Estimación por punto

ESCUELA POLITÉCNICA
SUPERIOR DE CÓRDOBA
Universidad de Córdoba
DEPARTAMENTO DE ESTADÍSTICA

Planteamiento general de un problema de estimación paramétrica

Sea X una característica medida en una población, siendo la función de densidad de X función de uno o varios parámetros θ desconocidos, es decir $f(x) = f(x;\theta)$, y cuya forma puede ser conocida (salvo θ) o no.

En la población se aplica un diseño experimental de muestreo aleatorio simple, es decir, se considera una muestra aleatoria genérica $\vec{X} = \left(X_1, X_2, \ldots, X_n\right)$. Al tomar una muestra concreta (realización), se obtienen valores numéricos $\left(x_1, x_2, \ldots, x_n\right)$ a partir de los cuales se realiza la estimación de θ.

Un estimador (por punto) es una función de la muestra genérica $\vec{X} = \left(X_1, X_2, \ldots, X_n\right)$, es decir un estadístico $\widehat{\theta}\left(\vec{X}\right)$ o simplemente $\widehat{\theta}$.

La estimación por punto tiene como objetivo proporcionar valores numéricos de algún parámetro (o de varios) desconocido de una población en la que se mide una característica X. Para ello, una vez obtenidos los valores numéricos de la realización de la muestra, basta sustituirlos en la expresión del estimador $\widehat{\theta}$.

Nota: Al realizar estimaciones puntuales no se controla el error que se puede cometer en dicha estimación. La calidad de la estimación se mide por las propiedades que cumple el estimador.

Propiedades de los estimadores

- Insesgadez

$\hat{\theta}$ es un estimador insesgado de θ (o centrado) sí verifica que:

$E\left(\hat{\theta}\right) = E\left[t\left(\vec{X}\right)\right] = \theta$ (θ parámetro poblacional desconocido)

Ej: Dado una m.a.s $\vec{X} = \left(X_1, X_2, \ldots, X_n\right)$ de una característica $X \in D\left(\mu, \sigma^2\right)$

- $\bar{X}$ estimador insesgado de μ $\quad E\left[\bar{X}\right] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E\left[X_i\right] = \frac{1}{n} n\mu = \mu$
- S^2 estimador sesgado de σ^2 $\quad E\left[S^2\right] = \sigma^2 - \frac{\sigma^2}{n}$ Sesgo
- $\bar{S}^2$ estimador insesgado de σ^2

$$E\left[\bar{S}^2\right] = \frac{n}{n-1} E\left[S^2\right] = \frac{n}{n-1}\left(\sigma^2 - \frac{\sigma^2}{n}\right) = \frac{n}{n-1}\left[\sigma^2\left(1 - \frac{1}{n}\right)\right] = \frac{\cancel{n}}{\cancel{n-1}}\sigma^2 \frac{\cancel{n-1}}{\cancel{n}} = \sigma^2$$

- $\hat{\theta}$ es un estimador asintóticamente insesgado de θ si:

$$E\left(\hat{\theta}\right)\xrightarrow[n\to\infty]{}\theta$$

Ejemplo:

$$E\left(S^2\right)=\sigma^2-\frac{\sigma^2}{n}\to\sigma^2$$

- Consistencia

Un estimador es consistente si al aumentar *n* indefinidamente, converge en probabilidad al parámetro estimado:

$$\hat{\theta}(X_1,X_2,\ldots,X_n)\xrightarrow[n\to\infty]{P}\theta \text{ es decir } P\left(\left|\hat{\theta}(\vec{X})-\theta\right|<\delta\right)\to 1$$

Caracterización:

Si $\hat{\theta}$ es estimador asintóticamente insesgado de θ y se verifica que $\mathrm{V}(\hat{\theta})\xrightarrow[n\to\infty]{}0$, se dice consistente (de θ).

Ejemplos:

$$E\left(S^2\right)\to\sigma^2;\quad V\left(S^2\right)=\frac{\mu_4-\sigma^4}{n}\to 0$$

$$E\left(\bar{X}\right)\to\mu;\quad V\left(\bar{X}\right)=\frac{\sigma^2}{n}\to 0$$

- Si dos estimadores $\hat{\theta}_1$ y $\hat{\theta}_2$ son estimadores insesgados de θ, se dice que $\hat{\theta}_1$ es más eficiente que $\hat{\theta}_2$ sí:

$$V(\theta_1) < V(\theta_2)$$

- Un estimador $\hat{\theta}$ que sea insesgado, consistente y el mas eficiente se denomina mejor estimador insesgado para θ. Suele denominarse también el estimador *BUE* (Best Unbiased Estimator).

ESCUELA POLITÉCNICA SUPERIOR DE CÓRDOBA
Universidad de Córdoba
DEPARTAMENTO DE ESTADÍSTICA

Cota de Frechet-Cramer-Rao (*FCR*)

Sea una muestra a.s. $\vec{X} = \left(X_1, X_2, \ldots, X_n\right)$ obtenida de una población con f.d.d. $f(x) = f(x,\theta)$ en las que se cumplen ciertas condiciones de tipo matemático (p.e. el espacio muestral no depende de θ). Sea $\hat{\theta} = t(\vec{X})$ un estimador insesgado de θ con

$$L(\theta) = L(\theta, \vec{X}) = f(x_1,\theta) f(x_2,\theta) \cdots f(x_n,\theta)$$

su función de verosimilitud.

