Una primera aproximación a la Estadística Paramétrica

Proba-2015

Presentación del problema

- Ingredientes: datos generados por un mecanismo aleatorio.
- Objetivo: inferir cuál es el mecanismo que genera los datos.
- Mecanismo aleatorio: Función de distribución.
  - Caso discreto: función de probabilidad puntual
  - Caso continuo: función de densidad
- Modus Operandi: hacer cuenta con los datos para *estimar* el parámetro que determina la distribución que genera los datos.

Estadística paramétrica

Asumimos que la función de distribución que genera los datos pertenece a una familia

$$\mathcal{M} = \{F(\cdot, \theta)\,, \theta \in \Theta\}\,, \text{(modelo)}$$

siendo $\Theta \subset \mathbb{R}^k$, para algún k.

- Caso discreto: $\mathcal{M} = \{p(\cdot, \theta)\,, \theta \in \Theta\}$
- Ejemplo: familia Poisson; $\mathcal{M} = \{\mathcal{P}(\theta) : \theta \in \Theta = \mathbb{R}_{>0}\}$
- Caso continuo: $\mathcal{M} = \{f(\cdot, \theta)\,, \theta \in \Theta\}$
- Ejemplo: familia normal;
  $\mathcal{M} = \{\mathcal{N}(\mu, \sigma^2) : \theta = (\mu, \sigma^2) \in \Theta = \mathbb{R} \times \mathbb{R}_{>0}\}$

Ejercicio

Para cada distribución, establezca los parámetros que la determinan y el espacio al cual estos pertenecen.

1. Bernoulli
2. Poisson
3. Exponencial
4. Normal
5. Gama
6. Uniforme

Muestra

1. $X_1, \cdots, X_n$ se dice una muestra si son variables aleatorias i.i.d. (independientes identicamente distribuidas)
2. En el contexto estadístico, los datos observados constituyen una realización de la muestra y los denotaremos por $\mathbf{x} = x_1, \ldots, x_n$.

Inferencia Estadística

- Hacer una cuenta con los datos.
- Propiedad razonable: A medida que el número de observaciones aumente, esperamos que la cuenta converja al valor que queremos estimar.
- Ejemplo: modelo $\mathcal{M} = \{F(\cdot, \theta)\ , \theta \in \Theta\}$
- $x_1, \ldots, x_n$ datos generados según $F(\cdot, \theta^*)$.
- Objetivo: Averiguar θ^* a partir de los datos, cualquiera sea θ^*.

Inferencia Estadística Paramétrica

- $x_1, \ldots, x_n$ datos generados según $F(\cdot, \theta^*)$.
- Objetivo: Averiguar θ^* a partir de los datos.
- Estimar:hacer una cuenta con los datos $h_n(x_1, x_2, \ldots, x_n)$
- como $\theta \in \Theta$, pedimos que $\widehat{\theta} = h_n(x_1, x_2, \ldots, x_n) \in \Theta$.
- queremos que $\widehat{\theta}_n \to \theta^*$, cualesquiera sean los valores observados en la realización de la muestra, cualquiera sea $\theta^* \in \Theta$.

Convergencia

Pedir que el procedimiento converja cualesquiera sean las observaciones, equivale a evaluar la función h_n en la muestra $X_1, \ldots, X_n$ y pedir que la nueva variables aleatoria $h_n(X_1, \ldots, X_n)$ converja en probabilidad.

Estadístico

- Toda función $h_n(X_1 \ldots, X_n)$ que depende exclusivamente de la muestra se llama estadístico.
- Un estadístico es una nueva variable aleatoria.
- Aquellos estadísticos que sean utilizados para estimar se llaman estimadores.
- Para estimar θ utilizamos $\widetilde{\theta}_n$ o $\widehat{\theta}_n$.
- Es decir, ponemos $\widehat{\theta}_n = h_n(X_1, \ldots, X_n)$ para denotar estadísticos con los que se desea estimar a θ.

Consistencia

Diremos que $\widehat{\theta}_n = h_n(X_1, \ldots, X_n)$ es un estimador consistente para θ si,

$h_n(X_1, \ldots, X_n) \underset{n \to \infty}{\to} \theta$, en probabilidad, cuando $X_i \sim F(\cdot, \theta)$, cualquiera s

En síntesis:

- Modelo $\mathcal{M} = \{F(\cdot, \theta) : \theta \in \Theta\}$
- Estimador: $\widehat{\theta} = \widehat{\theta}_n = h_n(X_1, \cdots, X_n)$
- Consistencia: $\widehat{\theta}_n = h_n(X_1, \cdots, X_n) \to \theta$ en probabilidad, cuando $X_i \sim F(\cdot, \theta)$, para todo $\theta \in \Theta$.

