The Maximum LikeliHood estimation.

Sea $A(x_1, \ldots x_n)$ un vector aleatorio cuya distribución depende de un parámetro θ

Definimos la función de verosimilitud de A es:

$$L(\theta) = f_{x_1, \ldots x_n}(x_1, \ldots x_n; \theta)$$

Que es la función de densidad evaluada en el Vector A es una función del parámetro. No necesariamente son identicamente distribuidas.

Cuando son independientes

$$\mathcal{L}(\theta) = f(x_1; \theta) f(x_2; \theta) \ldots f(x_n; \theta)$$
$$= \prod_{i=1}^{n} f(x_i, \theta)$$

El Método consiste en maximizar la función $L(\theta)$.
θ que maximiza $L(\theta)$ se conoce como estimación máximo verosímil

"Dicho Valor ~~fen~~ maximiza la probabilidad que el modelo describa a A

Sea $A(x_1, \dots x_n)$ $A \sim exp(\alpha)$

$$\Rightarrow \quad \mathcal{L}(\theta) = \prod_{i=1}^{n} \frac{1}{\alpha} e^{-x_i/\alpha} = \frac{1}{\alpha^n} e^{-\sum x_i/\alpha}$$

Tomamos Logaritmo natural. (Dado que $\ln(x)$ monótono)

$$\ln \mathcal{L}(\theta) = \ln\left(\left(\frac{1}{\alpha}\right)^n e^{-\sum x_i/\alpha}\right)$$

$$= n \ln\left(\frac{1}{\alpha}\right) - \sum \frac{x_i}{\alpha}$$

Para encontrar el máximo encontremos el máximo

$$\frac{d \ln \mathcal{L}(\alpha)}{d\alpha} = \frac{n}{1/\alpha}\cancel{\left(-\frac{1}{\alpha^2}\right)} - \cancel{\left(-\frac{1}{\alpha^2}\right)} \sum x_i = 0$$

$$\alpha = \frac{\sum x_i}{n} \quad \text{promedio muestral } \bar{x}$$

Se acostumbra a notar

$$\hat{\alpha} = \bar{x}$$

Ejemplo 2 Normal Sea $A(x_1, \dots x_n)$ con $A \sim N(\mu, \sigma^2)$

La función de Verosimilitud es:

$$\mathcal{L}(\mu, \sigma^2) = f(x_1; \mu, \sigma^2) \dots f(x_n; \mu, \sigma^2)$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_1-\mu)^2/2\sigma^2} \cdots \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_n-\mu)^2}{2\sigma^2}}.$$

$$= (2\pi\sigma^2)^{-n/2} \cdot \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

Tomando Logaritmos.

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

Ahora puntos críticos:

$$\frac{\partial}{\partial\mu}\ln L(\mu,\sigma^2) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) \qquad \frac{\partial}{\partial\sigma^2}\ln L(\mu,\sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{(2\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\Rightarrow \sum_{i=1}^{n}(x_i-\mu) = 0 \quad y \quad -n\sigma^2 + \sum_{i=1}^{n}(x_i-\mu)^2 = 0 \quad (2)$$

Resolver sistema de ecuaciones

De (1) $\hat{\mu} = \bar{x}$ de (2) $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$

Ejercicio Calcular la Matriz hessiana y Mostrar que es definida Negativa:

$$H(\mu,\sigma^2) = \begin{pmatrix} \frac{\partial^2}{\partial\mu^2}\ln(L(\mu,\sigma^2)) & \frac{\partial^2}{\partial\mu\partial\sigma^2}\ln L(\mu,\sigma^2) \\ \frac{\partial^2}{\partial\sigma^2\mu}\ln(L(\mu,\sigma^2)) & \frac{\partial^2}{\partial\sigma^2}\ln L(\mu,\sigma^2) \end{pmatrix}$$

$$\begin{pmatrix} -\frac{n}{\sigma^2} & -(\sigma^2)^{-2}\sum_{i=1}^{n}(x_i-\mu) \\ -(\sigma^2)^{-2}\sum_{i=1}^{n}(x_i-\mu) & \frac{n}{2}(\sigma^2)^{-2}-(\sigma^2)^{-3}\sum_{i=1}^{n}(x_i-\mu)^2 \end{pmatrix}$$

Evaluar en el punto crítico

$$H(\hat{\mu},\hat{\sigma}^2) = \begin{pmatrix} -n(\hat{\sigma}^2)^{-1} & 0 \\ 0 & -\frac{n}{2}(\hat{\sigma}^2)^{-2} \end{pmatrix}$$

Como $H_{11} < 0$ y el determinante $|H(\hat{\mu},\hat{\sigma}^2)| > 0$

Ejemplo 3 Poisson $\frac{e^{-\lambda}\lambda^x}{x!}$ $\lambda = \theta$

$$L(\theta) = \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!} = \frac{e^{-n\theta}\theta^{\sum x_i}}{\prod_{i=1}^{n} x_i!}$$

Tomando logaritmos

$$\mathcal{L}(\theta) = \ln L(\theta) = -n\theta + \sum x_i \ln\theta - \ln\prod_{i=1}^{n} x_i!$$

Derivando

$$\frac{\partial \mathcal{L}n(\theta)}{\partial\theta} = -n + \frac{\sum x_i}{\theta} = 0$$

$$\Rightarrow \hat{\theta} = \frac{\sum x_i}{n} = \overline{X}$$ Media muestral.

Varianza de estimadores

$V(\hat\theta) = E(\hat\theta^2) - E(\hat\theta)^2$ No calculable en general

Para la exponencial

$$\int \cdots \int \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2 \frac{1}{\theta} e^{-x_1/\theta} \cdots \frac{1}{\theta} e^{-x_n/\theta}\, dx_1 \ldots dx_n$$

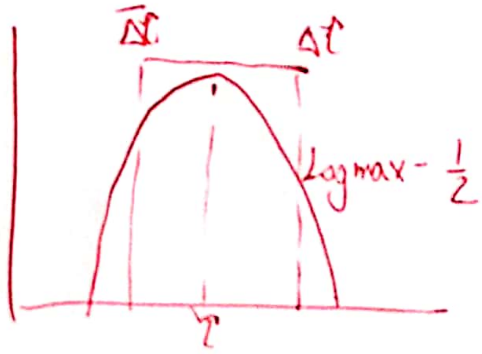
$$- \left(\int \cdots \int \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \frac{1}{\theta} e^{-x_1/\theta} \cdots \frac{1}{\theta} e^{-x_n/\theta}\, dx_1 \ldots dx_n\right)^2$$

$$= \frac{\theta^2}{n}$$

Expandiendo alrededor del Estimador

$$\log L(\theta) = \log L(\hat\theta) + \frac{\partial \log L}{\partial \theta}\Big|_{\theta=\hat\theta}(\theta-\hat\theta) + \frac{1}{2!}\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)_{\theta=\hat\theta}(\theta-\hat\theta)^2$$

$$\log L(\theta) = \log L_{max} - \frac{(\theta-\hat\theta)^2}{2\hat\sigma_\theta^2}$$



Si Evaluamos en $\theta = \hat\theta \pm \hat\sigma_\theta$

$$\log L(\hat\theta \pm \hat\sigma_\theta) = \log max - \frac{(\hat\theta \pm \hat\sigma_\theta - \hat\theta)^2}{2\hat\sigma_\theta^2}$$

$$= \log max - \frac{1}{2} \Rightarrow \hat\sigma^2 = -\frac{1}{\frac{\partial^2 \log L}{\partial \theta^2}\Big|_{\hat\theta}}$$