STT 1000 - STATISTIQVES

ARTHVR CHARPENTIER

Loi Gaussienne

Loi normale $\mathcal{N}(\theta,\theta)$

Soit x_i un échantillon Gaussien de loi $\mathcal{N}(\theta,\theta)$ (de variance θ) avec $\theta > 0$.

$$\log \mathcal{L}(\theta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\theta^2) - \frac{1}{2\theta^2}\sum_{i=1}^n (x_i-\theta)^2$$

$$\frac{\partial \log \mathcal{L}(\theta)}{\partial\theta} = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^n (x_i-\theta)^2 + \frac{1}{\theta}\sum_{i=1}^n (x_i-\theta)$$

$$\frac{\partial \log \mathcal{L}(\theta)}{\partial\theta} = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^n x_i^2 - \frac{n}{2}$$

Autrement dit, $\widehat{\theta}$ est solution de l'équation quadratique

$$\widehat{\theta}^2 + \widehat{\theta} - \overline{x^2} = 0, \text{ i.e. } \widehat{\theta} = -\frac{1}{2} + \sqrt{\frac{1}{4} + \overline{x^2}}$$

Notons que

$$\frac{\partial^2 \log \mathcal{L}(\theta)}{\partial\theta^2} = \frac{n}{\theta^3}\left(\frac{\theta}{2} - \widehat{\theta} - \widehat{\theta}^2\right) < 0$$

Loi normale $\mathcal{N}(\theta,\theta^2)$

Soit x_i un échantillon Gaussien de loi $\mathcal{N}(\theta,\theta^2)$ (de variance θ^2) avec $\theta > 0$.

$$\log \mathcal{L}(\theta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\theta) - \frac{1}{2\theta^2}\sum_{i=1}^{n}(x_i-\theta)^2$$

$$\frac{\partial \log \mathcal{L}(\theta)}{\partial \theta} = -\frac{n}{\theta} + \frac{n\overline{x^2}}{\theta^3} - \frac{n\overline{x}}{\theta^2}$$

Autrement dit, $\widehat{\theta}$ est solution de l'équation quadratique

$$\widehat{\theta}^2 + \overline{x}\widehat{\theta} - \overline{x^2} = 0, \text{ i.e. } \widehat{\theta} = -\frac{1}{2} + \sqrt{\frac{\overline{x^2}}{4} + \overline{x^2}}$$

qui est encore un maximum.

Loi normale multivariée $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$$f(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^k |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$

$$\log \mathcal{L}(\widehat{\boldsymbol{\mu}}, \boldsymbol{\Sigma}) = -\frac{dn}{2}\log(2\pi) - \frac{n}{2}\log|\boldsymbol{\Sigma}| - \frac{n-1}{2}\text{tr}(\boldsymbol{\Sigma}^{-1}\boldsymbol{S})$$

où

$$\boldsymbol{S} = \frac{1}{n-1}\sum_{i=1}^n (\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})^\top$$

On peut alors montrer que le maximum de la log-vraisemblance est atteint en

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{n}(n-1)\boldsymbol{S} = \frac{1}{n}\sum_{i=1}^n (\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})^\top$$