Universidad Rey Juan Carlos

T4. MÉTODOS DE ESTIMACIÓN

ESTADÍSTICA 2 – URJC – GRADO DE ADE Y ECO – PACORABADAN.COM

ESQUEMA DEL TEMA

- 4.0 Presentación.
- 4.1. Método de Estimación de la máxima veroslimilitud.
- 4.2. Propiedades de los estimadores máximo verosímiles
- 4.3. Método de estimación de los momentos.
- 4.4. Estimación por el método de los mínimos cuadrados.

4.0. PRESENTACIÓN

- En Tema 3, hemos visto cuales deben ser las propiedades que debe tener un estimador para considerarlo “bueno”.
- Ahora, Buscamos métodos que nos permitan elegir de los infinitos estimadores posibles, el más razonable (VEROSIMIL), para luego ver si cumple o no las propiedades del Tema 3.
- Vamos a ver tres métodos que se basan en distintos supuestos:
  1. **Máxima verosimilitud**: *Conocemos la distribución* de probabilidad de la población.
  2. **Método de los momentos**: conocemos los *momentos de la distribución* poblacional.
  3. **Mínimos cuadrados**: *minimizamos la distancia euclídea* entre el valor conocido de una observación y la estimación puntual (procedimiento similar a la regresión que estudiamos en Estadística 1).

4.1. MÉTODO DE ESTIMACIÓN POR MÁXIMA VEROSIMILITUD

- **Fundamento intuitivo**: de todos los posibles sucesos que pueden tener lugar, suponemos que aparecerá el más probable.
  - **INCONVENIENTE**: Pretende hallar la Moda de $\boldsymbol{L(X,\theta)}$, **estimador mas pobre que la media o la mediana… pero si n es grande la moda suele acercarse a la mediana (MARTÍN PLIEGO, P.116, 2005)**
- **Planteamiento operativo**: conocida la función de probabilidad $F(X,\theta)$ construimos la función de verosimilitud $L(X,\theta)$

$$\xi: F(X,\theta)\begin{cases} VAC \;\; F(X,\theta)=f(x)\;(densidad) \\ VAD \;\; F(X,\theta)=P(\xi=x)\;(cuantía)\end{cases}, siendo\ \boldsymbol{\theta\ desconocido\ con\ campo\ de\ variación\ \Theta}, \leftrightarrow \theta\in\Theta$$

- **Función de verosimilitud** $L(X,\theta)$ es la función conjunta de cuantía o densidad de la muestra (donde θ aparece como incógnita no probabilística).
- **Criterio de selección** del estimador $\hat{\theta}$ del parámetro θ:

$$\boxed{\hat{\theta}\ tal\ que\ L\left(X,\hat{\theta}\right)=\max_{\theta\in\Theta} L(X,\theta)}$$

4.1. MÉTODO DE ESTIMACIÓN POR MÁXIMA VEROSIMILITUD

$$\boxed{\hat{\theta} \; tal \; que \; L\left(X;\hat{\theta}\right) = \max_{\theta \in \Theta} L(X,\theta)}$$

- $L(X,\theta)$ es multiplicativa (por ser el resultado de la intersección de la aparición de los x_i elementos de X), para facilitar el método de optimización (derivar) tomamos logaritmos.

$$\boxed{\hat{\theta} \; tal \; que \; \ln L\left(X,\hat{\theta}\right) = \max_{\theta \in \Theta} \ln L(X,\theta)}$$

- Derivamos $\ln L(X,\theta)$ rescto de θ e igualamos a cero: **CN** de mínimo/máximo: despejamos y obtenemos una expresión de θ, en función de los valores muestrales (se prescinde de soluciones que sean constantes).

$$CN : \frac{\partial \ln L(X,\theta)}{\partial \theta} = 0 \rightarrow \hat{\theta}$$

- **CN*S*** de máximo: $\frac{\partial^2 \ln L(X,\theta)}{\partial^2 \theta} < 0 \rightarrow \hat{\theta}$ es estimador maximoverosimil

EJEMPLO 2

Calcúlese el estimador máximo verosímil del parámetro p de la distribución $B(1;\, p)$, en muestras aleatorias simples de tamaño n.

La función de verosimilitud es

$$L(\mathbf{X};\, p) = p^{\sum_{i=1}^{n} x_i} (1-p)^{n - \sum_{i=1}^{n} x_i}$$

Tomando logaritmos, derivando respecto a p e igualando la derivada a cero

$$\frac{\partial \ln L(\mathbf{X};\, p)}{\partial p} = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{n - \sum_{i=1}^{n} x_i}{1-p} = 0$$

la solución de la ecuación es

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

Comprobamos que es un máximo

$$\frac{\partial^2 \ln L(\mathbf{X};\, p)}{\partial p^2} = \frac{-\sum_{i=1}^{n} x_i}{p^2} - \frac{n - \sum_{i=1}^{n} x_i}{(1-p)^2}$$

expresión que particularizada para $\hat{p} = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ conduce a

$$\left[\frac{\partial^2 \ln L(\mathbf{X};\, p)}{\partial p^2}\right]_{p=\hat{p}} = -\frac{n}{\bar{x}} - \frac{n}{1-\bar{x}}$$

menor que cero, condición de máximo.

