Estimation of Parameters

Maximum Likelihood Estimation (MLE)

population r.v. X

Likelihood Function corresponding to a random sample of size n

$(x_1, x_2, \ldots, x_n)$ is

$(X_1, X_2, \ldots, X_n)$

$$L(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_k) = \begin{cases} f_{X_1}(\theta_1, \theta_2, \ldots, \theta_k)\, f_{X_2}(\theta_1, \theta_2, \ldots, \theta_k) \cdots f_{X_n}(\theta_1, \theta_2, \ldots, \theta_k) & \text{if } X \text{ has discrete distribution with p.m.f. } f_i \\ f_X(x_1; \theta_1, \theta_2, \ldots, \theta_k)\, f_X(x_2; \theta_1, \theta_2, \ldots, \theta_k) \cdots f_X(x_n; \theta_1, \theta_2, \ldots, \theta_k) & \text{if } X \text{ has continuous distribution with p.d.f } f_X \end{cases}$$

If L is globally maximum for $\hat{\theta}_i = \hat{\theta}_i(x_1, x_2, \ldots, x_n)$ $i = 1, 2, \ldots n$

then $\hat{\theta}_i$ is the maximum likelihood estimate of θ_i (if they exists) $i = 1, 2, \ldots, n$.

$$\boxed{\hat{\theta}_1 = \hat{\theta}_1(x_1, x_2, \dots, x_n)} \xrightarrow{\text{good?}} \theta_1$$

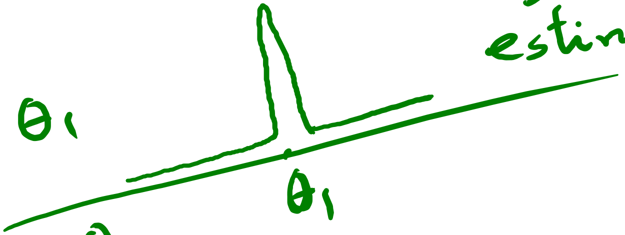
estimate

$\vdots$

$$\hat{\theta}_k = \hat{\theta}_k(x_1, \dots, x_n)$$

Sampling distribution

$\hat{\Theta}_1(X_1, X_2, \dots, X_n)$ is a good estimator of θ_1



(i) consistency $\hat{\Theta}_1 \xrightarrow{\text{in } P} \theta_1$

(ii) Unbiased $E(\hat{\Theta}_1) = \theta_1$

$N(m,\sigma)$ $\quad D=\left\{(m,\sigma): \begin{array}{l} m\in\mathbb{R} \\ \sigma\in\mathbb{R}^+ \end{array}\right\}$

Notes :

① Let L be continuously differentiable func. of $\theta_1,\theta_2,\dots,\theta_k$ in a domain $D=\{(\theta_1,\theta_2,\dots,\theta_k) : \theta_1\in R_1, \theta_2\in R_2, \dots, \theta_k\in R_k\}$

$R_i \subseteq \mathbb{R}$, $i=1,2,\dots,k$.

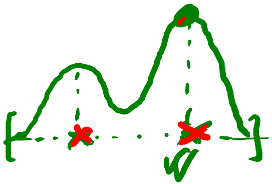
If L has "global maximum" at an interior point ⊛ $(\mathring{\theta}_1,\mathring{\theta}_2,\dots,\mathring{\theta}_k)$ of D, then

$$\not\Leftarrow \quad \boxed{\frac{\partial L}{\partial\theta_1}=0,\ \frac{\partial L}{\partial\theta_2}=0,\ \dots,\ \frac{\partial L}{\partial\theta_k}=0} \text{ at } (\mathring{\theta}_1,\mathring{\theta}_2,\dots,\mathring{\theta}_k).$$

(Likelihood eqn^{ns})

Algorithms

[Convex Optimization]



② $L > 0 \Rightarrow \log_e L$ is defined and $\log_e L$ attains the global maximum at the same point $(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k)$ where L attains global maximum. So $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k$ (if they exist) can be obtained by solving

$$\frac{\partial(\log_e L)}{\partial\theta_1} = 0, \quad \frac{\partial(\log_e L)}{\partial\theta_2} = 0, \ldots\ldots, \frac{\partial(\log_e L)}{\partial\theta_k} = 0$$

where $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k$ must satisfy

$$L(x_1, x_2, \ldots, x_n; \hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k) \geq L(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_k)$$

$$\forall\ (\theta_1, \theta_2, \ldots, \theta_k) \in D$$

global max!!

③ If L is global max$^{\underline{m}}$ at a point which is not an interior point of D (i.e. at a boundary point) or some or all the partial derivatives $\frac{\partial L}{\partial \theta_1}, \ldots\ldots, \frac{\partial L}{\partial \theta_k}$ do not exist at the points where L is global max$^{\underline{m}}$, then above method of calculus cannot be applied.