Método de los momentos: ejemplos

$$\overline{X}_n \xrightarrow{\mathcal{P}} \mu \text{ , cuando } X_i \sim \mathcal{N}(\mu, 1)\text{, cualquiera sea } \mu \in \Theta.$$

$$\overline{X}_n \xrightarrow{\mathcal{P}} \frac{1}{\lambda} \text{ , cuando } X_i \sim \mathcal{E}(\lambda)\text{, cualquiera sea } \lambda \in \Theta.$$

$$\overline{X}_n \xrightarrow{\mathcal{P}} \frac{\theta}{2} \text{ , cuando } X_i \sim \mathcal{U}(0, \theta)\text{, cualquiera sea } \theta \in \Theta.$$

Método de los momentos

- Sea $G(\theta) = E[X]$ cuando $X \sim F(\cdot, \theta)$.
- Ley de los grande números garantiza que

$$\overline{X}_n \xrightarrow{\mathcal{P}} G(\theta)\ \text{, cuando } X_i \sim F(\cdot, \theta), \text{ cualquiera sea } \theta \in \Theta.$$

- Si la función G es inversible con G^{-1} continua, tenemos que

$$G^{-1}(\overline{X}_n) \xrightarrow{\mathcal{P}} \theta\ \text{, cuando } X_i \sim F(\cdot, \theta), \text{ cualquiera sea } \theta \in \Theta,$$

- $\widetilde{\theta} := G^{-1}(\overline{X}_n)$ es un estimador consistente para θ.

Método de los momentos

- Sea $G(\theta) = E[X]$ cuando $X \sim F(\cdot, \theta)$.
- $\widetilde{\theta} := G^{-1}(\overline{X}_n)$ es un estimador consistente para θ.
- $\widetilde{\theta}$ es solución de
$$\overline{X}_n = G(\widetilde{\theta})$$

Verosimilitud: Ejemplo

- Monedas, cara$=1$, ceca$=0$.
- Moneda bolsillo derecho: equilibrada
- Moneda bolsillo izquierdo: probabilidad de cara es 0.8.
- Objetivo: identificar la moneda a partir de una muestra. en $n=100$ lanzamientos se observa la muestra

$$\mathbf{x} = \underbrace{1,\ldots,1}_{\mathbf{12veces}}\underbrace{0,\ldots,0}_{\mathbf{5veces}}\underbrace{1,\ldots,1}_{\mathbf{23veces}}\underbrace{0,\ldots,0}_{\mathbf{8veces}}\underbrace{1,\ldots,1}_{\mathbf{15veces}}\underbrace{0,\ldots,0}_{\mathbf{3veces}}$$
$$\underbrace{1,\ldots,1}_{11 veces}\underbrace{0,\ldots,0}_{4 veces}\underbrace{1,\ldots,1}_{13 veces}\underbrace{0,\ldots,0}_{6 veces}$$

- ¿Cuál de las dos monedas diría que está utilizando?

Propuesta de estimación por maxima verosimilitud

- Función de verosimilitud

- $L(0.8)$ = probabilidad de observar la muestra $\mathbf{x}$ con moneda de $p = 0.8$.
- $L(0.5)$ = probabilidad de observar la muestra $\mathbf{x}$ con moneda de $p = 0.5$ (equilibrada).

$$L(0.8) = \underbrace{0.8\ldots0.8}_{12veces}\underbrace{0.2\ldots0.2}_{5veces}\underbrace{0.8\ldots0.8}_{23veces}\underbrace{0.2\ldots0.2}_{8veces}\underbrace{0.8\ldots0.8}_{15veces}\underbrace{0.22\ldots}_{3veces}$$

$$\underbrace{0.8\ldots0.8}_{11veces}\underbrace{0.2\ldots0.2}_{4veces}\underbrace{0.8\ldots0.8}_{13veces}\underbrace{0.2\ldots0.2}_{6veces} =$$

$$= (0.8)^{74}(0.2)^{26} = 4.523128e - 26$$

siendo 74 el número de caras observadas en las $n = 100$ repeticiones. Análogamente, tenemos que

$$L(1/2) = (1/2)^{74}(1/2)^{26} = 7.888609e - 31 \,,$$

Y ahora?

- $L(0.8) = 4.523128e - 26$
- $L(1/2) \; = \; (1/2)^{74}(1/2)^{26} = 7.888609e - 31$
- ¿cuál moneda diría usted que está utilizando?

Propuesta de máxima verosimilitud verosimilitud

La propuesta de máxima verosimilitud consiste en pensar que la moneda que estamos utilizando es aquella para los cuales los valores observados resultan mas probables. Es decir, elegir la moneda que maximiza la probabilidad de los valores observados. Siendo que

$$L(0.8) \ > \ L(1/2)$$

concluimos que se está utilizando la moneda no equilibrada.

