu \in \left(\bar{X} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right)\right)$

Por tanto, el intervalo de confianza es $I = \left(\bar{X} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right)$

En la práctica se toma sólo una muestra, luego $\bar{X} = \bar{x}$, El intervalo de confianza es $I = (\bar{x} - E, \bar{x} + E)$, siendo el error $E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

Sea X es una v.a. y sea p la v.a. que representa la proporción de individuos de la población que tiene una determinada característica

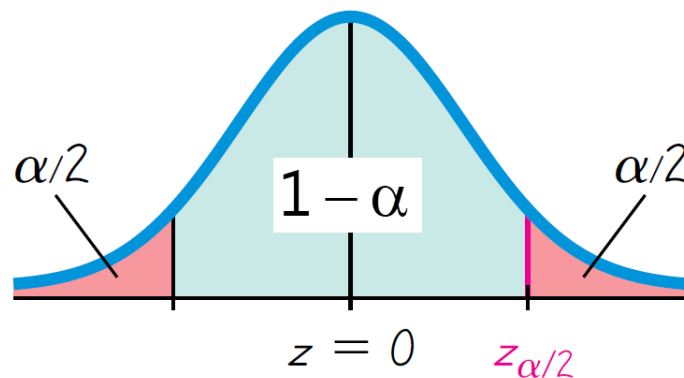
Si tomamos una muestra de tamaño n cuya proporción es $\hat{p}$ con $n\hat{p} \geq 5$ y $n(1-\hat{p}) \geq 5$,

$\hat{p}$ tiene una distribución $N(n\hat{p}, \sqrt{n\hat{p}(1-\hat{p})})$

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \rightarrow N(0,1)$$

Sea I el intervalo de confianza a nivel de significación α (el nivel de confianza es, entonces, $n_c = 1 - \alpha$)

Sea $z_{\alpha/2}$ el valor que cumple: $p(Z < z_{\alpha/2}) = 1 - \frac{\alpha}{2} \Rightarrow p(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha = n_c$



Parámetro a estimar	Situación	Intervalo de confianza	Error máximo admisible	Interpretación gráfica
Media	Sea X es una v.a. que sigue una distribución normal de la que desconocemos la media poblacional μ y σ es conocida. Queremos estimar cuánto vale μ . Tomamos una muestra de tamaño n cuya media es $\bar{x}$	$I = (\bar{x} - E, \bar{x} + E)$ El punto medio del intervalo es $\bar{x}$ y su amplitud es $2E$	$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$	 $p(\mu \in I) = 1 - \alpha$
Proporción	Si X una v.a. que sigue una distribución normal de la que desconocemos la proporción de individuos, p , de la población que tiene una cierta característica y queremos estimar su valor. Tomamos una muestra de tamaño n cuya proporción de individuos que cumple dicha característica es $\hat{p}$	$I = (\hat{p} - E, \hat{p} + E)$ El punto medio del intervalo es $\hat{p}$ y la amplitud es $2E$	$E = z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	 $p(p \in I) = 1 - \alpha$

$z_{\alpha/2}$ es el valor que cumple $p(Z < z_{\alpha/2}) = \frac{1+n_c}{2}$, siendo $n_c = 1 - \alpha$, el nivel de confianza

Propiedades de los intervalos de confianza

- El punto medio del intervalo de confianza para la media es $\bar{x}$ y para la proporción es $\bar{p}$

$$\text{punto medio} = \frac{\text{extremo superior} + \text{extremo inferior}}{2} \begin{cases} \text{para la media: punto medio} = \frac{(\bar{x} + E) + (\bar{x} - E)}{2} = \bar{x} \\ \text{para la proporción: punto medio} = \frac{(\bar{p} + E) + (\bar{p} - E)}{2} = \bar{p} \end{cases}$$
- La amplitud del intervalo de confianza, es $A = 2E$ y, por tanto, el error es $E = \frac{A}{2}$

$$A = \text{amplitud} = \text{extremo superior} - \text{extremo inferior} \begin{cases} \text{para la media: } A = (\bar{x} + E) - (\bar{x} - E) = 2E \\ \text{para la proporción: } A = (\bar{p} + E) - (\bar{p} - E) = 2E \end{cases}$$
- Cuanto mayor es el tamaño de la muestra, n , menor es el error E y por tanto más “estrecho” es el intervalo de confianza
- Cuanto mayor es el nivel de confianza, n_c , mayor es el error E y por tanto más amplio es el intervalo de confianza

Una empresa editorial está interesada en conocer la proporción de lectores habituales entre los adolescentes de 12 a 18 años, para ello se traslada a un centro de educación secundaria donde pregunta a 400 alumnos si leen habitualmente, de ellos 250 responden afirmativamente. Halla el intervalo en el que se estima que está la proporción de lectores con un nivel de confianza del 96%.

