Clase Auxilar 5: Métodos de Estimación y Propiedades de Estimadores

Profesora Nancy Lacourly
Auxiliares: Andrés Iturriaga J. - Héctor Olivero Q.

8 de abril de 2009

Estimación de Parámetros

Supongamos que estamos estudiando algún fenómeno de naturaleza aleatoria, por ejemplo el número de personas que llegan a una cola en un supermercado. Dado nuestro conocimiento del fenómeno con el que estamos lidiando, sabemos que sigue cierta distribución de probabilidad, la que conocemos salvo ciertos parámetros, en nuestro ejemplo, sabemos que la llegada de los clientes sigue una distribución de Poisson de parámetro λ que desconocemos.

El problema de estimación puntual se trata de encontrar los parámetros desconocidos dadas ciertas observaciones de la variable (fenómeno) que estamos estudiando.

Método de los Momentos

Una consecuencia de la ley fuerte de los grandes números es que si tenemos una muestra $\{X_n\}_{i=1}^n$ de una variable aleatoria X entonces se tiene que:

$$m_k = \bar{X}_n^k = \frac{1}{n}\sum_{i=1}^n X_i^k \xrightarrow{n\to\infty} \mathbb{E}(X^k) = \mu_k \quad \mathbb{P}\text{-c.s.}$$

Esto nos indica que un indicador razonable para el momento de orden k de la variable X es m_k. La idea anterior la sistematizamos en el siguiente método conocido como Método de los Momentos.

Método de los Momentos

Dada una variable X que sigue una distribución $F(\theta_1, \ldots, \theta_k)$ el método de los momentos se puede resumir en los siguientes pasos:

1. Relacionar los parámetros a estimar con los momentos teóricos de X. Esto generará ecuaciones de la forma:

$$\mu_i = H_i(\theta_1, \ldots, \theta_k) \text{ para } i = 1 \ldots k \tag{1}$$

2. Estimar μ_k por m_k es decir, plantear las ecuaciones:

$$m_i = H_i(\hat{\theta}_1, \ldots, \hat{\theta}_k) \text{ para } i = 1 \ldots k \tag{2}$$

3. Resolver el sistema (2) y encontrar $(\hat{\theta}_1, \ldots, \hat{\theta}_k)$.

Método de los Momentos: Ejemplo

Clase Auxilar 5

MA34B-01

Métodos de Estimación

Método de los Momentos

Método de Máxima Verosimilitud

Método Bayesiano

Propiedades Estimadores

Información de Fisher y Cota de Cramer Rao

Problemas Propuestos

Ejemplo:

Sea una m.a.s. $\{X_i\}_{i=1}^{n}$ obtenida de una distribución con densidad:

$$f(x,\theta) = \frac{1}{2-\theta}\chi_{[\theta,2]}$$

donde θ es desconocido y pertenece a $[0,2]$

1. Encuentre el estimador de los momentos de θ.

Método de los Momentos: Ejemplo

Solución

1. Calculamos el momento teórico:

$$\mathbb{E}(X) = \int_{\theta}^{2} \frac{x dx}{2-\theta} = \frac{2+\theta}{2} \tag{3}$$

2. Estimamos μ_1 por m_1 y tenemos:

$$m_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{2+\hat{\theta}}{2} \tag{4}$$

3. Despejamos:

$$\hat{\theta} = 2(\frac{1}{n}\sum_{i=1}^{n} X_i - 1) \tag{5}$$

Método de Máxima Verosimilitud

Clase Auxilar 5

MA34B-01

Métodos de Estimación

Método de los Momentos

Método de Máxima Verosimilitud

Método Bayesiano

Propiedades Estimadores

Información de Fisher y Cota de Cramer Rao

Problemas Propuestos

Sea $\{X_i\}_{i=1}^n$ una muestra i.i.d. de una variable X que sigue una distribución $F(\theta)$. Se define la función de verosimilitud como:

$$f_\theta(\vec{x}) = \begin{cases} \prod_{i=1}^n \mathbb{P}(X_i = x_i/\theta) & \text{Caso Discreto} \\ \prod_{i=1}^n f(x_i/\theta) & \text{Caso Continuo} \end{cases}$$

El estimador de máxima verosimilitud de θ, $\hat{\theta}_{\mathrm{MV}} = \hat{\theta}_{\mathrm{MV}}(\vec{x})$ es la solución del problema:

$$\max_{\delta(\vec{x})} f_\delta(\vec{x})$$

Dado que suele ser más fácil es frecuente resolver el problema equivalente:

$$\max_{\delta(\vec{x})} \log(f_\delta(\vec{x}))$$

Método Bayesiano

Hasta este punto no hemos hecho ningun supuesto sobre el comportamiento de los parámetros a estimar. El método Bayesiano de estimación se utiliza cuando tenemos conocimiento *a priori* del parámetro, reflejado en una distribución de probabilidad.

