Estimación de parametros

1. (**Theoretical**) Sea $\mathcal{A}(x_1, x_2, ..., x_n)$, donde $\mathcal{A} \sim \mathcal{N}(\mu, \sigma)$ con parámetros μ y σ. Muestre que los estimadores máximo verosímiles son:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{9.49}$$

La función de verosimilitud $(\mathcal{L}(\theta | x_1, x_2, ..., x_n))$ nos dice que tan bien un modelo se ajusta a los datos. Los estimadores máximo verosímiles son los parámetros del modelo que mejor se ajustan a los datos, es decir, los parámetros en los que la función de verosimilitud se maximiza. $\nabla \mathcal{L} = 0$.

$$\mathcal{L}(\vec{x}, \mu, \sigma) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$= \frac{1}{(\sigma\sqrt{2\pi})^n} \prod_{i=1}^{n} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$= \frac{1}{(\sigma\sqrt{2\pi})^n} e^{-\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$\boxed{\mathcal{L}(\vec{x}, \mu, \sigma) = \frac{1}{(\sigma\sqrt{2\pi})^n} e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}}}$$

$$\nabla \mathcal{L} = \begin{pmatrix} \frac{\partial \mathcal{L}}{\partial \mu} \\ \frac{\partial \mathcal{L}}{\partial \sigma} \end{pmatrix} = \vec{0}$$

$$\frac{\partial}{\partial \mu}\left(\mathcal{L}(\vec{x}, \mu, \sigma)\right) = \frac{1}{(\sigma\sqrt{2\pi})^n} e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}} \left[\frac{1}{2\sigma}\left(\sum_{i=1}^{n} 2(x_i-\mu)(-1)\right)\right] = 0 \longleftrightarrow \sum_{i=1}^{n}(x_i-\mu) = 0$$

$$\sum_{i=1}^{n}(x_i - \hat{\mu}) = x_1 - \hat{\mu} + x_2 - \hat{\mu} + \cdots + x_n - \hat{\mu} = 0$$

$$x_1 + x_2 + \cdots + x_n = n\hat{\mu}$$

$$\hat{\mu} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

$$\boxed{\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{X}}$$

Por las propiedades del logaritmo natural, los puntos críticos de una función $f(x)$ son los mismos que los de $l(x) = \ln(f(x))$

$$l(\vec{x},\hat{\mu},\sigma) = \ln(L(\vec{x},\hat{\mu},\sigma)) = \ln(1) - \left[\ln(\sigma^n) + \ln(\sqrt{2\pi}^n)\right] + \left(-\frac{\sum_{i=1}^{n}(x_i-\hat{\mu})^2}{2\sigma^2}\right)$$

$$\frac{\partial}{\partial\sigma}\left(l(\vec{x},\hat{\mu},\sigma)\right) = \frac{-1}{\sigma^n}(n\sigma^{n-1}) - \left(\frac{-1}{\sigma^3}\right)\sum_{i=1}^{n}(x_i-\hat{\mu})^2 = \frac{-n\sigma^2}{\sigma^3} + \frac{\sum_{i=1}^{n}(x_i-\hat{\mu})^2}{\sigma^3}$$

$$\frac{\partial}{\partial\sigma}\left(l(\vec{x},\hat{\mu},\sigma)\right) = 0 \longleftrightarrow n\sigma^2 = \sum_{i=1}^{n}(x_i-\hat{\mu})^2$$

$$\boxed{\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{X})^2}$$