Miércoles 20 de octubre de 2021.

Examen 2
- Oral : Martes 26/10/21
- Escrito : Viernes 29/10/21

5 ESTIMACION

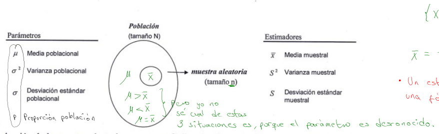
Objetivo: Aprender la construcción y características de los estimadores de los parámetros más comunes, de manera puntual y por intervalo.

Las mediciones hechas de una población reciben el nombre de parámetros, por ejemplo la media poblacional μ, la varianza poblacional σ^2 y la desviación estándar poblacional σ.
Conocer el valor de un parámetro implica realizar un censo lo cual además de ser costoso en ocasiones es casi imposible. Dado que es necesario tener idea de por donde está el valor del parámetro se acostumbra obtener un pedacito de la población, al cual se le llama muestra aleatoria (m.a), se busca que tal muestra sea lo más parecida posible a la población de donde salió para poder decir cosas sobre la población en base al estudio de la muestra. Las mediciones realizadas en una muestra aleatoria se llaman estimadores. Ejemplos de estimadores son la media muestral $\bar{x}$, la varianza muestral S^2 y la desviación estándar muestral S.

Con $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$, $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$ y $S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$, como se observa los estimadores son funciones de la m.a.

Una vez que se tiene la m.a y se sustituyen los valores en la fórmula de un estimador ya se tiene un valor numérico al cual se le llama estimación.

$\{X_1, X_2, \dots, X_n\}$ m.a $\longleftrightarrow$ Hay independencia



La elección de la muestra aleatoria no es tarea fácil, existe la teoría del muestreo la cual estudia los diferentes tipos de muestreo y su análisis.

$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$

- Un estimador es una fórmula.

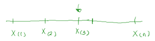
μ : parámetro

$\hat{\mu}$: estimador del parámetro

$\hat{\mu}_1=\bar{x}$ $\qquad$ $\hat{\mu}_4=\frac{x_{(1)}+x_{(n)}}{2}$

$\hat{\mu}_2=x_M$: Moda

$\hat{\mu}_3=x_m$: media



estadísticas de orden

$x_{(1)}\le x_{(2)}\le \cdots \le x_{(n-1)}\le x_{(n)}$

$x_1, x_2, x_3, \quad , x_7, x_8$
3, 1, 3, 2, 1, 1, 3, 2

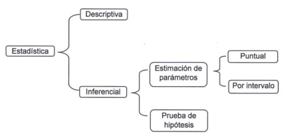
$x_{(1)}=1$ $\quad$ $x_{(2)}=1$ $\quad$ $x_{(3)}=1$ $\quad$ $x_{(4)}=2$

$x_{(5)}=2$ $\quad$ $x_{(6)}=3$ $\quad$ $x_{(7)}=3$ $\quad$ $x_{(8)}=3$

5.1 Estimación De Parámetros.

La parte de las matemáticas que trata de resolver problemas en base al análisis realizado sobre datos que han sido obtenidos de la manera más adecuada posible (en base a un buen diseño de muestreo) se conoce como **estadística**.

La estadística descriptiva organiza, resume y presenta los datos, mientras que la estadística inferencial toma decisiones respecto a una población en base a todo un estudio que hace sobre una parte de tal población, conocida como **muestra**.



Podemos proponer $\underbrace{\text{estimadores}}_{\text{fórmula}}$ SIN tener datos (la m.a)

Luego se hace el trabajo de campo para obtener la base de datos $\underbrace{(x_1, x_2, \dots, x_n)}_{\text{números}}$

Podemos evaluar las fórmulas (estimadores) y encontrar su valor numérico y éste se llama "estimación".

5.1.1 Estimación puntual

Generalmente los parámetros de las distribuciones que describen adecuadamente a un conjunto de datos son desconocidos por lo que surge la necesidad de "*estimarlos*".