Formalicemos, sea $\{X_i\}_{i=1}^n$ una muestra i.i.d. de una variable X que sigue una distribución $F(\theta)$. Supongamos que sabemos que θ sigue una distribución $\pi(\theta)$. Queremos incorporar la información contenida en la muestra sobre θ y para esto consideramos una nueva distribución para θ dada por:

$$\xi(\theta/(x_1,\ldots,x_n)) = \frac{f_n(x_1,\ldots,x_n/\theta)\pi(\theta)}{\int_\Theta f_n(x_1,\ldots,x_n/\theta)\pi(\theta)d\theta} \tag{6}$$

Método Bayesiano

Clase Auxilar 5

MA34B-01

Métodos de Estimación

Método de los Momentos

Método de Máxima Verosimilitud

Método Bayesiano

Propiedades Estimadores

Información de Fisher y Cota de Cramer Rao

Problemas Propuestos

Una vez que ya hemos calculado $\xi(\theta/(x_1,\ldots,x_n))$ encontramos el estimador bayesiano para θ minimizando alguna función de pérdida. Es decir tenemos que:

$$\hat{\theta}_{\text{Bayes}} = \underset{\delta(x_1,\ldots,x_n)}{\operatorname{argmin}} \, \mathbb{E}(L(\theta,\delta(x_1,\ldots,x_n))) \tag{7}$$

En concreto tenemos lo siguiente:

Cuadro: Función de Pérdida y Estimador Bayesiano

Función de Pérdida	Estimador
$L(\theta,\delta) = (\theta-\delta)^2$	$\hat{\theta}_{\text{Bayes}} = \mathbb{E}(\theta/\vec{x})$
$L(\theta,\delta) = \|\theta-\delta\|$	$\hat{\theta}_{\text{Bayes}} = a \ / \ \int_a^\infty \xi(\theta/\vec{x})d\theta = \frac{1}{2}$

Propiedades Estimadores

Con todo lo anterior dada una muestra de una v.a. somos capaces de construir distintos estimadores. Es natural preguntarse si algunos estimadores son mejores que otros. Esto nos lleva a la necesidad de tener criterios que nos permitan comparar estimadores.

Definición:

Se define el sesgo de un estimador como $\text{Sesgo}(\hat{\theta}) = \mathbb{E}(\hat{\theta}) - \theta$. Diremos que un estimador es insesgado si su sesgo es igual a cero, diremos que es asintóticamente insesgado si su sesgo tiende a cero con el tamaño de la muestra.

Propiedades Estimadores

Definición:

Se define el error cuadrático medio de un estimador como $\text{ECM}(\hat{\theta}) = \mathbb{E}(\hat{\theta} - \theta)^2$. Diremos que un estimador converge en media cuadrática al parámetro que está estimando si su ECM tiende a cero cuando el tamaño de la muestra se va a infinito. Se tiene la siguiente identidad fundamental:

$$\text{ECM}(\hat{\theta}) = \mathbb{V}ar(\hat{\theta}) + (\text{Sesgo}(\hat{\theta}))^2$$

Definición:

Sea n el tamaño de la muestra y $\hat{\theta}_n$ el estimador de θ asociado a dicha muestra. Se dice que $\hat{\theta}_n$ converge en probabilidad a θ si:

$$(\forall \epsilon > 0) \lim_{n \to \infty} \mathbb{P}(|\hat{\theta}_n - \theta| \leq \epsilon) = 1$$

Propiedades Estimadores

Definición:

Diremos que un estimador $\hat{\theta}_n$ es un estimador consistente para θ si $\hat{\theta}_n$ converge en probabilidad hacia θ.
Como la convergencia en media cuadrática implica la convergencia en probabilidad, tenemos que si un estimador converge en media cuadrática entonces es consistente.

