Apêndice 1 - Cálculo dos estimadores de máxima verossimilhança

Para uma sequência de V.A.s $X = [X_1, X_2, X_3]$ que representam amostras i.i.d de uma distribuição $f(x|\theta)$, temos que a função de verossimilhança pode ser escrita como

$$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

O estimador MLE é encontrado ao maximizar $L(\theta)$:

$$\hat{\theta}_{MLE} = \underset{\theta}{\arg\max}\ L(\theta)$$

1 - Normal com μ e σ desconhecidos

Temos a pdf da normal como

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

Assim, $L(\theta)$ é

$$L(\theta) = \prod_{i=1}^{n}\left(\frac{1}{\sqrt{2\pi}\,\sigma} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}}\right) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{\frac{-\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}}$$

Usamos o logaritmo de $L(\theta)$ para calcular o valor máximo e obtemos

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

Podemos agora calcular as derivadas parciais $\qquad x_i^2 - 2x_i\mu + \mu^2$

$$\frac{\partial \ell}{\partial \hat{\mu}} = -\sum_{i=1}^{n} \frac{(2\hat{\mu} - 2x_i)}{\cancel{2\sigma^2}} = 0$$

$$\therefore \sum_{i=1}^{n} \cancel{2}\hat{\mu} = \sum_{i=1}^{n} \cancel{2}x_i \quad \therefore n\hat{\mu} = \sum_{i=1}^{n} x_i \quad \therefore \hat{\mu} = \frac{\sum x_i}{n}$$

Para σ, temos:

$$\ell(\theta) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{\boxed{\sum_{i=1}^{n}(x_i-\mu)^2}}{2\sigma^2} \quad S$$

$$\frac{\partial \ell}{\partial \sigma} = -\frac{n}{\cancel{2}} \frac{\cancel{4}\cancel{\pi}\cancel{\sigma}}{\cancel{2}\cancel{\pi}\sigma^{\cancel{2}}} - \frac{S}{2}\left(\frac{-2}{\sigma^3}\right) = 0$$

Assim, temos:

$$\frac{n}{\sigma} = \frac{S}{\sigma^3} \quad \therefore \sigma^2 = \frac{S}{n} = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}$$

(Onde trocamos μ por $\hat{\mu}_{MLE}$

2 - Distribuição gamma

Temos a pdf da distribuição gama como:

$$f(x|\alpha,\lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$$

Calculamos a verossimilhança como:

$$L(\theta) = \prod_{i=1}^{n} \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x_i^{\alpha-1} e^{-\lambda x_i} = \left(\frac{\lambda^{\alpha}}{\Gamma(\alpha)}\right)^{n} e^{-\lambda \sum_n x_i} \prod_{i=1}^{n} x_i^{\alpha-1}$$

O logaritmo da verossimilhança:

$$S$$

$$\ell(\theta) = n\alpha \ln(\lambda) - n\ln(\Gamma(\alpha)) - \lambda \sum_{i=1}^{n} x_i + (\alpha-1)\sum_{i=1}^{n} \ln(x_i)$$

Calculamos as derivadas e igualamos a zero para obter os parâmetros:

$$\frac{\partial \ell}{\partial \lambda} = \frac{n\alpha}{\lambda} - S = 0$$

$$\frac{n\alpha}{\lambda} = S \quad \therefore \quad \hat{\lambda}_{MLE} = \frac{\hat{\alpha}}{\overline{X}}$$

Agora para α

$$\frac{\partial \ell}{\partial \alpha} = n\ln(\hat{\lambda}) - \frac{n}{\Gamma(\alpha)} \Gamma'(\alpha) + \sum_{i=1}^{n} \ln(x_i) = 0$$

Vamos chamar $\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}$ de $\phi(\alpha)$, que é conhecida como a função digamma.

Assim, substituindo λ por $\frac{\alpha}{\overline{X}}$, temos:

$$n\ln\left(\frac{\hat{\alpha}}{\overline{X}}\right) - n\phi(\hat{\alpha}) + \sum_{i=1}^{n} \ln(x_i) = 0$$

$$n \ln(\hat{\alpha}) - n \ln(\bar{X}) - n\,\phi(\alpha) + \sum_{i=1}^{n} \ln(x_i) = 0$$

fatorando n, temos

$$\ln(\hat{\alpha}) - \ln(\bar{X}) - \phi(\alpha) + \overline{\ln(x_i)} = 0$$

Observação: Temos que não existe forma analítica para encontrar os valores dos parâmetros da distribuição gamma. Sendo assim, o código encontrará os valores dos parâmetros via MLE através de otimização numérica.