

Auxiliar 2 Estadística Semestre Primavera 2005

Prof: Nancy Lacourly

Aux: Sebastián Court

Algunas definiciones.- Diremos que un método de estimación es **insesgado** si el valor esperado del estimador, tomado sobre todas las muestras posibles de tamaño n dado, es exactamente igual al valor verdadero de la población.

En esta auxiliar estaremos trabajando en muestras aleatorias sin reemplazo salvo se diga lo contrario.

Problema1.- Consideremos una población finita Y (Y_1, \dots, Y_N) de tamaño N con media poblacional $\mu = \frac{Y_1 + \dots + Y_N}{N}$. No conocemos el valor de μ , ya que no conocemos con exactitud la población Y , sin embargo deseamos estimarlo tomando una m.a.s de tamaño n (y_1, \dots, y_n) y calculando $\bar{y} = \frac{y_1 + \dots + y_n}{n}$. Demuestre que \bar{y} es un estimador insesgado para μ . Deduzca que $N\bar{y}$ es un estimador insesgado para $Y_1 + \dots + Y_N$.

Solución.- Notemos primero que podemos tomar un total de $\binom{N}{n}$ muestras distintas de tamaño n .

Dado que cada muestra que se tome tendrá un valor para \bar{y} distinto, tenemos que cada uno tiene una probabilidad $\frac{1}{\binom{N}{n}}$ de ocurrir.

Por lo tanto,

$$\mathbb{E}(\bar{y}) = \frac{\sum \bar{y}}{\binom{N}{n}} = \frac{\sum y_1 + \dots + y_n}{n \binom{N}{n}} \text{ donde la suma es sobre todas las muestras posibles.}$$

Estamos interesados ahora en evaluar $\sum y_1 + \dots + y_n$. Para eso buscamos cuantas muestras contienen a y_i cualquiera. Esta cantidad será $\binom{N-1}{n-1}$, luego:

$$\sum y_1 + \dots + y_n = \binom{N-1}{n-1} (Y_1 + \dots + Y_N)$$

Por lo tanto,

$$\mathbb{E}(\bar{y}) = \frac{\binom{N-1}{n-1}}{\binom{N}{n}} \frac{n!(N-n)!}{n \cdot N!} (Y_1 + \dots + Y_N) = \frac{Y_1 + \dots + Y_N}{N} = \mu.$$

Por último $\mathbb{E}(N\bar{y}) = N\mathbb{E}(\bar{y}) = N \frac{Y_1 + \dots + Y_N}{N} = Y_1 + \dots + Y_N$. \square

Problema2.- Con las mismas hipótesis del problema anterior, demuestre que la varianza de \bar{y} es

$$V(\bar{y}) = \frac{S^2}{n} \left(\frac{N-n}{N-1} \right)$$

$$\text{donde } S^2 = \frac{\sum_{i=1}^N (Y_i - \mu)^2}{N}.$$

Solución.- Notemos que:

$$n(\bar{y} - \mu) = (y_1 - \mu) + (y_2 - \mu) + \dots + (y_n - \mu) \quad (1)$$

y también que

$$\mathbb{E}((y_1 - \mu)^2 + (y_2 - \mu)^2 + \dots + (y_n - \mu)^2) = \frac{n}{N}((Y_1 - \mu)^2 + (Y_2 - \mu)^2 + \dots + (Y_N - \mu)^2). \quad (2)$$

Además

$$\begin{aligned} & \mathbb{E}((y_1 - \mu)(y_2 - \mu) + (y_1 - \mu)(y_3 - \mu) + \dots + (y_{n-1} - \mu)(y_n - \mu)) \\ &= \frac{n(n-1)}{N(N-1)}((Y_1 - \mu)(Y_2 - \mu) + (Y_1 - \mu)(Y_3 - \mu) + \dots + (Y_{N-1} - \mu)(Y_N - \mu)) \quad (3) \end{aligned}$$