Definición 1: Un **estimador** es una función de las variables aleatorias $X_1, X_2, \dots, X_n$. Es también una variable aleatoria y su función de distribución de probabilidad puede obtenerse a partir de la distribución conjunta de $X_1, X_2, \dots, X_n$.

Ejemplo 1: Supón que el parámetro de interés es θ. Entonces $\hat{\theta}=f(X_1, X_2, \dots, X_n)$ es un estimador de θ.

Existen dos tipos de estimadores; **puntual** si solo se presenta un número como posible valor del parámetro θ, y **por intervalo** si se presenta un conjunto de números dentro del cual puede estar el valor de θ.
Por ejemplo cuando decimos; edad promedio de 35 años, estatura promedio de 1.65m, calificación promedio de 9.3, etc. son estimaciones puntuales.
Una vez que se ha obtenido la muestra aleatoria y sustituido en la función del estimador se obtiene un número al cual se le conoce como estimación.

Población
Gpo 2CM16 PyE

Variable de interés : edad
N = 28
μ = 22.5 años

Muestra
n = 5

$\underbrace{\bar{x}=\frac{1}{n}\sum^{5}x_i}_{\text{estimador (fórmula)}}=\underbrace{\frac{1}{5}(108)=21.6\text{ años}}_{\text{Evaluación numérica del estimador}}$

estimación (número)

$\mu > \bar{x}$ Sub-estimo

$\mu < \bar{x}$ sobre-estimo

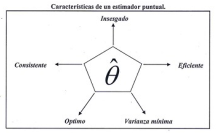
$\mu = \bar{x}$

Para otra m.a me sale que la estimación es

$\bar{x}$ = 23 años

Se sabe que la media muestral $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ es un estimador de la media poblacional μ. Si se ha seleccionado la m.a de tamaño 3 dada por: $x_1=20$, $x_2=18$ y $x_3=21$ se tiene que una estimación para la media poblacional es $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{20+18+21}{3}=19.67$.
A todo estimador puntual se le piden ciertas características para considerarlo bueno. Estas características sirven para decidir con cual estimador quedarnos en caso de ser necesario.

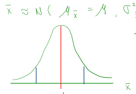
Características de un estimador puntual.



Como ya se dijo $\hat{\theta}$ es una v.a y tiene una distribución de probabilidad. Si $\hat{\theta}_1$ y $\hat{\theta}_2$ son estimadores diferentes del parámetro θ tales que $E(\hat{\theta}_1)=\theta$ y $E(\hat{\theta}_2)>\theta$.

¿Qué estimador es "mejor"?
Si tenemos que elegir entre uno de los dos, nos decidiríamos por el primero es decir $\hat{\theta}_1$ ya que en promedio le pega al verdadero valor del parámetro; esto significa que si se repite muchas veces el muestreo obteniendo estimaciones $\hat{\theta}_{11}, \hat{\theta}_{12}, \dots, \hat{\theta}_{1k}$ gran número de estos valores están cerca del parámetro de interés, sin embargo no ocurre lo mismo si usamos $\hat{\theta}_2$ para estimar al parámetro ya que sólo una pequeña proporción de estos valores están cerca de θ.
Pero si las distribuciones de los dos estimadores están centradas en θ con $V(\hat{\theta}_1)<V(\hat{\theta}_2)$. ¿Qué estimador es mejor? Bajo estas circunstancias $\hat{\theta}_1$ es mejor ya que tiene una menor dispersión, esto garantiza que la proporción de estimadores del parámetro que se acercan al parámetro es mayor que la proporción correspondiente de los valores de $\hat{\theta}_2$.

