极大似然估计

- Data: $X = (x_1, x_2, \dots, x_N)^T = \begin{pmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_N^T \end{pmatrix}_{N \times P}$

  $x_i \in \mathbb{R}^P$.

  $x_i \overset{iid}{\sim} N(\mu, \Sigma)$.

  $\theta = (\mu, \Sigma)$

点估计
- MLE: $\theta_{MLE} = \arg\max_\theta P(x|\theta)$

  令 $P=1$ $\quad \theta = (\mu, \delta^2)$.

$$\begin{cases} P(x) = \frac{1}{\sqrt{2\pi}\delta} e^{-\frac{(x-\mu)^2}{2\delta^2}} \\ P(x) = \frac{1}{(2\pi)^{P/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu) \Sigma^{-1} (x-\mu)\right) \end{cases}$$

$$\log P(x|\theta) = \log \prod_{i=1}^N P(x_i|\theta) = \sum_{i=1}^N \log P(x_i|\theta).$$

$$= \sum_{i=1}^N \log \frac{1}{\sqrt{2\pi}\delta} \exp\left(-\frac{(x-\mu)^2}{2\delta^2}\right).$$

$$= \sum_{i=1}^N \left[\log \frac{1}{\sqrt{2\pi}} - \frac{(x-\mu)^2}{2\delta^2} + \log \frac{1}{\delta}\right]$$

① $\mu_{MLE} = \arg\max_\mu \log P(x|\theta)$

$$= \arg\max_\mu \Sigma \frac{(x-\mu)^2}{2\delta^2}$$

$$= \arg\max_\mu \Sigma (x-\mu)^2$$

$$\frac{\partial (x-\mu)^2}{\partial \mu} = \Sigma\, 2(x_i - \mu)(-1) = 0$$

$$\Sigma (x_i - \mu) = 0 \Rightarrow \Sigma x_i - \Sigma \mu = 0$$

$$\boxed{\mu_{MLE} = \frac{1}{N} \sum_{i=1}^N x_i}$$ 无偏

$$E[\mu_{MLE}] = \frac{1}{N} \Sigma E[x_i]$$

$$= \frac{1}{N} \Sigma E[x_i] = \frac{1}{N} \Sigma \mu = \mu$$

$$var[\mu_{MLE}] = var\left(\frac{1}{N} \Sigma x_i^2\right] = \frac{1}{N^2} Var(\Sigma x_i^2) = \frac{1}{N^2} \Sigma \delta^2 = \frac{1}{N} \delta^2$$

② $\sigma^2_{MLE} = \arg\max\limits_{\sigma} P(x|\theta)$

$$= \arg\max \sum \left[\log \frac{1}{\sigma} - \frac{(x-\mu)^2}{2\sigma^2} \right]$$

$$= \arg\max \underbrace{\sum \left[-\log \sigma - \frac{(x-\mu)^2}{2\sigma^2} \right]}_{L}$$

$$\frac{\partial L}{\partial \sigma} = \sum \left[-\frac{1}{\sigma} + (x_i-\mu)^2 \sigma^{-3} \right] = 0$$

$$\sum \left[-\sigma^2 + (x_i-\mu)^2 \right] = 0$$

$$N\sigma^2 = \sum (x_i-\mu)^2$$

$$\boxed{\sigma^2_{MLE} = \frac{1}{N}\sum (x_i-\mu)^2}$$ → 有偏估计.

$E[\sigma^2_{MLE}] = \frac{N-1}{N}\sigma^2$ ← 真正的参数.

无偏：$\hat{\sigma} = \frac{1}{N-1}\sum (x_i - \mu_{MLE})^2$

$$\begin{aligned}
E[\sigma^2_{MLE}] &= E\left[\frac{1}{N}\sum x_i^2 - \mu^2_{MLE}\right] \\
&= E\left[\frac{1}{N}\sum x_i^2 - \mu^2 - (\mu^2_{MLE} - \mu^2)\right] \\
&= E\left[\frac{1}{N}\sum x_i^2 - \mu^2\right] - E\left[\mu^2_{MLE} - \mu^2\right] \\
&= E\left[\frac{1}{N}\sum (x_i^2 - \mu^2)\right] - \left[E[\mu^2_{MLE}] - E(\mu^2)\right] \\
&= \frac{1}{N}\sum E(x_i^2 - \mu^2) \quad - \left[E(\mu^2_{MLE}) - \mu^2\right] \\
&= \frac{1}{N}\sum \left[E[x_i^2] - \mu^2\right] \quad - var[\mu_{MLE}] \\
&= \frac{1}{N}\sum Var(x_i) \quad - \frac{1}{N}\sigma^2 \\
&= \sigma^2 - \frac{1}{N}\sigma^2 \\
&= \boxed{\frac{N-1}{N}}\sigma^2
\end{aligned}$$

有偏：比真实的小. (估计小了)