ESTIMACIÓN DE PARÁMETROS

1.- INTRODUCCION

Los estimadores son estadísticos que vamos a utilizar para averiguar el valor de un parámetro desconocido, además, son variables aleatorias y por tanto tendrán media y varianza.

$$\mu_{\hat{\theta}} = \text{E } \hat{\theta}$$

$$\sigma^2_{\hat{\theta}} = Var_{\hat{\theta}}$$

Para demostrar las propiedades de los estimadores ($\hat{\theta}$) calcularemos su media y varianza: $\mu_{\hat{\theta}}$ y $\sigma^2_{\hat{\theta}}$

2.- PROPIEDADES DESEABLES DE LOS ESTIMADORES

✓ SUFICIENCIA

Se dice que $\hat{\theta}$ es suficiente si al reducir la muestra a un solo valor no se pierde la información. Es suficiente si contiene toda la muestra.

✓ INSESGADEZ

Se dice que $\hat{\theta}$ es centrado o INSESGADO si E $\hat{\theta} = \theta$

Si E $\hat{\theta} = \theta$ + b ($\hat{\theta}$) se dice que $\hat{\theta}$ es SESGADO

Sesgo $\Rightarrow$ b ($\hat{\theta}$) = E $\hat{\theta}$ - θ

Un $\hat{\theta}$ sesgado podrá ser asintóticamente insesgado cuando:

para muestras grandes n $\cong \infty$

$$\lim_{n \to \infty} \text{b}(\hat{\theta}) = 0$$

$$\lim_{n \to \infty} \text{E}(\hat{\theta}) = \lim_{n \to \infty} (\theta + \text{b}(\hat{\theta})) = \theta$$

✓ CONSISTENCIA

Se dice que $\hat{\theta}$ es consistente si $\lim_{n \to \infty} ECM_{\hat{\theta}} = 0$

$$\text{Donde: } ECM_{\hat{\theta}} = E(\hat{\theta} - \theta)^2 = Var\ \hat{\theta} + b(\hat{\theta})^2$$

✓ EFICIENCIA

Es una propiedad que la utilizaremos para comparar dos estimadores.

- Si dos estimadores son insesgados, será más eficiente el que tenga menor varianza
- Si dos estimadores son sesgados o uno de ellos es sesgado utilizaremos el criterio del error cuadráticos medio y será más eficiente el que menor error cuadrático medio tenga.

PROPIEDADES DE $\bar{x}$ COMO ESTIMADOR DE μ

- $\bar{x}$ siempre es un estimador **insesgado** de μ ya que: E ($\bar{x}$) = μ

$$\hat{\mu} = \bar{x}$$

$$E(\hat{\mu}) = E\left(\frac{x_1 + x_2 + \ldots + x_n}{n}\right) = \left(\frac{Ex_1 + Ex_2 + \ldots + Ex_n}{n}\right) = \left(\frac{\mu + \mu + \ldots + \mu}{n}\right) = \left(\frac{n\,\mu}{n}\right) = \mu$$

Como $E(\hat{\mu}) = \mu$, podemos decir que $\bar{x}$ es un estimador insesgado.

- $\bar{x}$ siempre es un estimador **consistente** de μ ya que: Var ($\bar{x}$) = $\frac{\sigma^2}{n}$

Por tanto, será consistente si

$$\lim_{n \to \infty} ECM_{\hat{\theta}} = 0$$

$$\text{Como: } ECM_{\hat{m}} = Var_{\hat{\mu}} + b(\hat{\mu})^2 = \frac{\sigma^2}{n} + 0^2$$

$$\lim_{n \to \infty} ECM_{\hat{\mu}} = \lim_{n \to \infty} Var_{\hat{\mu}} = \lim_{n \to \infty} \frac{\sigma^2}{n} = 0$$

3.- ESTIMACION DE PARAMETROS POR PUNTO

Si desconocemos algún parámetro que defina una determinada población, para trabajar con dicho colectivo será necesario que estimemos los parámetros desconocidos a través de una m.a.s..

Los métodos más habituales para estimar parámetros son:

- ✓ **Método de máxima verosimilitud**

  La lógica subyacente a este método es calcular los estimadores que maximicen la distribución de probabilidades conjunta para una muestra determinada.

- ✓ **Método de los momentos**

  La lógica subyacente a este método es igualar momentos poblacionales a momentos muestrales y de esta ecuación, despejaremos el parámetro a estimar.

Ejemplos de algunos parámetros y sus estimadores puntuales:

Media de una población (μ), su estimador, $\hat{\mu} = \bar{x}$

Proporción de una población (π), su estimador, $\hat{\pi} = \mathrm{P} = \frac{x}{n}$

Varianza de una población (σ^2), su estimador, $\hat{\sigma}^2 = \mathbf{S}^2$

Nota: Como estimador de la varianza, vamos a utilizar la **cuasivarianza muestral**, puesto que es un estimador **insesgado y consistente** de la varianza pobalcional.

$$\mathbf{S}^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1} n_i$$

4.- ESTIMACION DE PARAMETROS POR INTERVALO, (I.C.)