Prob 6: Find the MLE, say $\hat{p}$, of p of Binomial (N, p) population. Prove that $\hat{p}$ is consistent and unbiased estimate of p.

Sol. Assume N is known. p is the only unknown parameter where $0 < p < 1$. Let, X be the population random variable

$$P(X = x) = \binom{N}{x} p^x (1-p)^{N-x}, \quad x = \underbrace{0, 1, 2, \ldots, N}.$$

Let, $(x_1, x_2, \ldots, x_n)$ be a random sample of size n from the population of X.

$X_i \sim \text{Binomial}(N, p)$, $i = 1, 2, \ldots, n$

Likelihood function

$$\boxed{L} = L(x_1, x_2, \ldots, x_n ; p)$$

$$= \binom{N}{x_1} p^{x_1} (1-p)^{N-x_1} \binom{N}{x_2} p^{x_2} (1-p)^{N-x_2} \cdots \cdots \binom{N}{x_n} p^{x_n} (1-p)^{N-x_n}$$

$$= \binom{N}{x_1}\binom{N}{x_2} \cdots \binom{N}{x_n} p^{n\bar{x}} (1-p)^{nN - n\bar{x}}$$

where $\bar{x} = \dfrac{x_1 + x_2 + \cdots + x_n}{n}$.

$$\boxed{\log_e L} = \log_e\left\{\binom{N}{x_1} \cdots \cdots \binom{N}{x_n}\right\} + (n\bar{x})\log p + (nN - n\bar{x})\log(1-p).$$

$0 < p < 1$

For a fixed random sample $\log_e L$ is a fn of p.

For local maximum :

$$\frac{d}{dp}\left(\log_e L\right) = 0$$

$$\Rightarrow \frac{n\bar{x}}{p} - \frac{nN - n\bar{x}}{1-p} = 0$$

$$\Rightarrow \bar{x}(1-p) - (N-\bar{x})p = 0$$

$$\Rightarrow p = \frac{\bar{x}}{N}.$$

(critical point of $\log_e L$)

$$\frac{d^2(\log_e L)}{dp^2} = -\frac{n\bar{x}}{p^2} - \frac{nN - n\bar{x}}{(1-p)^2}$$

$$\left.\frac{d^2(\log_e L)}{dp^2}\right|_{p=\frac{\bar{x}}{N}} = -n\left[\frac{N^2}{\bar{x}} + \frac{N^2}{(N-\bar{x})}\right] < 0$$

Assuming $\bar{x} \neq 0$, $\bar{x} \neq N$

$\Rightarrow p = \frac{\bar{x}}{N}$ is a local maximum.

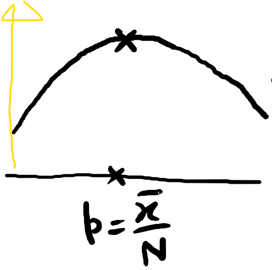
$$\frac{d}{dp}\left(\log_e L\right) = \frac{n\bar{x}}{p} - \frac{nN - n\bar{x}}{1-p} \quad > 0 \text{ if } 0 < p < \frac{\bar{x}}{N}$$

and

$$< 0 \text{ if } \frac{\bar{x}}{N} < p < 1.$$

$\Rightarrow$ $\log_e L$ is mon. strictly increasing in $0 < p \leq \frac{\bar{x}}{N}$

and mon. strictly decreasing in $\frac{\bar{x}}{N} \leq p \leq 1.$



graph of $\log_e L$.

$\Rightarrow$ $\log_e L$ has global max^{m} at $p = \frac{\bar{x}}{N}$

$\Rightarrow$ MLE: $\hat{p} = \frac{\bar{x}}{N}$

Check! (i) $\hat{p} \xrightarrow[\text{in } P]{} p$ (Apply LLNs of equal components)

(ii) $E(\hat{p}) = p$

$$\boxed{E(\hat{p}) = p}$$

Sol. (i) $\hat{p} = \frac{\bar{X}}{N} = \frac{X_1 + X_2 + \cdots + X_n}{nN}$

$$= \frac{\frac{X_1}{N} + \frac{X_2}{N} + \cdots + \frac{X_n}{N}}{n}$$

(i) $\frac{X_1}{N}, \frac{X_2}{N}, \cdots, \frac{X_n}{N}$ are mutually independent r.v.s

(ii) all have the same distribution as $\frac{X}{N}$

so, $E\left(\frac{X_i}{N}\right) = E\left(\frac{X}{N}\right) = \frac{E(X)}{N} = \frac{Np}{N} = p$

$$\boxed{\begin{gathered} X \sim \text{Binomial}(N, p) \\ E(X) = Np \end{gathered}}$$

$$E(\hat{p}) = E\left(\frac{\bar{X}}{N}\right) = E\left(\frac{X_1 + X_2 + \cdots + X_n}{nN}\right)$$

$$= \frac{E(X_1) + E(X_2) + \cdots + E(X_n)}{nN}$$

$$= \frac{\cancel{N} p \cancel{n}}{\cancel{n} \cancel{N}} = p$$

$\Rightarrow \hat{p}$ is an unbiased estimate of p.