Entonces se verifica que la varianza de $\hat{\theta}$ está acotada inferiormente por la cota de *FCR(*θ*)*.

$$V(\theta) \geq \frac{1}{E\left[\left(\frac{\partial \ln L(\theta, \vec{X})}{\partial \theta}\right)^2\right]} \equiv FCR(\theta)$$

Si existe un estimador cuya varianza coincida con esta cota, se dice que $\hat{\theta}$ es el estimador mas eficiente de θ.

Obtención de estimadores. Método de la máxima verosimilitud

Dada una muestra aleatoria simple $\vec{X}$ de una población, el método consiste en maximizar su función de verosimilitud.

$$f(\vec{X}) = L(\theta) = L(\theta, \vec{X}) = f(x_1, \theta) f(x_2, \theta) \cdots f(x_n, \theta)$$

siendo $f(\vec{X})$ la f.d.d. de la v.a. $\vec{X}$.

Para ello, basta con obtener la solución de:

$$\frac{\partial L(\theta, \vec{X})}{\partial \theta} = 0$$

Pero si una función cualquiera alcanza un máximo en un punto, la función **ln** de ésta también alcanza un máximo en el mismo punto, con lo que basta resolver:

$$\frac{\partial \ln L(\theta, \vec{X})}{\partial \theta} = 0$$

Apéndices - 1

La media muestral es el estimador de máxima verosimilitud para la media poblacional de una variable normalmente distribuida.

$$L\left(\bar{X},\theta\right) = f(x_1,\theta)f(x_2,\theta)\cdots f(x_n,\theta)$$

$$L\left(\vec{X},\mu\right) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\frac{\left(x_1-\mu\right)^2}{\sigma^2}}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\frac{\left(x_2-\mu\right)^2}{\sigma^2}}\cdots\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\frac{\left(x_n-\mu\right)^2}{\sigma^2}} = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{\left(x_i-\mu\right)^2}{\sigma^2}}$$

$$\ln L\left(\vec{X},\mu\right) = \ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n - \frac{1}{2}\sum_{i=1}^{n}\frac{\left(x_i-\mu\right)^2}{\sigma^2}$$

$$\frac{\partial \ln L\left(\vec{X},\mu\right)}{\partial\mu} = +\frac{1}{2}2\sum_{i=1}^{n}\frac{\left(x_i-\mu\right)}{\sigma^2} = 0 \Leftrightarrow \sum_{i=1}^{n}\left(x_i-\mu\right) = 0 \Leftrightarrow \sum_{i=1}^{n}x_i - n\mu = 0 \Leftrightarrow \mu = \frac{1}{n}\sum_{i=1}^{n}x_i$$

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

Apéndice - 2

La media muestral es el estimador de máxima verosimilitud para el parámetro λ en una población de tipo Poisson.

$$L\left(\vec{X},\lambda\right)=f(x_1,\lambda)f(x_2,\lambda)\cdots f(x_n,\lambda)$$

$$L\left(\vec{X},\lambda\right)=\frac{e^{-\lambda}\lambda^{x_1}}{x_1!}\frac{e^{-\lambda}\lambda^{x_2}}{x_2!}\cdots\frac{e^{-\lambda}\lambda^{x_n}}{x_n!}=\frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n}x_i}}{\prod_{i=1}^{n}x_i!}$$

$$\ln L\left(\vec{X},\lambda\right)=-n\lambda+\sum_{i=1}^{n}x_i\ln\lambda-\sum_{i=1}^{n}\ln x_i!$$

$$\frac{\partial\ln L\left(\vec{X},\lambda\right)}{\partial\lambda}=-n+\frac{1}{\lambda}\sum_{i=1}^{n}x_i=0\Leftrightarrow\sum_{i=1}^{n}x_i=n\lambda\Leftrightarrow\lambda=\frac{1}{n}\sum_{i=1}^{n}x_i$$

Con lo que $\hat{\lambda}=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{x}$

Ejemplos

La duración de una consulta en un cajero automático es una magnitud aleatoria X cuya distribución es aproximadamente Normal. Se toman datos eligiendo al azar cinco consultas, resultando los tiempos de uso (en segundos) siguientes: 44, 36, 40, 45 y 25. Se desea estimar la duración media de una consulta.

Puesto que, como hemos visto, el estimador de máxima verosimilitud para estimar μ en una población Normal es la media muestral, una estimación del tiempo de consulta medio es la media de los datos muestrales, es decir 38 segundos.

La distribución de Poisson está asociada a sucesos raros, o sea cuya probabilidad de ocurrencia es pequeña. En una plantación de tabaco se seleccionan doce parcelas del mismo tamaño y en cada una se observa el número de plantas afectadas por moho azul, obteniéndose los siguientes resultados: 5, 2, 10, 6, 3, 3, 5, 8, 1, 10, 4 y 3, y se desea estimar el número medio de plantas enfermas.

Como también acabamos de ver, la media muestral es un estimador del parámetro λ de una distribución de Poisson, además, también sabemos que en una distribución de Poisson, la media coincide con el parámetro, con lo que solo bastará calcular la media de la muestra, en este caso 5 plantas enfermas, para obtener una estimación del parámetro desconocido λ.

Estimación por punto

ESCUELA POLITÉCNICA SUPERIOR DE CÓRDOBA
Universidad de Córdoba
DEPARTAMENTO DE ESTADÍSTICA