Estimación Puntual

Cuando se trata de caracterizar a una población de acuerdo con cierto parámetro (o parámetros) con base en una muestra aleatoria, usualmente se desconoce la distribución de probabilidad de la variable de interés. Para aproximar, a partir de la muestra, el valor del parámetro poblacional se utilizan los valores colectados para, a través de una función, construir que se denomina estimador puntual.

Estimador puntual
Sea $\{X_i, X_2, \dots, X_n\}$ una muestra aleatoria de cierta distribución de probabilidad común con cierto parámetro θ, denotado como $X_i \sim D_\theta$. Un **estimador puntual** del parámetro θ es una función de la muestra: $$\hat{\Theta} = g(X_1, X_2, \dots, X_n)$$ Una **estimación puntual** de θ es el valor particular al evaluar la función, esto es, $$\theta = g(x_1, x_2, \dots, x_n)$$

Un parámetro es un valor fijo y usualmente desconocido. Un estimador es una variable aleatoria, por tanto, el valor que tome en una estimación está en función de la muestra con la que se obtiene.

Dado que un estimador es una función de variables aleatorias, entonces es en sí mismo una variable aleatoria. Dos propiedades deseables en un estimador que sea consistente y sea insesgado. La consistencia se refiere a que conforme el tamaño de muestra se incremente, el valor converge al parámetro verdadero. Ser insesgado se refiere a que la esperanza del estimador sea igual al parámetro de la distribución.

Ejemplo 1. *Una moneda tiene probabilidad p de ocurrir águila cuando se lanza y el resultado es aleatorio. En este caso el parámetro de interés es $\theta = p$.*

Si la moneda se lanza n veces y si x es la cantidad de águilas observadas, un estimador de p es la proporción de águilas observadas, esto es:

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

Si al lanzar la moneda cada vez el resultado se codifica como 1 si es águila y 0 si es sol, entonces:

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

Donde:

$$x_i \begin{cases} 1 \text{ si es águila} \\ 0 \text{ si es sol} \end{cases}$$

por tanto:

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

Esto es, la proporción es un promedio de una muestra cuyosvalores son 0 o 1.

Si en $n = 200$ lanzamientos se obtienen $x = 108$ águilas, entonces:

$$\hat{p} = \frac{108}{200} = 0.54$$

es la probabilidad estimada de que ocurra águila.

Si en el ejemplo 1 se repite el experimento m veces en cada uno obtendremos una estimación puntual y esto permite construir una distribución empírica de $\hat{p}$. Para mostrar esto, mediante simulación, se repite el experimento m = 1000 veces con n = 500. El histograma resultante se muestra en la figura 1.

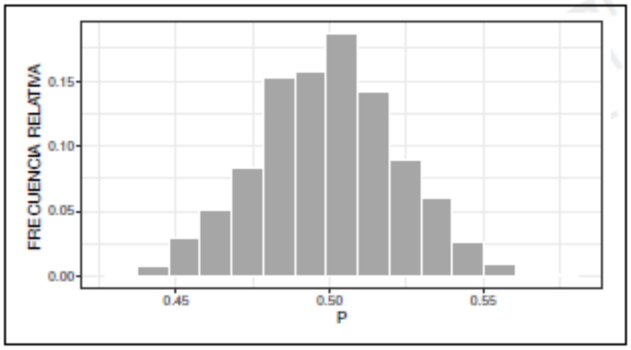

Figura 1 Distribución empírica de $\hat{p}$.

Estimación puntual de los parámetros p, μ y σ^2

Tres de los parámetros más utilizados para caracterizar una población son la proporción, la media y la varianza.

Estimador puntual de p, μ y σ^2
Sea $\{X_i, X_2, \ldots, X_n\}$ una muestra aleatoria donde $x_i \sim x$ cuya densidad es $$p(x;p) = p^x(1-p)^{1-x}1_{\{0,1\}}, p \in (0,1)$$ El estimador puntual de la proporción p es $$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} x_i$$ Sea $\{X_i, X_2, \ldots, X_n\}$ una muestra aleatoria de una variable aleatoria continua. Los estimadores puntuales de la media $\boldsymbol{\mu}$ y varianza $\boldsymbol{\sigma^2}$ son, respectivamente

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

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

El estimador de la proporción es también un promedio, solo que los valores provienen de una variable aleatoria discreta que toma solo los valores de 0 o 1. Si bien se definen los estimadores de media y varianza de una variable aleatoria continua, en el caso de la discreta también es factible obtener estas estimaciones.

Ejemplo 2. *En cierto sector residencial se pretende construir una plaza con áreas de restaurantes, ropa y gimnasio. Se realiza una encuesta basada en una muestra de* $n = 150$ *personas profesionistas en un rango de edad de 25 a 45 años, de los cuales* $n_m = 80$ *son mujeres y* $n_h = 70$ *son hombres. La encuesta consiste en tres preguntas: si se está de acuerdo en la construcción, el ingreso mensual.*

$x_m = 70$ *de las mujeres y* $x_h = 56$ *de los hombres están de acuerdo en la construcción de la plaza. La proporción estimada para cada caso es, respectivamente,*

$$\hat{p}_m = \frac{70}{80} = 0.875$$

$$\hat{p}_h = \frac{56}{70} = 0.8$$

En el siguiente cuadro se resumen los resultados obtenidos en relación con la proporción de hombres y mujeres que están de acuerdo, el promedio y desviación estándar del ingreso.

PAR	M	H
p	0.875	0.8
μ	25,000.00	30,000.00
σ	4,000.00	3,000.00

Tabla 1 Resultados encuesta

En el ejemplo 2 se tienen las estimaciones puntuales de dos segmentos de población, a saber, mujeres y hombres. La estimación de la proporción de los que están a favor es mayor, numéricamente, la de mujeres que la de hombres. Se registra un ingreso mayor medio en hombres que en mujeres.

Referencias:

Hoel, P. G. (1984). Elementary Statistics. John Wiley & Sons.

Valdés, J. R. (2019). Estadística. Notas de clase.