1. Propiedades de los estimadores

1.1. Eficiencia relativa. Definición: Dados dos estimadores insesgados, $\hat{\theta}_1$ y $\hat{\theta}_2$, de un parámetro θ, con varianzas $V(\hat{\theta}_1)$ y $V(\hat{\theta}_2)$, entonces la eficiencia (eff) de $\hat{\theta}_1$ respecto a $\hat{\theta}_2$, se define como la razón

$$\mathrm{eff}(\hat{\theta}_1, \hat{\theta}_2) = \frac{V(\hat{\theta}_2)}{V(\hat{\theta}_1)}.$$

Si $\mathrm{eff}(\hat{\theta}_1, \hat{\theta}_2) > 1$ entonces se prefiere al estimador $\hat{\theta}_1$, mientras que si $\mathrm{eff}(\hat{\theta}_1, \hat{\theta}_2) < 1$ entonces se prefiere al estimador $\hat{\theta}_2$; ya que estamos interesados en tener el estimador de menor varianza.

Ejemplo. Como en un ejercicio anterior, consideramos una muestra aleatoria de tamaño tres de una distribución exponencial con media θ, la cual establece que $\hat{\theta}_1 = Y_1$, $\hat{\theta}_2 = (Y_1 + Y_2)/2$, $\hat{\theta}_3 = (Y_1 + 2Y_2)/3$ y $\hat{\theta}_5 = \bar{Y}$ son estimadores insesgados de θ. Encuentre la eficiencia de $\hat{\theta}_1$ respecto a $\hat{\theta}_5$, de $\hat{\theta}_2$ en relación con $\hat{\theta}_5$ y de $\hat{\theta}_3$ respecto a $\hat{\theta}_5$.

Solución.
Como calculamos en un ejercicio anterior las varianzas, entonces las eficiencias pedidas son:

$$\mathrm{eff}(\hat{\theta}_1, \hat{\theta}_5) = \frac{\frac{1}{3}\theta^2}{\theta^2} = \frac{1}{3}.$$

Como la eficiencia da menor que 1, entonces $\hat{\theta}_5$ es mas eficiente que $\hat{\theta}_1$.

$$\mathrm{eff}(\hat{\theta}_2, \hat{\theta}_5) = \frac{\frac{1}{3}\theta^2}{\frac{1}{2}\theta^2} = \frac{2}{3}.$$

Nuevamente, como la eficiencia es menor que 1, entonces $\hat{\theta}_5$ es mas eficiente que $\hat{\theta}_2$.

$$\mathrm{eff}(\hat{\theta}_3, \hat{\theta}_5) = \frac{\frac{1}{3}\theta^2}{\frac{5}{3}\theta^2} = \frac{1}{5}.$$

Otra vez, se obtiene que $\hat{\theta}_5$ es el estimador más eficiente.

Ejemplo. Suponga que Y_1, Y_2, ..., Y_n denota una muestra aleatoria de una distribución uniforme en el intervalo $(\theta, \theta + 1)$. Sean

$$\hat{\theta}_1 = \bar{Y} - \frac{1}{2} \text{ y } \hat{\theta}_2 = Y_{(n)} - \frac{n}{n+1}.$$

a Demuestre que $\hat{\theta}_1$ y $\hat{\theta}_2$ son estimadores insesgados de θ.
b Encuentre la eficiencia de $\hat{\theta}_1$ respecto a $\hat{\theta}_2$.

<u>Solución</u>.
a) Primero mostremos que ambos estimadores son insesgados, es decir, que $E[\hat{\theta}] = \theta$.

$$E[\hat{\theta}_1] = E\left[\bar{Y} - \frac{1}{2}\right] = \frac{1}{n}E\left[\sum_{i=1}^{n} Y_i\right] - \frac{1}{2} = \frac{1}{n}\sum_{i=1}^{n} E[Y_i] - \frac{1}{2}$$

Ahora, $E[Y_i] = \dfrac{\theta + \theta + 1}{2} = \theta + \dfrac{1}{2}$, por lo tanto

$$E[\hat{\theta}_1] = \frac{1}{n}n\left(\theta + \frac{1}{2}\right) - \frac{1}{2} = \theta.$$