Tomada, al azar, una muestra de 120 estudiantes de una Universidad, se encontró que 54 de ellos hablaban inglés. Halle, con un nivel de confianza del 90%, un intervalo de confianza para estimar la proporción de estudiantes que hablan el idioma inglés entre los estudiantes de esa Universidad. Con los datos del ejercicio anterior, se pretende repetir la experiencia para conseguir que la cota del error que se comete al estimar, por un intervalo de confianza, la proporción de alumnos que hablan inglés en esa Universidad no sea superior a 0,05, con un nivel de confianza del 99%. ¿Cuántos alumnos tendríamos que tomar, como mínimo, en la muestra?

Se toma una muestra de 250 casas de una población de edificios antiguos para estimar la proporción de casas de este tipo. Supongamos que el 30% de todos los edificios son antiguos. Hallar la probabilidad de que la proporción de edificios antiguos esté entre 0,25 y 0,35.

Tomada una muestra aleatoria de 300 personas mayores de edad de una gran ciudad, se obtuvo que 105 habían votado a un determinado partido X. Halle, con un nivel de confianza del 90%, un intervalo de confianza que permita estimar la proporción de votantes del partido X en esa ciudad.

En una muestra aleatoria de 300 personas mayores de edad de una gran ciudad se encontró que 105 leían un determinado periódico X. A la vista de esos datos se pretende seleccionar una nueva muestra para conseguir una cota de error de 3 centésimas como máximo, con un nivel de confianza del 95%, para la estimación de la proporción de lectores de ese periódico por medio de un intervalo de confianza. Deduzca el número de individuos de la población que, como mínimo, debe tener la muestra.

Se desea estimar, por medio de un intervalo de confianza, la proporción p de individuos daltónicos de una población a través del porcentaje observado en una muestra aleatoria de individuos de tamaño n . Si el porcentaje de individuos daltónicos en una muestra aleatoria es igual al 30%, calcule el valor mínimo de n para que, con un nivel de confianza del 95%, el error que se cometa en la estimación sea inferior a 0,031.

Para estimar, por medio de un intervalo de confianza, la proporción p de individuos miopes de una población, se ha tomado una muestra de 80 individuos con la que se ha obtenido un porcentaje de individuos miopes del 35%. Determine, usando un nivel de confianza del 99%, el correspondiente intervalo de confianza para la proporción de miopes de toda la población.

En una encuesta realizada a 500 mujeres adultas de una población se encontró que 300 de ellas están casadas actualmente. Construya con estos datos un intervalo de confianza, con un nivel del 90%, para la proporción de mujeres adultas actualmente casadas en esa población.

Una muestra aleatoria de automóviles tomada en una zona turística ha permitido obtener un intervalo de confianza, al nivel del 95%, para estimar de la proporción de matrículas extranjeras de esa zona, siendo sus extremos 0,232 y 0,368. Determine el valor de la proporción estimada a través de esa muestra y una cota del error de estimación a este nivel de confianza.

Utilizando el mismo nivel de confianza, ¿cuál sería la cota de error, si esa misma proporción se hubiera observado en una muestra de 696 matrículas?

Para conocer la audiencia de uno de sus programas (proporción de televidentes que lo prefieren), una cadena de TV ha encuestado a 1000 personas elegidas al azar obteniendo una proporción muestral del 33% de personas favorables a ese programa. Calcule una cota del error de estimación, por medio de un intervalo de confianza, con un nivel del 92%.

Se va a tomar una muestra aleatoria de 600 recién nacidos en este año en una ciudad para estimar la proporción de varones entre los recién nacidos de esa ciudad, mediante un intervalo de confianza con un nivel del 95%. ¿Cuál será el error de estimación a ese nivel si se observan 234 varones en la muestra?