(Martín Pliego,p.112, 2005)

EJEMPLO 3

En la distribución $N(\mu;\sigma)$, con varianza conocida, el parámetro μ desconocido se estima mediante el método de la máxima verosimilitud, en muestras aleatorias simples de tamaño n. La función de verosimilitud es

$$L(\mathbf{X};\mu) = f(x_1;\mu,\sigma)\cdots f(x_n;\mu,\sigma) =$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \cdots \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x_n-\mu)^2}{2\sigma^2}} =$$

$$= \frac{1}{(\sigma^2)^{\frac{n}{2}}(2\pi)^{\frac{n}{2}}}\, e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}}$$

su logaritmo es igual a

$$\ln L(\mathbf{X};\mu) = -\frac{n}{2}\ln\sigma^2 - \frac{n}{2}\ln 2\pi - \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}$$

y su derivada respecto al parámetro μ

$$\frac{\partial \ln L(\mathbf{X};\mu)}{\partial\mu} = \frac{\sum_{i=1}^{n}(x_i-\mu)}{\sigma^2} = 0.$$

Resuelta la ecuación llegamos a que el estimador máximo verosímil de μ es la media de la muestra

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

La condición de máximo se verifica, pues

$$\left[\frac{\partial^2 \ln L(\mathbf{X};\mu)}{\partial\mu^2}\right]_{\mu=\hat{\mu}} = \frac{-n}{\sigma^2} < 0.$$

(Martín Pliego,pp.112-113; 2005)

4.2. PROPIEDADES DE LOS ESTIMADORES MV (MAXIMOVEROSÍMILES)

PROPIEDAD	Comportamiento	Observaciones
INSESGADEZ	$\hat{\theta}_{MV}$ asintóticamente insesgados siempre	(en algunos casosextrictamente)
CONSISTENCIA	Sí. Por eso, son asintóticamente insesgados.	
EFICIENCIA	No siempre, pero…	**CN: de estimador eficiente** que se haya obtenido por MV
NORMALIDAD Y EFICIENCIAS ASINTÓTICAS	$\hat{\theta}_{MV}$ son asintóticamente normales $\hat{\theta}_{MV}$ son asintóticamente eficientes	
SUFICIENCIA	Si T(X) es estadístico suficiente de $\theta \rightarrow \hat{\theta}_{MV} = f\big(T(X)\big)$	
INVARIANZA	$\hat{\theta}_{MV}$ bajo una transformación del parámetro/s.	

Repasar con Tema 3

4.3. ESTIMACIÓN POR EL MÉTODO DE LOS MOMENTOS

- Fundamento teórico:
  - $\boldsymbol{E(a_r) = \alpha_r}$ y además
  - **teorema de Khintchine** (a_r es estimador consistente de α_r).
- Planteamiento operativo: igualar los momentos poblacionales a los muestrales.
- En consecuencia:
  - Necesitamos forma un sistema con tantas ecuaciones como parámetros tenga la función de probabilidad
  - Cada ecuación del sistema corresponderá con la definición de momento.

118 ■ FUNDAMENTOS DE INFERENCIA ESTADÍSTICA

Planteamos el sistema de k ecuaciones con k incógnitas, número de parámetros, igualando cada momento poblacional, $\alpha_r(\theta_1, \ldots, \theta_k)$, a su correspondiente momento muestral a_r

$$\alpha_1(\theta_1, \ldots, \theta_k) = a_1$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$\alpha_k(\theta_1, \ldots, \theta_k) = a_k$$

la solución del sistema proporciona los estimadores de los parámetros

$$\theta_1^* = \theta_1^*(a_1, \ldots, a_k)$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$\theta_1^* = \theta_1^*(a_1, \ldots, a_k)$$

4.3. ESTIMACIÓN POR EL MÉTODO DE LOS MOMENTOS

EJEMPLO 6

Estimación del parámetro p en la distribución de Poisson $P(\lambda)$, en muestras aleatorias simples tamaño n

$$\alpha_1 = \lambda; \qquad a_1 = \overline{x}$$

luego

$$\lambda^* = \overline{x}\,.$$

4.3. ESTIMACIÓN POR EL MÉTODO DE LOS MOMENTOS

EJEMPLO 7

Estímese, por el método de los momentos, el parámetro θ, límite superior del intervalo de variación de la distribución $U(0;\theta)$, en muestras aleatorias simples de tamaño n.

La función de densidad de esta distribución uniforme es

$$f(x;\theta) = \frac{1}{\theta}\cdot$$

El momento de orden uno con respecto al origen es $\alpha_1 = \dfrac{\theta}{2}$, y la media muestral $\bar{x}$, igualando ambos momentos, el estimador resulta $\theta^* = 2\bar{x}$.

4.3. ESTIMACIÓN POR EL MÉTODO DE LOS MOMENTOS

EJEMPLO 8

En la distribución $N(\mu, \sigma)$ estimamos la media y la varianza por el método de los momentos. Tenemos que estimar dos parámetros, por lo cual necesitamos dos ecuaciones, resultado de igualar los momentos poblacionales de primer y segundo orden respecto al origen a los muestrales

$$\alpha_1(\mu, \sigma^2) = a_1$$

$$\alpha_2(\mu, \sigma^2) = a_2$$

El primer momento poblacional en la distribución normal es μ, y el segundo α_2 que, expresado en función de la varianza queda $\alpha_2 = \sigma^2 + \mu^2$. El sistema de ecuaciones es

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

$$\sigma^2 + \mu^2 = \frac{\sum_{i=1}^{n} x_i^2}{n}$$

y la solución

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

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

4.4. ESTIMACIÓN POR MÍNIMOS CUADRADOS.

- Fundamento teórico: No hay conocimiento de la población, y se pretende un “ajuste funcional (regresión)” a la nube de puntos muestral.
- Fundamento operativo: regresión tipo I (ver estadística 1), donde la variable no aleatoria son los parámetros y se pretende reducir el error cuadrático medio… Ver (Martín Pliego, p.120-121)

BIBLIOGRAFÍA

- Martín Pliego, Fundamentos de Inferencia Estadística, Editorial AC, 3ªEd, 2005
- Otros recursos en www.pacorabadan.com