Por lo tanto, $\hat{\theta}_1$ es insesgado. Para estudiar $\hat{\theta}_2$ necesitamos la densidad de $Y_{(n)}$. La función de distribución es

$$F(y) = P(Y \leq y) = \int_{\theta}^{y} 1 dt = y - \theta.$$

Por lo tanto la densidad será

$$f_{Y_{(n)}}(y) = n[F(y)]^{n-1} f(y) = n[y - \theta]^{n-1}.$$

Ahora calculemos el valor esperado del máximo

$$E[Y_{(n)}] = \int_{\theta}^{\theta+1} yn(y - \theta)^{n-1} dy$$

Si realizamos el cambio de variable $t = y - \theta$, entonces nos queda

$$= \int_0^1 (t + \theta)nt^{n-1}dt = \int_0^1 nt^n dt + \int_0^1 \theta nt^{n-1}dt = n\frac{t^{n+1}}{n+1}|_0^1 + \theta t^n|_0^1 = \frac{n}{n+1} + \theta.$$

Por lo tanto

$$E[\hat{\theta}_2] = E\left[Y_{(n)} - \frac{n}{n+1}\right] = \theta.$$

Con lo que se demuestre que nuestro segundo estimador también es insesgado.

b) Para calcular la eficiencia, necesitamos las varianzas de ambos estimadores. Para ello asumiremos que la muestra es independiente.

$$V(\hat{\theta_1}) = V(\bar{Y} - 1/2) = V\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(Y_i).$$

Y $V(Y_i) = \dfrac{(\theta+1-\theta)^2}{12} = \dfrac{1}{12}$, por lo que queda

$$V(\hat{\theta_1}) = \frac{1}{12n}.$$

Por otro lado es claro que $V(\hat{\theta_2}) = V(Y_{(n)})$, entonces debemos calcular $E[Y_{(n)}^2]$.

$$E[Y_{(n)}^2] = \int_{\theta}^{\theta+1} y^2 n(y-\theta)^{n-1}dy = \int_0^1 (t+\theta)^2 nt^{n-1}dt = \frac{n}{n+2} + \frac{2\theta n}{n+1} + \frac{n\theta^2}{n}.$$

Luego

$$V(Y_{(n)}) = E[Y_{(n)}^2] - E[Y_{(n)}]^2 = \frac{n}{n+2} + \frac{2\theta n}{n+1} + \frac{n\theta^2}{n} - \left(\theta + \frac{n}{n+1}\right)^2$$

$$= \frac{n}{n+2} + \frac{2\theta n}{n+1} + \theta^2 - \theta^2 - \frac{2n\theta}{n+1} - \frac{n^2}{(n+1)^2} = \frac{n}{(n+1)^2(n+2)}.$$

Finalmente, entonces

$$\operatorname{eff}(\hat{\theta_1}, \hat{\theta_2}) = \frac{V(\hat{\theta_2})}{V(\hat{\theta_1})} = \frac{\dfrac{n}{(n+1)^2(n+2)}}{\dfrac{1}{12n}} = \frac{12n^2}{(n+1)^2(n+2)}.$$

1.2. Consistencia.

<u>Definición</u>: Se dice que $\hat{\theta_n}$ es un estimador consistente de θ si, para cualquier número positivo ε,

$$\lim_{n\to\infty} P(|\hat{\theta_n} - \theta| \leq \varepsilon) = 1$$

o

$$\lim_{n\to\infty} P(|\hat{\theta_n} - \theta| \geq \varepsilon) = 0.$$

<u>Teorema</u>: Un estimador insesgado $\hat{\theta_n}$ de θ constituye un estimador consistente de θ si

$$\lim_{n\to\infty} V(\hat{\theta_n}) = 0.$$

<u>Ejemplo</u>. Suponga que Y_1, Y_2, ..., Y_n denota una muestra aleatoria de una distribución uniforme en el intervalo $(\theta, \theta+1)$. Sean

$$\hat{\theta_1} = \bar{Y} - \frac{1}{2} \text{ y } \hat{\theta_2} = Y_{(n)} - \frac{n}{n+1}.$$

Demuestre que $\hat{\theta_1}$ y $\hat{\theta_2}$ son estimadores consistentes de θ.

<u>Solución</u>.
Como ya tenemos las varianzas de ambos estimadores, calculemos sus límites para demostrar que son consistentes

$$\lim_{n\to\infty} V(\hat{\theta_1}) = \lim_{n\to\infty} \frac{1}{12n} = 0.$$

y

$$\lim_{n\to\infty} V(\hat{\theta_1}) = \lim_{n\to\infty} \frac{n}{(n+1)^2(n+2)} = 0.$$

Por lo tanto ambos estimadores son consistentes.

