1 Função de máxima verossimilhança (maximum likelihood function)

.

O método de estimativa por máxima verossimilhança tem o objetivo de encontrar estimativas para os parâmetros de uma determinada distribuições tomando como entrada uma amostra dessa população.

Parece razoável que uma boa estimativa do parâmetro desconhecido θ de uma distribuição seja o valor que maximize a probabilidade, isto é, a verossimilhança, de obter os dados que observamos. Resumidamente, esta é a idéia por trás do método de estimativa de máxima verossimilhança.

Mas como implementaremos o método na prática? Suponha que tenhamos uma amostra aleatória $X_1, X_2, \ldots, X_m$ para a qual a função de densidade de probabilidade (ou massa) de cada X_i é $f(x^{(i)};\theta)$. Então, a função de densidade de probabilidade conjunta de $X_1, X_2, \ldots, X_m$, que chamaremos $L(\theta)$ é:

$$L(\theta) = P(X_1 = x^{(1)}, X_2 = x^{(2)}, \ldots, X_m = x^{(m)}) = f(x^{(1)};\theta) \cdot f(x^{(2)};\theta) \cdots f(x^{(m)};\theta) = \prod_{i=1}^{m} f(x^{(i)};\theta)$$

2 Estimadores para a média e para a variância de uma função de densidade gaussiana monovariada

Sejam $x^{(1)}, x^{(2)}, \ldots, x^{(m)}$ uma amostra aleatória de uma distribuição normal univariada com média μ e desvio padrão σ. Vamos provar as expressões para os estimadores da média e da variância da população.

Vamos escrever a função de densidade como uma função de μ e de σ^2:

$$f(x^{(i)};\theta_1,\theta_2) = \frac{1}{\sqrt{\theta_2}\sqrt{2\pi}} \exp\left[-\frac{(x^{(i)} - \theta_1)^2}{2\theta_2}\right]$$

Agora, vamos definir a função de verossimilhança (likelihood function):

$$L(\theta_1,\theta_2) = \prod_{i=1}^{m} f(x^{(i)};\theta_1,\theta_2) = \theta_2^{-m/2}(2\pi)^{-m/2} \exp\left[-\frac{1}{2\theta_2}\sum_{i=1}^{m}(x^{(i)} - \theta_1)^2\right]$$

Definimos também o log da função de verossimilhança:

$$\log L(\theta_1,\theta_2) = -\frac{m}{2}\log\theta_2 - \frac{m}{2}\log(2\pi) - \frac{\sum(x^{(i)}-\theta_1)^2}{2\theta_2}$$

Tomando a derivada parcial do log da função de verossimilhança com relação a θ_1 e igualando a 0:

$$\frac{\partial \log L(\theta_1,\theta_2)}{\partial\theta_1} = \frac{-2\sum(x^{(i)}-\theta_1)(-1)}{2\theta_2} = 0$$

Agora, multiplicando por θ_2 e distribuindo a soma, temos:

$$\sum x^{(i)} - m\theta_1 = 0$$

Agora, resolvendo para θ_1:

$$\hat{\theta}_1 = \hat{\mu} = \frac{\sum x^{(i)}}{m}$$

Agora, para θ_2. Tomando a derivada parcial do log da função de verossimilhança com relação a θ_2 e igualando a 0:

$$\frac{\partial \log L(\theta_1,\theta_2)}{\partial\theta_2} = \frac{-m}{2\theta_2} + \frac{(\sum(x^{(i)}-\theta_1))^2}{2\theta_2^2} = 0$$

Assim, obtemos:

$$-m\theta_2 + \sum(x^{(i)}-\theta_1)^2 = 0$$

Resolvendo para θ_2:

$$\hat{\theta}_2 = \hat{\sigma}^2 = \frac{\sum(x^{(i)}-\bar{x})^2}{m}$$

Sendo assim, como queríamos provar:

$$\hat{\mu} = \frac{\sum x^{(i)}}{m}$$

$$\hat{\sigma}^2 = \frac{\sum(x^{(i)}-\hat{\mu})^2}{m}$$