Unidad 6. INTERVALOS DE CONFIANZA

ESTIMACIÓN POR INTERVALOS

La estimación puntual presenta un problema evidente. Si damos un único punto como estimación puntual del parámetro θ, esta estimación difícilmente se aproximará al verdadero valor de θ. Recordemos que nuestro objetivo es conseguir que la estimación se aproxime lo más posible al verdadero valor del parámetro. Para llevar a cabo esta idea vamos a construir un intervalo (a, b) de forma que la probabilidad de que el verdadero valor del parámetro esté dentro de (a, b) sea la que nosotros fijemos de antemano. A este tipo de intervalos los llamamos **intervalos de confianza**.

Definición de intervalos de confianza.-

Se llama intervalo de confianza para el parámetro θ a un nivel de confianza $1 - \alpha$, con $0 < \alpha < 1$, a un intervalo de números reales (a , b) de forma que la probabilidad de que el parámetro θ esté en el intervalo (a , b) sea $1 - \alpha$: $P(a \leq \theta \leq b) = 1 - \alpha$

Al parámetro α le llamamos nivel de significación

La idea que encierra esta definición es la siguiente: elegimos un nivel de confianza $1 - \alpha$ próximo a 1 (por ejemplo tomamos $1 - \alpha$ y exigimos que el intervalo de confianza contenga el verdadero valor del parámetro en el 95% de los casos posible, es decir, el 95% de las estimaciones son buenas y el 5% restantes malas.

Construcción de algunos intervalos de confianza.-

1.- Intervalo de confianza para la media de una población normal con ***varianza conocida***

Supongamos que tenemos una población X que sigue una distribución normal N(μ , σ) donde la media es desconocida y la varianza conocida.
Si : $x_1, ...,x_n$ es una muestra aleatoria simple de tamaño n de la población X, un intervalo de confianza para la media μ a un nivel de confianza de $1 - \alpha$ viene dado por la siguiente expresión :

$$I = \left(\bar{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} , \bar{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right)$$

donde : $\overline{X}$ es la media muestral
σ es la desviación típica poblacional
n es el tamaño de la muestra
$Z_{\frac{\alpha}{2}}$ es un valor que obtendremos de la tabla de la normal N (0, 1); este valor que toma la abscisa de forma que el área que queda a su derecha es igual a α/2.

3.- Intervalo de confianza para una ***proporción***

Consideremos una población X que se distribuye según una binomial de parámetro p desconocido.

Unidad 6. INTERVALOS DE CONFIANZA

Si de la población extraemos una muestra aleatoria simple $x_1, \ldots\ldots, x_n$ de tamaño n > 30, un intervalo de confianza para el parámetro p a un nivel de confianza 1 - α viene dado por:

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

donde: $\hat{p} = \dfrac{n^{o}.de.veces.que.ocurre.el.suceso.en.la.muestra}{n}$ es la frecuencia relativa muestral.

TAMAÑO DE LA MUESTRA

1.-En la estimación de la media (conocida la varianza poblacional)

Un intervalo de confianza para la media de una población es: $I = \left(\bar{X} - Z_{\frac{\alpha}{2}} \dfrac{\sigma}{\sqrt{n}}, \bar{X} + Z_{\frac{\alpha}{2}} \dfrac{\sigma}{\sqrt{n}} \right)$

Así pues el error máximo que podemos conocer a un nivel de significación α es: $E = Z_{\frac{\alpha}{2}} \dfrac{\sigma}{\sqrt{n}}$

Despejando de esta igualdad n obtenemos el tamaño muestral $n = \dfrac{Z_{\frac{\alpha}{2}}^{2} \cdot \sigma^{2}}{E^{2}}$

2.-Para la estimación de una proporción

Un intervalo de confianza para una proporción es: $I = \left(\hat{p} - Z_{\frac{\alpha}{2}} \sqrt{\dfrac{\hat{p} \cdot q}{n}}, \hat{p} + Z_{\frac{\alpha}{2}} \sqrt{\dfrac{\hat{p} \cdot q}{n}} \right)$

El error máximo que podemos cometer a un nivel de significación α es: $E = Z_{\frac{\alpha}{2}} \sqrt{\dfrac{\hat{p} \cdot q}{n}}$

Despejando de esta igualdad n obtenemos el tamaño muestral $n = \dfrac{Z_{\frac{\alpha}{2}}^{2} \cdot \hat{p} \cdot q}{E^{2}}$

PROBLEMAS.-

1.- De una población normal N(μ, σ) de la que conocemos su desviación típica σ = 3, extraemos una muestra aleatoria simple de tamaño 9. Sabiendo que la media muestral es igual a 20, obtener un intervalo de confianza del 95%.

2.- Se sabe que el peso de las jugadoras de la liga de fútbol profesional se distribuye normalmente con desviación típica de 6 kg. Para estudiar el peso medio de las jugadoras se extrae una muestra de tamaño 8, obteniendo los siguientes resultados : 63,7 ; 48 ; 43,5 ; 65 ; 82 ; 70,3 ; 56,5 ; 50

Calcula un intervalo de confianza a un nivel de significación del 10% para el peso medio de las jugadoras.

3.- Consideremos una población normal N(μ, σ) de la cual desconocemos tanto la media como la desviación típica. De la población extraemos una muestra aleatoria simple de tamaño 144 resultando una media muestral de 2 y una desviación típica muestral de 0,6. Calcular un intervalo de confianza para la media poblacional a un nivel de confianza del 99%.

4.- Una empresa dedicada al embotellamiento de vino quiere establecer los límites máximos y mínimos de vino contenido en botellas de 1 litro. Para ello selecciona 100 botellas resultando que el contenido medio de vino es de 998 cm^3, con desviación típica de 5 cm^3 Si suponemos que la cantidad de vino introducida en una botella sigue una distribución normal, calcular dichos límites (máximo y mínimo) de confianza al 98%.

5.- De una muestra aleatoria de 267 usuarios de ordenadores personales, 205 manifestaron que prefieren el entorno de trabajo Macintosh al entorno Windows. Con base a esta muestra, calcula un intervalo de confianza del 98% para el porcentaje de usuarios que prefieren el entorno Macintosh.

6.- Se pretende realizar un estudio sobre el tiempo medio que tarda un antivirus en reparar un archivo infectado. Determinar el tamaño de la muestra necesario para poder realizar dicha estimación con un error medio de media hora a un nivel de confianza del 95%. Se sabe por estudios realizados anteriormente que la desviación típica poblacional es σ = 3.