$\bar{x}\approx N\left(\mu_{\bar{x}}=\mu, \sigma^2_{\bar{x}}=\frac{\sigma^2}{n}\right)$



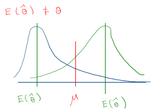
Definición 2: $\hat{\theta}$ es un estimador **insesgado** de θ si $E(\hat{\theta})=\theta$. $\longrightarrow$ No significa que $\hat{\theta}=\theta$

Definición 3: ***i)*** El error cuadrático medio del estimador está dado por $ECM(\hat{\theta})=E(\theta-\hat{\theta})^2$.

ii) La eficiencia de $\hat{\theta}_1$ relativa a $\hat{\theta}_2$ se define como $Er=\frac{ECM(\hat{\theta}_1)}{ECM(\hat{\theta}_2)}$.

Entonces $\hat{\theta}$ es un estimador sesgado cuando

$E(\hat{\theta})\neq\theta$



Observaciones:

1) El error cuadrático medio es un medio para comparar estimadores, así, si $ECM(\hat{\theta}_1)<ECM(\hat{\theta}_2)$ decimos que $\hat{\theta}_1$ es "mejor" que $\hat{\theta}_2$.

2) Si $Er<1$, se dice que $\hat{\theta}_1$ es más **eficiente** que $\hat{\theta}_2$.

3) $\boxed{Sesgo(\hat{\theta})=E(\hat{\theta})-\theta}$

si $Sesgo(\hat{\theta})=0 \rightarrow E(\hat{\theta})-\theta=0 \rightarrow E(\hat{\theta})=\theta$
si $Sesgo(\hat{\theta})\neq 0 \rightarrow E(\hat{\theta})-\theta\neq 0 \rightarrow E(\hat{\theta})\neq\theta$

$$
\begin{aligned}
ECM(\hat{\theta}) &= E(\theta-\hat{\theta})^2\\
&= E\left[\theta-\hat{\theta}+E(\hat{\theta})-E(\hat{\theta})\right]^2\\
&= E\left[\left(E(\hat{\theta})-\hat{\theta}\right)-\left(E(\hat{\theta})-\theta\right)\right]^2\\
&= E\left[\left(E(\hat{\theta})-\hat{\theta}\right)^2-2\left(E(\hat{\theta})-\hat{\theta}\right)\left(E(\hat{\theta})-\theta\right)+\left(E(\hat{\theta})-\theta\right)^2\right]\\
&= E\left[E(\hat{\theta})-\hat{\theta}\right]^2-2E\left[E(\hat{\theta})-\hat{\theta}\right]Sesgo(\hat{\theta})+E\left[Sesgo(\hat{\theta})\right]^2\\
&= V(\hat{\theta})-2Sesgo(\hat{\theta})\left[E(\hat{\theta})-E(\hat{\theta})\right]+Sesgo^2(\hat{\theta})
\end{aligned}
$$

$$\boxed{ECM(\hat{\theta})=V(\hat{\theta})+Sesgo^2(\hat{\theta})}$$

Si el estimador es insesgado, se tiene que $Sesgo(\hat{\theta})=0$ y entonces $ECM(\hat{\theta})=V(\hat{\theta})$.

El estimador insesgado de varianza mínima es aquel que tiene varianza igual a la cota de Cramér Rao, CCR. Los demás estimadores tienen varianza mayor a dicha cota.

$$\boxed{V(\hat{\theta})\ge\frac{1}{nE\left[\frac{\partial}{\partial\theta}\ln f_X(x;\theta)\right]^2}=CCR}$$

Si $V(\hat{\theta})=\frac{1}{nE\left[\frac{\partial}{\partial\theta}\ln f_X(x;\theta)\right]^2}$ se dice que se tiene el estimador insesgado de **varianza mínima** para el parámetro $\hat{\theta}$.

Ejemplo 2: Si X es una v.a con media μ y varianza σ^2 y $X_1, X_2, \dots, X_n$ es una m.a de tamaño n de X, se tiene que la media muestral $\overline{X}$ y la varianza muestral S^2 son estimadores puntuales de la media y la varianza poblacional respectivamente. ¿Son insesgados estos estimadores?

Solución

La media muestral está dada por la expresión $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$. Entonces su esperanza es:

$E(\overline{X})=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\mu \therefore$ $\overline{X}$ sí estima de manera insesgada a la media poblacional.

