Parameter Estimation

The problem:

The distribution of the population is known. The parameters are unknown. The estimation is based on a collection of 'n' experimental outcomes: $x_1, x_2, \ldots, x_n$. Each outcome is value of a r.v X_i. Then the set of R.V $X_1, X_2, \ldots X_n$ is called a sample.

Method 1: Method of Moments:

Equate the k^{th} moment of the sample with the k^{th} moment of the population.

$$M'_k = \sum_{i=1}^{n} \frac{X_i^k}{n}, \quad k = 1, 2, \ldots n$$

Sample Moment

&

k^{th} population moment

$$\mu'_k = E[X^k]$$

Eg let

$$f(x) = \begin{cases} (k+1)\, x^k & 0 < x < 1, \ k > 0 \\ 0 & \text{otherwise} \end{cases}$$

Estimate k

Soln : ∵ Only one parameter is to be estimated we equate only the first moment.

The population 1^{st} moment (mean)

$$\mu_1' = \int_0^1 (k+1)\,x^k \cdot x\,dx$$

$$= (k+1)\left[\int_0^1 x^{k+1}\,dx\right]$$

$$= (k+1)\left[\frac{x^{k+2}}{k+2}\right]_0^1$$

$$= (k+1)\left[\frac{1}{k+2}\right]$$

$$\Rightarrow \mu_1' = (k+1)/(k+2) \quad \text{---} (i)$$

Sample mean

$$M_1' = \sum_{i=1}^{n} \frac{x_i}{n} = \bar{x} \quad \text{---} (ii)$$

$$\therefore \bar{x} = \frac{k+1}{k+2}$$

$$\Rightarrow (k+2)\,\bar{x} = k+1$$

$$\Rightarrow k\bar{x} + 2\bar{x} = k+1$$

$$\Rightarrow k\bar{x} - k = 1 - 2\bar{x}$$

$$\Rightarrow k(\bar{x}-1) = 1 - 2\bar{x}$$

$$\Rightarrow \boxed{\hat{k} = \frac{1-2\bar{x}}{\bar{x}-1} = \frac{2\bar{x}-1}{1-\bar{x}}}$$

$\therefore$ As a numerical example

let $S = \{0.25,\ 0.45,\ 0.55,\ 0.75,$

$0.85,\ 0.85,\ 0.95,\ 0.90\}$

$\therefore\ n = 8$

$\Rightarrow\ \bar{x} = 5.55/8 = 0.69375$

$$\Rightarrow\ \hat{k} = \frac{2 * 0.69375 - 1}{1 - 0.69375}$$

$$\Rightarrow\ \boxed{\hat{k} = 1.265306}$$

rv

Maximum Likelihood Estimation
(M.L.)

Let $\vec{x} = (x_1, x_2, \ldots, x_n)$ be n observations from a population with distribution

$$f(x; \theta_1, \theta_2, \ldots, \theta_k). \quad \text{—— (1)}$$

Each of the 'n' draws constitute an observation from the population given by (1).

Hence each observation may be thought of as a random variable.

$\therefore$ x_1 is outcome of r.v. X_1
x_2 " " " " X_2
$\vdots$
x_n " " " " X_n

We will use $x_1, x_2, \ldots x_n$ instead of $X_1, X_2, \ldots X_n$ to denote r.v also.

Also all the r.v $x_1, x_2, \ldots x_n$ follow the distribution (1) i.e

$$\left.\begin{aligned} x_1 &\sim f(x; \theta) \\ x_2 &\sim f(x; \theta) \\ &\vdots \\ x_n &\sim f(x; \theta) \end{aligned}\right] \text{if only one parameter '}\theta\text{' is to be estimated.}$$

$$\therefore x_i \sim f(x; \theta) \quad \forall\, i = 1, 2, \ldots, n$$

$\because$ The sample $\vec{x}$ taken from 'f' is

$$f(\vec{x};\theta) = f(x_1, x_2, \dots, x_n;\theta) \quad \text{--- (2)}$$

But $\because$ $x_1, x_2, \dots x_n$ are independent

$\therefore$ $f(x_1, x_2, \dots, x_n;\theta)$

$= f(x_1;\theta) . f(x_2;\theta) \dots f(x_n;\theta)$

$$\Rightarrow \boxed{f(\vec{x};\theta) = \prod_{i=1}^{n} f(x_i;\theta)}$$

Joint PDF of the random sample $\vec{x}$.

The motivation behind M.L

The value of the parameter θ ~~is~~ should be such that the probability of obtaining $\vec{x}$ from $f(\vec{x};\theta)$ is maximized.

i.e. θ^* is the optimal val. of $\vec{\theta}$ if

$$f(\vec{x}, \theta^*) > f(\vec{x}, \theta) \quad \forall \theta$$

$\therefore$ The joint pdf $f(\vec{x_i}; \theta)$ is called the likelihood f^n.

$$\therefore \boxed{L(\vec{x}; \theta) = \prod_{i=1}^{n} f(\vec{x_i}; \theta)}$$

likelihood f^n.

Objective of ML is to maximize $L(\vec{x}; \theta)$ w.r.t. θ.

This value of θ is w/a θ^*

This val. of θ^* will be a f^n of given observations.

$\therefore$ We must have

(i) $\frac{\partial}{\partial \theta} L(\vec{x}; \theta) = 0$ &

(ii) $\frac{\partial^2}{\partial^2 \theta} L(\vec{x}; \theta) < 0$ at $\theta = \theta^*$

Ex. Poisson Distribution:

$$f(x_i) = \frac{e^{-\lambda} \cdot \lambda^{x_i}}{x_i!}, \quad x_i = 1, 2, 3 \ldots$$

i.e. $f(x_i) = P(\lambda) \; \forall \; i = 1, 2, \ldots n$

$\therefore L(\bar{x};\lambda) = f(x_1;\lambda) \cdot f(x_2;\lambda) \cdots f(x_n;\lambda)$

$$= \frac{e^{-\lambda} \cdot \lambda^{x_1}}{x_1!} \cdot \frac{e^{-\lambda} \cdot \lambda^{x_2}}{x_2!} \cdots \frac{e^{-\lambda} \cdot \lambda^{x_n}}{x_n!}$$

$$= \frac{e^{-n\lambda} \lambda^{x_1+x_2+\cdots+x_n}}{\prod_{i=1}^{n} x_i!}$$

$$\therefore \boxed{L(\bar{x};\lambda) = \frac{e^{-n\lambda} \cdot \lambda^{\sum_i x_i}}{\prod_i x_i!}}$$

Taking log.

$$\log L = -n\lambda + \left(\sum x_i\right) \log \lambda - \sum_{i=1}^{n} \log(x_i!)$$

$$\therefore \frac{\partial L}{\partial \lambda} = -n + \frac{\sum x_i}{\lambda}$$

Equating to 0

$$-n\lambda + \sum x_i = 0$$

$$\Rightarrow \boxed{\lambda^* = \frac{1}{n} \sum_{i=1}^{n} x_i}$$

$$\Rightarrow \lambda^* = \bar{x} = \text{Sample Mean}$$