Para estimar la proporción de familias con un solo hijo en una ciudad, se ha tomado una muestra de familias al azar, de las cuales el 30% tiene un solo hijo. ¿Cuál es el mínimo tamaño muestral necesario para que, con esos datos, un intervalo de confianza de esa proporción a un nivel del 95% tenga una cota de error de 0,06, como máximo?

Una cadena de TV quiere saber si la audiencia de uno de sus programas sigue manteniéndose en el 25% de los espectadores.

¿Cuántos espectadores se deberían encuestar al azar, como mínimo, para tener un nivel de confianza del 90% de que el error en la estimación de la proporción actual sea igual o inferior a 0,03?

Calcule el tamaño mínimo de una muestra aleatoria de jóvenes entre 18 y 25 años para tener una confianza del 95% de que el error que se cometerá al estimar la proporción de fumadores entre esas edades no sea superior a 0,05, sabiendo que en una encuesta previa se ha encontrado un 32% de fumadores entre estos jóvenes.

Se va a tomar una muestra aleatoria de 600 recién nacidos en este año en una ciudad para estimar la proporción de varones entre los recién nacidos de esa ciudad, mediante un intervalo de confianza con un nivel del 95%. ¿Con qué proporción estimada será máxima la amplitud de ese intervalo? ¿Cuál es la amplitud máxima?

Para estimar la proporción de consumidores que prefieren un determinado refresco, por medio de un intervalo de confianza, se ha tomado una muestra al azar de 1075 consumidores, entre los que se han

encontrado 516 que lo prefieren. Determine una cota del error cometido para esa estimación a un nivel de confianza del 95%.

En una muestra aleatoria de 600 coches de una ciudad, 120 son de color blanco. Construya un intervalo de confianza de la proporción de coches de color blanco con un nivel de confianza del 98%.

Se estima la proporción de varones adultos, residentes en una población, con obesidad severa ($30 < \text{IMC} \leq 40$), mediante una muestra aleatoria de tamaño 500. Se obtiene una estimación de varones con obesidad severa del 18%. Utilizando un nivel de confianza del 98%, ¿cuál es el error máximo que se cometerá al estimar, por medio de un intervalo de confianza, esa proporción?

Se desea estimar la proporción de adultos que leen un determinado diario local por medio de un intervalo de confianza. Obtenga el tamaño mínimo de la muestra que garantice, aún en la situación más desfavorable, un error de la estimación inferior a 0.03, con un nivel de confianza del 95%.

Se estima, por un intervalo de confianza, la proporción de hogares con conexión a Internet utilizando una muestra aleatoria y con un nivel de confianza del 96%. Se obtiene así, una proporción estimada del 28%, con un error máximo del 6%. ¿Cuál es el tamaño mínimo de la muestra utilizada?

Mediante una muestra aleatoria de tamaño 400 se estima la proporción de residentes en Sevilla que tienen intención de asistir a un partido de fútbol entre el Betis y el C.F. Sevilla. Si para un nivel de confianza del 95% resulta un error máximo en la estimación del 3%. Obtenga el valor de la estimación, sabiendo que es inferior a 0,25.

En el diario "CÓRDOBA" del día 20 de Enero de 2.004 se publicó el resultado de un sondeo sobre intención de voto en las elecciones al Parlamento Andaluz del 14 de marzo de 2.004. Según la ficha técnica de la encuesta, el tamaño de la muestra fue de 5000 individuos, el nivel de confianza utilizado del 95%, y el error máximo de la estimación de los que no tienen decidido el voto del 1%. En la página 2, del mencionado diario, se estima que el 13,3% de los andaluces no tienen decidido el voto.

Analice la coherencia del resultado de la estimación con la ficha técnica de la encuesta, si se utiliza un muestreo aleatorio simple.

En una investigación de mercado se pregunta a 600 personas sobre el interés en consumir un determinado producto, si éste se comercializara en la ciudad. De ellas 55% manifiestan su intención de consumirlo. Con posterioridad a la encuesta, el fabricante del producto comercial exige que el error de la estimación sea inferior al 3%, con una confianza del 98%. ¿Cumple la investigación los requisitos exigidos por el fabricante?

En caso negativo, ¿cuál es el valor mínimo del tamaño de la muestra para cumplir con las exigencias del fabricante?

Tomada al azar una muestra de 90 alumnos de un Instituto se encontró que un tercio habla inglés. Halle, con un nivel de confianza del 97%, un intervalo de confianza para estimar la proporción de alumnos de ese Instituto que habla inglés.