Generalizando

- $X_i \sim \mathcal{B}(1, \theta)$, $\theta \in [0, 1]$.
- 74 caras en $n = 100$ repeticiones.
- θ?
- vamos a elegir aquel valor para el cuál los valores observados tienen mas probabilidad.

Función de verosimilitud:

- $L(\theta, \mathbf{x})$: Mide cuál es la probabilidad de observar nuestra realización cuando la probabilidad de cara es θ.

- en nuestro caso, con 74 caras en $n = 100$ repeticiones tenemos que
$$L(\theta) \;=\; \theta^{74}(1-\theta)^{26}$$

- El estimador de máxima verosimilitud está dado por $\hat{\theta}$ que maximiza $L(\theta)$.

- $\leftrightarrow$ maximizar $l(\theta) = \ln(L(\theta))$

- $l(\theta) = 74\ln(\theta) + 26\ln(1-\theta)$, se maximiza en $\hat{\theta} = 74/100$.

- Tenemos así que la probabilidad estimada de obtener cara con la moneda que está utilizando es $\hat{\theta} = 0.74$.

Máxima verosimilitud: caso discreto

- Modelo: $\mathcal{M} = \{p(\cdot,\theta)\ ,\theta \in \Theta\}$.

- $\mathbf{x}= x_1, \cdots, x_n$ realización de $X_1, \cdots, X_n$ i.i.d.

- Función de verosimilitud asociada a $\mathbf{x}= x_1, \cdots, x_n$:

$$L(\cdot\ ;\ \mathbf{x}\ )\ : \Theta \to \mathcal{R}$$

$$\begin{aligned} L(\theta\ ;\ \mathbf{x}\ ) &= P(X_1, \cdots, X_n = x_1, \ldots, x_n) \\ &= p_{X_1,\cdots,X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} p(x_i, \theta)\ , X_i \sim p(\cdot, \theta)\ . \end{aligned}$$

- Propuesta de máxima verosimilitud:

$$h_n(x_1, \ldots, x_n) = \underset{\theta \in \Theta}{\operatorname{argmax}}\ L(\theta\ ;\ \mathbf{x}\ )\ .$$

Observación Importante

1. Con $x_1, \ldots, x_n$ obtenemos $h_n(x_1, \cdots, x_n)$
2. Haciendo $h_n(X_1, \cdots, X_n)$ tenemos el estimador $\widehat{\theta}_n$.
3. Ejemplo caso binomial: $h_n(x_1, \cdots, x_n) = \cdots$, $\widehat{\theta}_n = \cdots$.

Estimador de Máxima verosimilitud: caso discreto

Bajo el modelo $\mathcal{M} = \{p(\cdot, \theta)\ , \theta \in \Theta\}$, el estimador de máxima verosimilitud para θ a la variable aleatoria $\widehat{\theta}_n = h_n(X_1, \ldots, X_n)$.

Ejercicio

Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución

- $Bi\,(1, \theta)\,,\ 0 \leq \theta \leq 1.$
- $\mathcal{P}\,(\theta)\,,\ \theta > 0.$
- $\mathcal{G}\,(\theta)\,,\ 0 \leq \theta \leq 1.$

En cada uno de estos casos, encontrar el estimador de máxima verosimilitud de θ.

Máxima verosimilitud: Caso continuo

- Modelo: $\mathcal{M} = \{f(\cdot, \theta), \theta \in \Theta\}$, con $f(\cdot, \theta)$ función de densidad.
- $\mathbf{x}= x_1, \cdots, x_n$ realización correspondiente a $X_1, \ldots, X_n$
- Función de verosimilitud: $L(\cdot\,;\,\mathbf{x}) : \Theta \to \mathbb{R}$

$$L(\theta\,;\,\mathbf{x}) = f_{X_1,\cdots,X_n}(x_1, \ldots, x_n)$$
$$= \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} f(x_i, \theta), X_i \sim f(\cdot, \theta)$$

- Propuesta de máxima verosimilitud:

$$h_n(x_1, \ldots, x_n) = \underset{\theta \in \Theta}{\operatorname{argmax}}\; L(\theta\,;\,\mathbf{x}).$$

Ejercicio

Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución

- $\mathcal{E}(\theta)$, $\theta > 0$.
- $\mathcal{U}[0, \theta]$, $0 < \theta$.
- $\mathcal{N}(\theta, 1)$, $\theta \in \mathbb{R}$.
- $\mathcal{N}(\mu, \sigma^2)$

En cada uno de estos casos, encontrar el estimador de máxima verosimilitud de los parámetros que identifican al modelo.

Ejercicio

Sea $X_1, \ldots, X_n$ una m.a. de una distribución exponencial desplazada, cuya densidad es

$$f\left(x\right) = e^{-(x-\theta)} I_{[\theta,\infty)}\left(x\right).$$

Encontrar el estimador de máxima verosimilitud de θ.