Donde la suma de los productos se extiende en todos los productos cruzados $(y_i - \mu)(y_j - \mu)$ con $i < j$. Además el factor $\frac{n(n-1)}{N(N-1)}$ aparece de el hecho que el lado izquierdo tiene $\frac{n(n-1)}{2}$ términos y el lado derecho $\frac{N(N-1)}{2}$ términos. Si elevamos al cuadrado (1), tomando esperanza sobre todas las muestras simples aleatorias y usando (2) y (3) obtenemos:

$$n^2 \mathbb{E}((\bar{y} - \mu)^2) = \frac{n}{N} \left\{ \sum_{i=1}^N (Y_i - \mu)^2 + \frac{2(n-1)}{N-1} ((Y_1 - \mu)(Y_2 - \mu) + (Y_1 - \mu)(Y_3 - \mu) + \dots + (Y_{N-1} - \mu)(Y_N - \mu)) \right\}$$

Completando los cuadrados respectivos:

$$n^2 \mathbb{E}((\bar{y} - \mu)^2) = \frac{n}{N} \left\{ \left(1 - \frac{n-1}{N-1}\right) \sum_{i=1}^N (Y_i - \mu)^2 + \frac{n-1}{N-1} ((Y_1 - \mu) + (Y_2 - \mu) + \dots + (Y_N - \mu))^2 \right\}$$

donde el segundo término dentro de la llave es 0. Por lo que si dividimos la expresión anterior por n^2 obtenemos:

$$\mathbb{E}((\bar{y} - \mu)^2) = \frac{N-n}{nN(N-1)} \sum_{i=1}^N (Y_i - \mu)^2 = \frac{S^2}{n} \frac{N-n}{N-1}. \quad \square$$

Problema3.- (*Teorema Central del Límite*) Supóngase que tenemos cierto número de voltajes con ruido independiente, $V_i, i = 1, \dots, n$, que se reciben en lo que se llama un "sumador". Sea V la suma de los voltajes recibidos, es decir, $V = V_1 + \dots + V_n$. Supóngase que cada una de las variable aleatorias V_i esta distribuida uniformemente en el intervalo $[0, 10]$. Calcule la probabilidad de que el voltaje de entrada sobrepase los 105 Volts, para el caso $n = 20$ y $n = 100$. Además entregue una cota para la probabilidad usando la desigualdad de Tchebyshev.

Solución.- Tenemos que $\mathbb{E}(V_i) = 5$ y que $V(V_i) = 100/12$.

Para $n = 20$:

De acuerdo al teorema central del límite:

$$S = \frac{(V-5n)\sqrt{12}}{10\sqrt{n}} \rightarrow N(0, 1)$$

Luego:

$$\mathbb{P}(V > 105) = \mathbb{P}\left(\frac{V-100}{(10/\sqrt{12})\sqrt{20}} > \frac{105-100}{(10/\sqrt{12})\sqrt{20}}\right) = 1 - \Phi(0,388) = 0,352.$$

Para $n = 100$:

De acuerdo al teorema central del límite:

$$S = \frac{(V-5n)\sqrt{12}}{10\sqrt{n}} \rightarrow N(0, 1)$$

Luego:

$$\mathbb{P}(V > 105) = \mathbb{P}\left(\frac{V-500}{(10/\sqrt{12})\sqrt{100}} > \frac{105-500}{(10/\sqrt{12})\sqrt{100}}\right) = 1 - \Phi(-13,683) \approx 1.$$

Por otro lado, recordemos que la desigualdad de Tchebyshev dice que:

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

Luego,

$$\mathbb{P}(V \geq 105) \leq \frac{\mathbb{E}(V)}{105}$$

donde,

$$\mathbb{E}(V) = \mathbb{E}(V_1 + \dots + V_n) = \mathbb{E}(V_1) + \dots + \mathbb{E}(V_n) = 5n$$

$$\text{Si } n = 20: \mathbb{P}(V \geq 105) \leq \frac{100}{105} = 0,952$$

$$\text{Si } n = 100 : \mathbb{P}(V \geq 105) \leq \frac{500}{105} = 4,762 \quad \square$$