1.3. Suficiencia.

<u>Definición</u>: Sea Y_1, Y_2, ..., Y_n una muestra aleatoria de una distribución de probabilidad con un parámetro θ cuyo valor se desconoce. Se dice que el estadístico $U = g(Y_1, Y_2, ..., Y_n)$ es suficiente para θ si la distribución condicional de Y_1, Y_2, ..., Y_n, dado U, no depende de θ.

<u>Definición</u>: Sean y_1, y_2, ..., y_n observaciones muestrales tomadas de las variables aleatorias correspondientes Y_1, Y_2, ..., Y_n, cuya distribución depende un parametro θ. Si Y_1, Y_2, ..., Y_n son variables discretas, la verosimilitud de la muestra, $L(y_1, y_2, ..., y_n|\theta)$, se define como la probabilidad conjunta de y_1, y_2, ..., y_n. Si Y_1, Y_2, ..., Y_n son variables aleatorias continuas, la verosimilitud $L(y_1, y_2, ..., y_n|\theta)$ se define como la densidad conjunta evaluada en y_1, y_2, ..., y_n.

<u>Teorema</u>: Si U es un estadístico basado en la muestra aleatoria Y_1, Y_2, ..., Y_n, entonces U es un estadístico suficiente para la estimación del parámetro θ si y sólo si la verosimilitud $L(\theta) = L(y_1, y_2, ..., y_n|\theta)$ se puede factorizar en dos funciones no negativas

$$L(y_1, y_2, ..., y_n|\theta) = g(u, \theta)h(y_1, y_2, ..., y_n)$$

donde $g(u, \theta)$ es una función de u y θ, mientras que $h(y_1, y_2, ..., y_n)$ no es función de θ.

<u>Ejemplo</u>. Sea $Y_1, Y_2, ..., Y_n$ una muestra aleatoria de una distribución de Poisson con parámetro λ. Demuestre mediante condicionamiento que $\sum_{i=1}^{n} Y_i$ es suficiente para λ.

<u>Solución</u>.
La verosimilitud no es mas que

$$L(y_1, y_2, ..., y_n|\lambda) = p(y_1, y_2, ..., y_n|\lambda) = p(y_1|\lambda)p(y_2|\lambda)...p(y_n|\lambda)$$

$$= \frac{\lambda^{y_1}e^{-\lambda}}{y_1!}\frac{\lambda^{y_2}e^{-\lambda}}{y_2!}...\frac{\lambda^{y_n}e^{-\lambda}}{y_n!} = \frac{\lambda^{\sum y_1}e^{-n\lambda}}{\prod y_i!}.$$

Entonces, tomando $g(\sum y_1, \lambda) = \lambda^{\sum y_1}e^{-n\lambda}$, y $h(y_1, y_2, ..., y_n) = \frac{1}{\prod y_1}$, entonces por el teorema anterior, entonces $\hat{\lambda} = \sum y_i$ es un estimador suficiente para λ.

<u>Ejemplo</u>. Si Y_1, Y_2, ..., Y_n denotan variables aleatorias independientes, que tienen una distribución idéntica que pertenece a la familia de distribuciones tipo potencia con parámetros α y θ, entonces, si $\alpha, \theta > 0$,

$$f(y|\alpha, \theta) = \begin{cases} y^{\alpha-1}\dfrac{\alpha}{\theta^\alpha} & , \quad 0 \leq y \leq \theta \\ 0 & , \quad \text{en otro punto} \end{cases}$$

Si se conoce el valor de θ, demuestre que $\prod_{i=1}^{n} Y_i$ es suficiente para α.

<u>Solución</u>.
La verosimilitud para α y θ es

$$L(y_1, y_2, ..., y_n|\alpha) = f(y_1, y_2, ..., y_n|\alpha) = f(y_1|\alpha)f(y_2|\alpha)...f(y_n|\alpha)$$

$$= y_1^{\alpha-1}\frac{\alpha}{\theta^\alpha}y_2^{\alpha-1}\frac{\alpha}{\theta^\alpha}...y_n^{\alpha-1}\frac{\alpha}{\theta^\alpha} = \left(\prod y_i\right)^{\alpha-1}\frac{\alpha^n}{\theta^{n\alpha}}.$$

Ahora si se toma $g(\prod y_i, \alpha) = (\prod y_i)^{\alpha-1}\frac{\alpha^n}{\theta^{n\alpha}}$ y $h(y_1, y_2, ..., y_n) = 1$, entonces se tiene que $\prod_{i=1}^{n} Y_i$ es un estimador suficiente para α.