

$$\hat{\theta}_{ML} = \arg \max_{\theta} \left(\prod_{k=1}^N p(\vec{x}_k; \theta) \right)$$

$$L(\theta) = \ln \left(\prod_{k=1}^N p(\vec{x}_k; \theta) \right)$$

$$\frac{\partial L(\theta)}{\partial \theta} = \sum_{k=1}^N \frac{1}{p(\vec{x}_k; \theta)} \frac{\partial}{\partial \theta} (p(\vec{x}_k; \theta)) = 0$$

(P1) Let $p(x_k, \sigma^2) = \frac{1}{\sqrt{2\pi} \sigma} \exp\left(-\frac{(x_k - \mu)^2}{2\sigma^2}\right)$. Compute $\hat{\sigma}_{ML}^2$

$$p(x_k, \sigma^2) = (2\pi)^{-1/2} (\sigma^2)^{-1/2} \exp\left(-\frac{(x_k - \mu)^2}{2\sigma^2}\right)$$

$$L(\sigma^2) = \ln \left(\prod_{k=1}^N (2\pi)^{-1/2} (\sigma^2)^{-1/2} \exp\left(-\frac{(x_k - \mu)^2}{2\sigma^2}\right) \right)$$

$$= \sum_{k=1}^N \ln \left((2\pi \sigma^2)^{-1/2} \right) - \sum_{k=1}^N \frac{(x_k - \mu)^2}{2\sigma^2}$$

$$= -\frac{N}{2} \ln(2\pi \sigma^2) - \frac{1}{2\sigma^2} \sum_{k=1}^N (x_k - \mu)^2$$

$$\frac{\hat{\sigma}_{ML}^2}{ML} : \frac{dL(\sigma^2)}{d\sigma^2} = 0$$

$$\frac{dL(\sigma^2)}{d\sigma^2} = -\left(\frac{N}{2}\right) \left(\frac{1}{2\pi \sigma^2}\right) (2\pi) + \frac{1}{2} \left(\frac{1}{\sigma^4}\right) \sum_{k=1}^N (x_k - \mu)^2$$

$$\frac{dL(\sigma^2)}{d\sigma^2} = 0$$

$$\left(\frac{N}{2}\right) \left(\frac{1}{\sigma^2}\right) = \left(\frac{1}{2}\right) \left(\frac{1}{\sigma^4}\right) \left(\sum_{k=1}^N (x_k - \mu)^2\right)$$

$$\Rightarrow \boxed{\hat{\sigma}_{ML}^2 = \frac{1}{N} \sum_{k=1}^N (x_k - \mu)^2}$$

Properties

(1) ML estimate is asymptotically unbiased

$$\lim_{N \rightarrow \infty} E[\hat{\theta}_{ML}] = \theta_0$$

(2) ML estimate is asymptotically consistent

$$\lim_{N \rightarrow \infty} \text{prob}(\|\hat{\theta}_{ML} - \theta_0\| \leq \epsilon) = 1$$

Alternatively, a strong condition:

$$\lim_{N \rightarrow \infty} E[\|\hat{\theta}_{ML} - \theta_0\|^2] = 0$$

(3) ML estimate is asymptotically efficient that it achieves Cramer-Rao lower bounds.

(4) The pdf of ML estimate as $N \rightarrow \infty$ approaches the Gaussian distribution with mean θ_0 .