Información de Fisher

Definición:

Se llama cantidad de información de Fisher dada por X sobre el parámetro θ a la cantidad:

$$I(\theta) = \mathbb{E}\left[\left(\frac{\partial \ln(f_X)}{\partial \theta}\right)^2\right]$$

Si el dominio de X no depende de θ, entonces:

$$I(\theta) = -\mathbb{E}\left[\left(\frac{\partial^2 \ln(f_X)}{\partial \theta^2}\right)\right]$$

Si esta cantidad existe.

Información de Fisher de una Muestra

Definición:

La cantidad de información de Fisher dada por una muestra aleatoria $\{x_i\}_i$ sobre el parámetro θ es:

$$I_n(\theta) = \mathbb{E}\left[\left(\frac{\partial \ln(L_\theta)}{\partial \theta}\right)^2\right]$$

Si el dominio de X no depende de θ, entonces:

$$I_n(\theta) = -\mathbb{E}\left[\left(\frac{\partial^2 \ln(L_\theta)}{\partial \theta^2}\right)\right]$$

Si esta cantidad existe.

Información de Fisher de una Muestra

Teorema

Si la muestra aleatoria $\{x_i\}_i$ es independiente, entonces:

$$I_n(\theta) = nI(\theta)$$

Teorema

Si el dominio de X no depende del parámetro θ, para todo estimador T insesgado de θ se tiene:

$$\mathbb{V}ar(T) \geq \frac{1}{I_n(\theta)}$$

Además si T es un estimador insesgado de $h(\theta)$, entoces $\mathbb{V}ar(T) \geq \frac{h'(\theta)^2}{I_n(\theta)}$

Propiedades Estimadores

Definición:

Diremos que un estimador $\hat{\theta}_n$ es un estimador consistente para θ si $\hat{\theta}_n$ converge en probabilidad hacia θ.
Como la convergencia en media cuadrática implica la convergencia en probabilidad, tenemos que si un estimador converge en media cuadrática entonces es consistente.

Problemas

Problema

Se considera una variable aleatoria X con distribución de *Poisson* de parámetro λ, es decir,:

$$\mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

En particular $\mathbb{E}(X) = \mathbb{V}ar(X) = \lambda$. Se considera que a priori, λ sigue una distribución $Gamma(2, 5)$.

1. Encuentre el estimador de Bayes de λ para una función de pérdida cuadrática.
2. Calcule su esperanza.¿El estimador es insesgado?

Recuerde que si X sigue una $Gamma(r, s)$ entonces $f_X(t) = \frac{t^{r-1}e^{\frac{-t}{s}}}{s^r\Gamma(r)}$ para $t > 0$. Además $\mathbb{E}(X) = rs$ y $\mathbb{V}ar(X) = rs^2$

Problemas

Problema

Sea una muestra aleatoria $(X_1, \ldots, X_n)$ de una variable aleatoria X con función de densidad dada por:

$$f_X(t/\alpha) = \begin{cases} \frac{\alpha 2^\alpha}{t^{\alpha+1}} & t \geq 2 \\ 0 & t < 2 \end{cases}$$

1. Calcule el estimador de Máxima Verosimilitud para α.
2. Calcule el estimador de los Momentos para α.

Problemas

Problema

Sea una m.a.s. $\{X_i\}_{i=1}^n$ obtenida de una distribución con densidad:

$$f_X(x/\theta) = \begin{cases} \frac{1}{2-\theta} & t \in [\theta, 2] \\ 0 & t \notin [\theta, 2] \end{cases}$$

Donde θ es desconocido y pertenece a $[0, 2]$

1. Encuentre el estimador de Máxima Verosimilitud de θ, calcule su esperanza y varianza y estudie sesgo y consistencia.

Problemas

Problema

Sea X una variable aleatoria con función de densidad dada por:

$$f_X(x/\alpha) = \begin{cases} \frac{2}{\alpha} x e^{\frac{-x^2}{\alpha}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Con $\alpha > 0$.

1. Encuentre el estimador de máxima verosimilitud para α. Calcule su esperanza y varianza.
2. Encuentre la cota de Cramer-Rao para la varianza de los estimadores insesgados de α.¿Qué puede decir sobre el estimador de máxima verosimilitud?