♣ 最大概似法：

若將一分佈之p.d.f.$f(x|\theta)$視爲一$\theta$之函數, 假設$\theta \in \Omega$, 而將$x$固定, 且以$L(\theta|x)$表此函數, 並稱此爲**概似函數**(likelihood function)。一般而言, 設有隨機變數$X_1, \cdots, X_n$, 令$\boldsymbol{X} = (X_1, \cdots, X_n)$, $\boldsymbol{x} = (x_1, \cdots, x_n)$。若將其聯合p.d.f.$f(x_1, \cdots, x_n|\theta) = f(\boldsymbol{x}|\theta)$ 視爲一$\theta$之函數, 則得其概似函數$L(\theta|x_1, \cdots, x_n) = L(\theta|\boldsymbol{x})$。**最大概似法**(method of maximum likelihood), 就是找參數$\theta$之估計值$\hat{\theta}(\boldsymbol{x})$, 使得在$\theta = \hat{\theta}(\boldsymbol{x})$之下, 會**最可能**(most likely)產生數據$\boldsymbol{x}$。本來要先給出$\theta$, $\boldsymbol{X}$的分佈才完全決定。現由於觀測值$\boldsymbol{x}$的出現, 我們倒回去想, 怎樣的$\theta$, 才會使此$\boldsymbol{x}$在諸多可能的$\boldsymbol{x}$中拔得頭籌? 即當$\boldsymbol{X} = \boldsymbol{x}$, 尋找$\hat{\theta}(\boldsymbol{x})$, 使其滿足

$$\begin{aligned} L(\hat{\theta}(\boldsymbol{x})|\boldsymbol{x}) &= f(\boldsymbol{x}|\hat{\theta}(\boldsymbol{x})) = \max\{f(\boldsymbol{x}, \theta), \theta \in \Omega\} \\ &= \max\{L(\theta, \boldsymbol{x}), \theta \in \Omega\} \text{。} \end{aligned}$$

若如此的$\hat{\theta}(\boldsymbol{x})$存在, 便稱爲$\theta$之**最大概似估計值**(maxinum likelihood estimate 縮寫爲MLE)。至於$\hat{\theta}(\boldsymbol{X})$則爲**最大概似估計量**(maximum likelihood estimator), 縮寫亦爲MLE。又注意$\theta$在此可以是一向量。即有$\boldsymbol{\theta} = (\theta_1, \cdots, \theta_k)$的型式。

底下列出常見分佈之p.d.f.及最大概似估計量。

1. 伯努力分佈, $\mathcal{B}er(\theta)$

   $$f(x) = \theta^x(1-\theta)^{1-x}, \quad 0 < \theta < 1, \quad x = 0, 1$$

   最大概似估計量

   $$\hat{\theta} = \overline{X}_n$$

2. 二項分佈, $\mathcal{B}(m, p)$, 假設$m$已知, $p$未知

   $$f(x) = \binom{m}{x} p^x(1-p)^{m-x}, \quad m \geq 1, \quad 0 < p < 1, \quad x = 0, 1, \cdots, n$$

   最大概似估計量

   $$\hat{p} = \frac{\overline{X}_n}{m}$$

3. 幾何分佈(自0開始), $\mathcal{G}e(\theta)$

   $$f(x) = \theta(1-\theta)^x, \quad 0 < \theta < 1, \quad x = 0, 1, \cdots$$

   最大概似估計量

   $$\hat{\theta} = \frac{1}{1 + \overline{X}_n}$$

4. 負二項分佈, $\mathcal{NB}(r,p)$, 假設$r$已知, $p$未知

$$f(x) = \binom{x+r-1}{x} p^r (1-p)^x, \quad r \geq 1, \quad 0 < p < 1, \quad x = 0, 1, \cdots$$

最大概似估計量

$$\hat{p} = \frac{r}{r + \overline{X}_n}$$

5. 波松分佈, $\mathcal{P}(\lambda)$

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad \lambda > 0, \quad x = 0, 1, \cdots$$

最大概似估計量

$$\hat{\lambda} = \overline{X}_n$$

6. 離散型均勻分佈, $\mathcal{D}\text{-}\mathcal{E}(1,\theta)$

$$f(x) = \frac{1}{\theta}, \quad \theta = 1, 2, \cdots, \quad x = 1, \cdots, \theta$$

最大概似估計量

$$\hat{\theta} = X_{(n)}$$

7. 均勻分佈, $\mathcal{U}(\alpha,\beta)$, $\alpha,\beta$皆未知

$$f(x) = \frac{1}{\beta - \alpha}, \quad \alpha < x < \beta$$

最大概似估計量

$$\hat{\alpha} = X_{(1)}, \quad \hat{\beta} = X_{(n)}$$

8. 指數分佈, $\mathcal{E}(\lambda)$

$$f(x) = \lambda e^{-\lambda x}, \quad \lambda > 0, \quad x > 0$$

最大概似估計量

$$\hat{\lambda} = \frac{1}{\overline{X}_n}$$

9. 常態分佈, $\mathcal{N}(\mu,\sigma^2)$, $\mu,\sigma^2$皆未知

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma)}, \quad \mu \in R, \quad \sigma > 0, \quad x \in R$$

最大概似估計量：

$$\hat{\mu} = \overline{X}_n, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

10. 雙指數分佈, $\mathcal{D}\text{-}\mathcal{E}(\lambda)$

$$f(x) = \frac{\lambda}{2} e^{-\lambda|x|}, \quad \lambda > 0, \quad x \in R$$

最大概似估計量

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