

Estimación puntual y por intervalo

1.2 Propiedades deseables de los estimadores puntuales

Se conoce la familia de distribuciones a partir de la cual se obtiene la muestra, pero no puede identificarse el miembro de interés de esta, ya que no se conoce el valor del parámetro. Este último tiene que estimarse con base en los datos de la muestra.

El estimador de un parámetro θ debe tener una distribución de muestreo concentrada alrededor de θ y la varianza del estimador debe ser la menor posible. Para ampliar las propiedades anteriores, considérese la siguiente. Sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n proveniente de una distribución con función de densidad $f(x, \theta)$, y sea $T = u(X_1, X_2, \dots, X_n)$ cualquier estadística. El problema es encontrar una función u que sea la que proporcione la mejor estimación de θ . Al buscar el mejor estimador de θ se hará uso de una cantidad muy importante que recibe el nombre de error cuadrático medio de un estimador.

Definición 1.1 Sea T cualquier estimador de un parámetro desconocido θ . Se define el error cuadrático medio de T como el valor esperado del cuadrado de la diferencia entre T y θ .

Para cualquier estadística T , se denotará el error cuadrático medio por $ECM(T)$; de esta forma

$$ECM(T) = E(T - \theta)^2.$$

Puede verse la razón del porque el error cuadrático medio es una cantidad importante para enjuiciar a los posibles estimadores de θ mediante el desarrollo de $ECM(T) = E(T - \theta)^2$; este es,

$$\begin{aligned} ECM(T) &= E(T^2 - 2\theta T + \theta^2) \\ &= E(T^2) - 2\theta E(T) + \theta^2 \\ &= \text{Var}(T) + E^2(T) - 2\theta E(T) + \theta^2 \\ &= \text{Var}(T) + [\theta - E(T)]^2. \end{aligned}$$

El error cuadrático medio de cualquier estimador es la suma de dos cantidades no negativas: una es la varianza del estimador y la otra es el cuadrado del sesgo del estimador. Estas dos cantidades se encuentran relacionadas en forma directa con las propiedades deseables de un estimador. La varianza de un estimador debe ser lo más pequeña posible mientras que la distribución de muestreo debe concentrarse del valor del parámetro. Por lo que debemos seleccionar el mejor estimador de θ ; es decir, la estadística que tenga el error cuadrático medio más pequeño posible entre todos los estimadores factibles de θ .

1.2.1 Estimadores insesgados

El término $\theta - E(T)$ recibe el nombre de sesgo del estimador.

Definición 1.2 Se dice que la estadística $T = u(X_1, X_2, \dots, X_n)$ es un estimador insesgado del parámetro θ , si $E(T) = \theta$ para todos los posibles valores de θ . De esta forma, para cualquier estimador insesgado de θ , la distribución de muestreo de T se encuentra centrada alrededor de θ y $ECM(T) = Var(T)$.

En la sección 7.4, demostramos que sin importar la distribución de la población de interés, $E(\bar{X}) = \mu$. Por lo tanto, la media muestral es un estimador insesgado de la media de la población μ para todo los valores de μ . De hecho, si X_1, X_2, \dots, X_n es una muestra aleatoria de la distribución de X con media μ , entonces cualquier X_i de la muestra es un estimador insesgado de μ , dado que $E(X_i) = \mu$ para todo $i = 1, 2, \dots, n$. Además si una estadística T es cualquier combinación lineal de las variables aleatorias de la muestra de manera tal que

$$T = a_1 X_1 + a_2 X_2 + \dots + a_n X_n$$

en donde $\sum_{i=1}^n a_i = 1$, entonces T es un estimador insesgado de μ dado que

$$\begin{aligned} E(T) &= \sum_{i=1}^n a_i E(X_i) \\ &= \sum_{i=1}^n a_i \mu \\ &= \mu. \end{aligned}$$

En la sección 7.5 se demostró que si la varianza muestral S^2 está dada por $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2 / (n - 1)$. Entonces, cuando se muestrea una distribución normal, $E(S^2) = \sigma^2$. Lo que demostraremos que es insesgado sin importar cuál sea la distribución de la población de interés. Sea X_1, X_2, \dots, X_n una muestra aleatoria de alguna distribución con una función de densidad no específica. De esta manera, $E(X_i) = \mu$ y $Var(X_i) = \sigma^2$ para toda $i = 1, 2, \dots, n$. Entonces,

$$\begin{aligned} E(S^2) &= E \left[\sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n - 1} \right] \\ &= (n - 1)^{-1} E \left\{ \sum_{i=1}^n [(X_i - \mu) - (\bar{X} - \mu)]^2 \right\} \\ &= (n - 1)^{-1} E \left\{ \sum_{i=1}^n [(X_i - \mu)^2 - 2(X_i - \mu)(\bar{X} - \mu) + (\bar{X} - \mu)^2] \right\} \\ &= (n - 1)^{-1} \left[\sum_{i=1}^n E(X_i - \mu)^2 - 2n E(\bar{X} - \mu)^2 \right] \end{aligned}$$

pero por definición $E(X_i - \mu)^2 = Var(X_i) = \sigma^2$ y $E(\bar{X} - \mu)^2 = Var(\bar{X}) = \sigma^2 / n$. Por lo tanto,

$$\begin{aligned} E(S^2) &= (n - 1)^{-1} \left[\frac{n\sigma^2 - (n\sigma^2)}{n} \right] \\ &= \frac{\sigma^2(n - 1)}{n - 1} \\ &= \sigma^2. \end{aligned}$$

En otras palabras, S^2 es un estimador insesgado de σ^2 sólo cuando el divisor es igual a $n - 1$. Es la razón por la que se divide por $n - 1$ en lugar de n . S no hará un estimador insesgado de σ .

1.3 Estimadores consistentes