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

¿Qué hago para averiguar si S^2 es un estimador insesgado?

Ejemplo 3: Supón que se tiene una m.a de tamaño $2n$ de una población denotada por X con media μ y varianza σ^2 y que $\hat{\mu}_1=\frac{1}{2n}\sum_{i=1}^{2n}X_i$ y $\hat{\mu}_2=\frac{1}{n}\sum_{i=1}^{n}X_i$ son dos estimadores para la media poblacional μ.
¿Cuál estimador recomendarías? Hasta ahora solo tengo para decidir la característica del sesgo.

Solución

$= \frac{2n\mu}{2n}$

$E(\hat{\mu}_1)=E\left[\frac{1}{2n}\sum_{i=1}^{2n}X_i\right]=\mu$ $\rightarrow$ $\hat{\mu}_1$ es insesgado

$E(\hat{\mu}_2)=E\left[\frac{1}{n}\sum_{i=1}^{n}X_i\right]=\mu$ $\rightarrow$ $\hat{\mu}_2$ es insesgado

Considerando solo el criterio del sesgo los 2 son igual de buenos.
Para decidir hay que usar otro criterio.

Ambos estimadores son insesgados, entonces hay que considerar otro criterio. Veamos cuales son sus errores cuadráticos medios.

$$ECM(\hat{\mu}_1)=V(\overline{X})=\frac{\sigma^2}{2n}$$

$$ECM(\hat{\mu}_2)=\frac{n\sigma^2}{n^2}=\frac{\sigma^2}{n}$$

$$Er=\frac{\sigma^2}{2n}\frac{n}{\sigma^2}=\frac{1}{2}<1$$

Entonces recomendaríamos $\hat{\mu}_1$ porque estima de manera más eficiente a la media poblacional.

Ejemplo 4: Si $X_1, X_2, \dots, X_7$ es una m.a de una población que tiene media μ y varianza σ^2.
Considera los siguientes estimadores de μ:

$\hat{\theta}_2=\frac{1}{2}(2X_1-X_6+X_4)$ $\qquad$ $\hat{\theta}_1=\frac{1}{7}\sum_{i=1}^{7}X_i$ $\qquad$ Considerando el criterio del sesgo

a) ¿Alguno de estos estimadores es insesgado?
b) ¿Cuál es mejor? ¿Por qué?

Solución

$=E\left(\frac{1}{2}(2X_1-X_6+X_4)\right)=\frac{1}{2}\left[2E(X_1)-E(X_6)+E(X_4)\right]$
$=\frac{1}{2}\left[2\mu-1+\mu\right]=\mu$

Como los 2 son insesgados necesito otro criterio.

a) $E(\hat{\theta}_1)=\frac{1}{7}(7\mu)=\mu$ y $E(\hat{\theta}_2)=\frac{1}{2}(2-1+1)\mu=\mu$. Ambos estiman de manera insesgada a la media poblacional.

b) $ECM(\hat{\theta}_1)=V(\overline{X})=\frac{\sigma^2}{7}$

$ECM(\hat{\theta}_2)=V(\hat{\theta}_2)=\frac{1}{4}(6\sigma^2)=\frac{3}{2}\sigma^2$

Entonces $\hat{\theta}_1$ es más eficiente que $\hat{\theta}_2$.

Ejemplo 5: Considera que se toman tres muestras aleatorias de tamaños 10, 8 y 6 de una población con media μ y varianza σ^2. S_1^2, S_2^2 y S_3^2 las varianzas correspondientes de las muestras. ¿Será S^2 un estimador insesgado de la varianza poblacional? $S^2=\frac{1}{24}\left(10S_1^2+8S_2^2+6S_3^2\right)$

Solución

Ya vimos que $E(S^2)=\sigma^2$ entonces $E(S^2)=\frac{1}{24}(10+8+6)\sigma^2=\sigma^2$, por tanto sí es un estimador insesgado del la varianza poblacional.