La estimación por intervalos de confianza consiste en acotar un parámetro desconocido entre dos valores y además permite conocer la fiabilidad de los resultados obtenidos, ya que indica el grado de confianza (1 - α) que se tiene de que el intervalo incluya entre sus límites el verdadero valor del parámetro a estimar. Por ejemplo, si construimos un intervalo de confianza del 90%, significa que de 100 intervalos construidos a partir de distintas muestras, aproximadamente 90 contendrán el parámetro a estimar

DEFINICIONES Y TERMINOLOGIA

- Confianza (1 - α)
- Nivel de significación α .
- Margen de error o precisión . Es la semiamplitud del I.C.
- Error estandar. Es la desviación del estimador.

DESARROLLO DEL I.C. PARA "μ", CUANDO EL COLECTIVO SEA NORMAL Y LA VARIANZA CONOCIDA.

Sabemos que: $\dfrac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \in$ N (0 , 1)

$$P\left[-z_{\frac{\alpha}{2}} \le \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \le z_{\frac{\alpha}{2}}\right] = 1-\alpha$$

$$P\left[-z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \le \bar{x}-\mu \le z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right] = 1-\alpha$$

$$P\left[-\bar{x}-z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \le -\mu \le -\bar{x}+z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right] = 1-\alpha$$

Multiplico por (-1) $\Rightarrow P\left[\bar{x}+z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \ge \mu \ge \bar{x}-z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right] = 1-\alpha$

Ordenamos $\Rightarrow P\left[\bar{x}-z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x}+z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right] = 1-\alpha$

Por lo tanto el $IC_{1-\alpha}(\mu) = \left[\bar{x}-z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} , \bar{x}+z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$

$$IC_{1-\alpha}(\mu) = \left[\bar{x} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

DESARROLLO DEL I.C. PARA "μ", CUANDO EL COLECTIVO SEA NORMAL Y LA VARIANZA DESCONOCIDA.

Sea un colectivo X $\in$ N (m , σ^2) del cuál extraemos una m.a.s. de tamaño "n" con la que creamos los siguientes estadísticos:

$$\bar{x} \in N\left(\mu, \frac{\sigma^2}{n}\right)$$

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

Con estos dos estadísticos formamos un nuevo estadístico:

$$\frac{\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)\cdot s^2}{\sigma^2}\Big/ n-1}} = \frac{\bar{x}-\mu}{s/\sqrt{n}} \in t_{n-1|} \equiv \frac{N(0,1)}{\sqrt{\dfrac{\chi^2_{n-1|}}{n-1}}}$$

Usaremos este estadístico para construir IC para la μ en los casos en los que σ^2 sea desconocida.

Sabemos que $\dfrac{\bar{x}-\mu}{s/\sqrt{n}} \in t_{n-1|}$

$$P\left[-t_{n-1|\frac{\alpha}{2}} \leq \frac{\bar{x}-\mu}{s/\sqrt{n}} \leq t_{n-1|\frac{\alpha}{2}}\right] = 1-\alpha$$

.
.
.

$$IC_{1-\alpha}(\mu) = \left[\bar{x} - t_{n-1|\frac{\alpha}{2}}\frac{s}{\sqrt{n}}, \bar{x} + t_{n-1|\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\right]$$

$$IC_{1-\alpha}(\mu) = \left[\bar{x} \pm t_{n-1|\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\right]$$

DESARROLLO DEL I.C. PARA LA PROPORCION "π",

Sea X un colectivo binario, extraemos una m.a.s. los suficiente grande ($n \geq 30$). Aplicando el Teorema de Moivre, podemos afirmar que la v.a. X "Nº de éxitos" se aproximará a una distribución normal.

$$X = \text{"Nº de éxitos"} = \Sigma x_i \rightarrow N(\mu = n\pi,\ \sigma^2 = n\pi(1-\pi))$$

Definimos el estadístico, proporción muestral: $p = \dfrac{X}{n} = \dfrac{N^o\, exitos}{N^o\, intentos}$

La distribución aproximada del estadístico proporción muestral, será:

$$P = \frac{x}{n} \rightarrow N\left(\pi,\ \sigma^2 = \frac{\pi(1-\pi)}{n}\right) \quad \Rightarrow \quad \frac{\frac{x}{n} - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \rightarrow N(0,1)$$

Utilizaremos este estadístico para construir IC para la proporción "π".

Sabemos que: $\dfrac{\frac{x}{n} - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \rightarrow N(0,1)$

$$P\left[-z_{\frac{\alpha}{2}} \leq \frac{\frac{x}{n} - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \leq z_{\frac{\alpha}{2}}\right] \approx 1-\alpha$$

.
.
.

$$P\left[\frac{x}{n} - z_{\frac{\alpha}{2}}\sqrt{\frac{\pi(1-\pi)}{n}} \leq \pi \leq \frac{x}{n} + z_{\frac{\alpha}{2}}\sqrt{\frac{\pi(1-\pi)}{n}}\right] \approx 1-\alpha$$

No podemos acotar π en función de π, por lo tanto vamos a sustituir π por su estimador máximo verosímil:

$$\hat{\pi} = p = \frac{x}{n}$$

$$IC_{\approx 1-\alpha}(\pi) = \left[p \pm z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right]$$

Cuando tengamos como dato la precisión o semiamplitud y nos pregunten por el tamaño muestral "n", si no tenemos información sobre p, utilizaremos el caso más desvaforable, que es: $p = \dfrac{1}{2}$