M378K: April 7th, 2025.

Maximum Likelihood Estimation.

Likelihood.

Def'n. Given a random sample $Y_1, Y_2, \dots, Y_n$ from a discrete distribution D w/ an unknown parameter θ, the likelihood function is defined as

$$L(\theta; y_1, \dots, y_n) = p^{\theta}_{Y_1,\dots,Y_n}(y_1,\dots,y_n) = p^{\theta}_{Y_1}(y_1)\, p^{\theta}_{Y_2}(y_2) \cdots p^{\theta}_{Y_n}(y_n)$$
$$= p^{\theta}(y_1)\, p^{\theta}(y_2) \cdots p^{\theta}(y_n)$$

where p^{θ} is the pmf of D.

Example. Bernoulli.

$$Y_1, Y_2, \dots, Y_n \sim B(p) \qquad \boxed{p \longleftrightarrow \theta}$$

pmf of $B(p)$: $p(y) = \begin{cases} p & \text{for } y=1 \\ 1-p & \text{for } y=0 \end{cases}$

$$p(y) = p^{y}(1-p)^{1-y} \quad \text{for } y = 0,1.$$

$$L(p; y_1, y_2, \dots, y_n) = p^{y_1}(1-p)^{1-y_1} \cdot p^{y_2}(1-p)^{1-y_2} \cdot \dots \cdot p^{y_n}(1-p)^{1-y_n}$$
$$= \prod_{i=1}^{n} p^{y_i} \cdot \prod_{i=1}^{n} (1-p)^{1-y_i}$$
$$= p^{\sum_{i=1}^{n} y_i} \cdot (1-p)^{n - \sum_{i=1}^{n} y_i}$$

For computational reasons, take the $\ln(\cdot)$ $\boxed{\text{increasing}}$, get the log-likelihood, i.e.,

$$\ell(p; y_1, \dots, y_n) = \left(\sum_i y_i\right)\ln(p) + \left(n - \sum_i y_i\right)\ln(1-p)$$

Next, we differentiate with respect to p

$$\ell'(p; y_1, \dots, y_n) = \left(\sum_i y_i\right) \cdot \frac{1}{p} + \left(n - \sum_i y_i\right) \cdot (-1) \cdot \frac{1}{1-p}$$

We equate the derivative to zero:

$$\left(\sum_i y_i\right)\frac{1}{p} - \left(n-\sum_i y_i\right)\cdot\frac{1}{1-p} = 0 \qquad /\cdot p(1-p)$$

$$(1-p)\left(\sum_i y_i\right) - \left(n-\sum_i y_i\right)\cdot p = 0$$

$$\sum_i y_i - \cancel{p\left(\sum_i y_i\right)} - n\cdot p + \cancel{\left(\sum_i y_i\right)\cdot p} = 0$$

$$\boxed{\hat{p}_{MLE} = \frac{\sum y_i}{n} = \bar{y}}$$

Def'n. If $Y_1,\dots,Y_n$ come from a continuous dist'n D w/ pdf f^Θ, then the likelihood f'tion is

$$L(\Theta; y_1,\dots,y_n) = f^\Theta_{Y_1,\dots,Y_n}(y_1,\dots,y_n) = f^\Theta(y_1)\cdot f^\Theta(y_2)\cdots f^\Theta(y_n)$$

Example. $\boxed{\text{Normal.}}$

$$Y_1,\dots,Y_n \sim N(\mu,\sigma)$$

The pdf: $f(y) = \dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(y-\mu)^2}{2\sigma^2}}$ for all $y\in\mathbb{R}$

$$L(\mu; y_1,\dots,y_n) = \prod_{i=1}^{n}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(y_i-\mu)^2}{2\sigma^2}}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_i (y_i-\mu)^2}$$

$$\propto e^{-\frac{1}{2\sigma^2}\sum_i (y_i-\mu)^2}$$

proportional to

$$\ell(\mu; y_1,\dots,y_n) = \text{constant} - \frac{1}{2\sigma^2}\sum_i (y_i-\mu)^2$$

$$\ell'(\mu; y_1, \dots, y_n) = -\frac{1}{\cancel{2}\sigma^2} \sum_i \cancel{2}(y_i - \mu)(-1) = 0$$

$$\sum_{i=1}^{n}(y_i - \mu) = 0$$

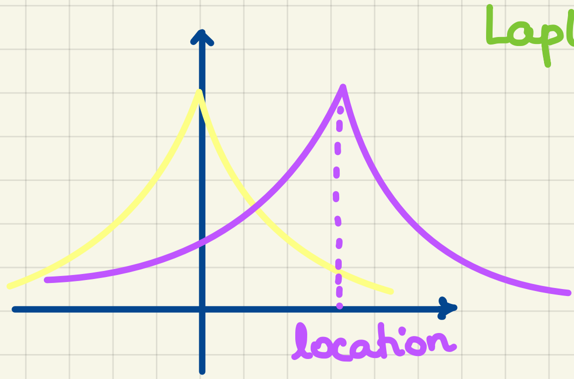
$$\sum_{i=1}^{n} y_i - n \cdot \mu = 0$$

$$\mu_{MLE} = \frac{\sum_{i=1}^{n} y_i}{n} = \bar{y}$$

$\square$

Laplace / Double Exponential

Example.



Look up on Wikipedia ☺