Applying LLN (of equal components)

$$\frac{\frac{X_1}{N} + \frac{X_2}{N} + \cdots + \frac{X_n}{N}}{n} \xrightarrow[\text{in } P]{} p \quad \text{as } n \to \infty$$

$$\Rightarrow \hat{p} = \frac{\bar{X}}{\cancel{n}N} \xrightarrow[\text{in } P]{} p \quad \text{as } n \to \infty$$

$\Rightarrow$ $\hat{p}$ is a <u>consistent estimate</u> of p.

(7) Find the MLE of m and σ^2 for a normal (m,σ) population.

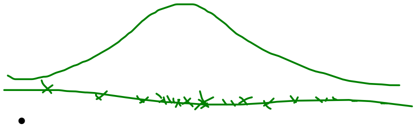
$$f_X(x;m,\sigma)=\frac{1}{\sqrt{2\pi}\,\sigma}\,e^{-\frac{1}{2\sigma^2}(x-m)^2}$$

<u>Sol.</u> $(x_1,x_2,\dots,x_n)$ be a random sample of size n.

$$L=L(x_1,x_2,\dots,x_n;\boxed{m,\sigma})=\frac{1}{(\sqrt{2\pi})^n\sigma^n}\,e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-m)^2}$$

for a fixed random sample, L is a fn of m and σ where $m\in(-\infty,\infty)$ and $\sigma\in(0,\infty)$.

to find $\hat{m}$ and $\hat{\sigma}$ for which L is globally maximum.



$$\log_e L = -\frac{n}{2}\log_e 2\pi - n\log_e \sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - m)^2$$

Likelihood eqn^s

$$\frac{\partial \log L}{\partial m} = 0, \qquad \frac{\partial \log L}{\partial \sigma} = 0$$

$$\Rightarrow \frac{\cancel{2}}{\cancel{2}\sigma^2}\sum_{i=1}^{n}(x_i - m) = 0$$

$$\Rightarrow \frac{n}{\sigma^2}(\bar{x} - m) = 0$$

$$\Rightarrow \hat{m} = \bar{x} \quad \checkmark$$

$$-\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i - m)^2 = 0$$

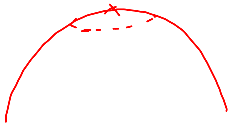
$$\Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2 = S^2 \quad \checkmark$$

Check! $\det\begin{bmatrix} \frac{\partial^2 \log L}{\partial m^2} & \frac{\partial^2 \log L}{\partial m\,\partial \sigma} \\ \frac{\partial^2 \log L}{\partial m\,\partial \sigma} & \frac{\partial^2 \log L}{\partial \sigma^2} \end{bmatrix} > 0$ is −ve definite

$\frac{\partial^2 \log L}{\partial m^2} < 0$

$f : \mathbb{R}^2 \to \mathbb{R}$

$$H_f(a,b) = \begin{bmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{xy}(a,b) & f_{yy}(a,b) \end{bmatrix}$$



$$D_1 = f_{xx}(a,b)$$

$$D_2 = \det H_f(a,b)$$

I. $f_x(a,b) = 0 = f_y(a,b) \Rightarrow (a,b)$ is a C.P. of f

II. If $D_1 < 0, \; D_2 > 0$: Maximum

If $D_1 > 0, \; D_2 > 0$: Minimum

If $D_2 < 0$: Saddle.

Since above is the only soln of of likelihood eqns

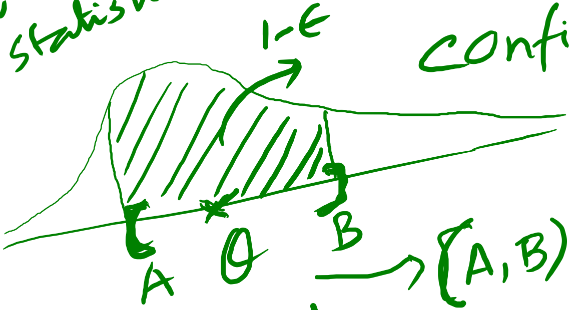
MLE $\hat{m} = \bar{x}$, $\hat{\sigma}^2 = S^2$

Interval Estimation

$A < \theta < B$

statistic 1

statistic 2



confidence coefficient $= 1-\epsilon$ $(\epsilon > 0)$

$[A, B] \to (1-\epsilon)\times 100\%$ confindence interval