If $x_1 = 0.2$, $x_2 = 0.6$, $x_3 = 0.5$, $x_4 = 0.3$ is a random sample of size 4, then what is the estimate of θ?

Solution:

To find an estimator, we shall equate the first population moment $E(x)$ to the first sample moment.

The population moment is given by

$$E(X) = \int_0^1 x \cdot f_X(x;\theta)\, dx$$

$$= \int_0^1 x \cdot \theta\, x^{\theta-1}\, dx$$

$$= \frac{\theta}{\theta+1}\left[x^{\theta+1}\right]_{x=0}^{x=1}$$

$$E(X) = \frac{\theta}{\theta+1}$$

WKT the first sample moment $M_1 = \overline{X}$

$$\Rightarrow E(X) = M_1$$

$$\Rightarrow \frac{\theta}{\theta+1} = \overline{x} \Rightarrow \theta = \frac{\overline{X}}{1-\overline{x}}$$

$$\Rightarrow \boxed{\hat{\theta} = \frac{\overline{X}}{1-\overline{X}}}$$ is an estimator of the parameter θ

Since $x_1 = 0.2$, $x_2 = 0.6$, $x_3 = 0.5$, $x_4 = 0.3$, we have $\overline{x} = \frac{1}{4}\sum_{i=1}^{4} x_i$

$\overline{x} = 0.4$

$\therefore$ The estimate of θ is $\frac{\overline{x}}{1-\overline{x}} = \frac{0.4}{1-0.4} = \frac{2}{3}$

Exercise:

P1) Suppose $X_1, X_2, \ldots, X_7$ is a random sample from a population X with pdf,

$$f(x;\beta) = \begin{cases} \dfrac{x^6 e^{-x/\beta}}{\Gamma 7 \, \beta^7}, & \text{if } 0 < x < \infty \\ 0 & \text{Otherwise} \end{cases}$$

Find an estimator of β by the Moment Method

Hint: Use $\Gamma\alpha = \int_0^{\infty} x^{\alpha-1} \cdot e^{-x} \, dx \, , \quad \alpha > 0$

Ans: $\hat{\beta} = \frac{1}{7}\bar{X}$.

P2). Suppose $X_1, \ldots X_n$ is a random sample from a Population X with density function.

$$f(x;\theta) = \begin{cases} \dfrac{1}{\theta}, & \text{if } 0 < x < \theta \\ 0 & \text{Otherwise.} \end{cases}$$

Find an estimator of θ by the Moment Method.

Ans: $\hat{\theta} = 2\bar{X}$

Maximum Likelihood Method: [R.A. Fisher, 1912]

As the name implies, the estimator will be the value of the parameter that maximizes the likelihood function.. Let us take a random sample of size 'n' $X_1, X_2, \ldots, X_n$ from a discrete probability distribution represented by $f_{\underset{\sim}{X}}(x_1, \ldots, x_n; \theta)$, where θ is a single Parameter of the distribution.

Let $x_1, x_2, \ldots x_n$ be the observed values in a random sample of size n. Then the likelihood function of the sample is.

$$L(\theta) = L(x_1, x_2, \ldots x_n ; \theta) = f_{\underline{x}}(x_1 \ldots x_n ; \theta)$$

$$= f(x_1 ; \theta) f(x_2 ; \theta) \ldots f(x_n ; \theta)$$

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

Since x_i's are independent

Notation: $f_{\underline{x}}(x_1, \ldots x_n ; \theta) \rightarrow$ represents
Joint PMF $P(x_1, \cdots x_n ; \theta)$ for Discrete
Joint PDF $f(x_1, \ldots x_n ; \theta)$ for Continuous Distributions.

Note that the likelihood function is now a function of only the unknown parameter θ. The maximum likelihood estimator (MLE) of θ is the value of θ that maximizes the likelihood function $L(\theta)$.

Remark: [Interpretation]

In the case of discrete r.v, the likelihood function of the sample $L(\theta)$ is just the probability

$$P(X_1 = x_1, X_2 = x_2, \ldots X_n = x_n)$$

i.e, $L(\theta)$ is just the probability of obtaining the sample values $x_1, x_2, \ldots, x_n$. Here the MLE is an estimator that maximizes the probability of occurrence of the sample values.

Note:

To maximize this likelihood, we consider the critical or stationary point(s) by taking derivatives with respect to all unknown parameter(s) and equating them to zero. A nice shortcut is to take logarithms first. Differentiating the sum

$$\ln \prod_{i=1}^{n} f(x_i;\theta) = \sum_{i=1}^{n} \ln f(x_i;\theta)$$

is easier than differentiating the product $\prod_{i=1}^{n} f(x_i;\theta)$

Besides, since logarithm is an increasing function, the log-likelihood $\ln L(\theta)$ is maximized at exactly the same point as the likelihood ~~of~~ $L(\theta)$.

P1. Poisson Distribution: MLE

Consider a poisson distribution with pmf

$$P(x;\alpha) = \frac{e^{-\alpha}\alpha^{x}}{x!}, \quad x = 0, 1, 2, \ldots$$

Suppose that a random sample $x_1, x_2, \ldots, x_n$ is taken from the distribution. What is the Maximum Likelihood Estimator of α?

Scanned with CamScanner

Solution:

The likelihood function $L(\alpha)$ of a random Sample of size n is

$$L(\alpha) = L(x_1, \dots x_n; \alpha) = \prod_{i=1}^{n} f(x_i; \alpha)$$

$$= \prod_{i=1}^{n} \frac{e^{-\alpha} \alpha^{x_i}}{x_i!}$$

$$= \frac{e^{-n\alpha} \alpha^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$$

Take natural Logarithm on both sides,

$$\ln L(\alpha) = \ln\left(e^{-n\alpha} \cdot \alpha^{\sum_{i=1}^{n} x_i}\right) - \ln\left(\prod_{i=1}^{n} x_i!\right)$$

$$= \ln\left(e^{-n\alpha}\right) + \ln\left(\alpha^{\sum_{i=1}^{n} x_i}\right) - \ln\left(\prod_{i=1}^{n} x_i!\right)$$

$$\ln L(\alpha) = -n\alpha + \sum_{i=1}^{n} x_i \cdot \ln\alpha - \ln\left(\prod_{i=1}^{n} x_i!\right)$$

Diff. w.r.t. 'α' on both sides.

$$\frac{d}{d\alpha} \ln L(\alpha) = -n + \sum_{i=1}^{n} x_i \cdot \frac{1}{\alpha}$$

Equating this to zero and soliving for the parameter α yields. $-n + \sum_{i=1}^{n} x_i \frac{1}{\alpha} = 0 \Rightarrow \alpha = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$

The second derivative of the $\ln L(\alpha)$ is negative.

$$\frac{d^2}{d\alpha^2} \ln L(\alpha) = -\frac{1}{\alpha^2}\sum_{i=1}^{n} x_i < 0$$

Scanned with CamScanner

which implies that the solution above indeed is a maximum. [using second derivative test]

$\therefore$ the maximum Likelihood estimator of α

is $\boxed{\hat{\alpha}_{MLE} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}}$

(P2) Exponential Distribution: MLE $f_X(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x < 0 \end{cases}$

Let X be exponentially distributed with parameter λ. The likelihood function of a random sample of size n is

$$L(\lambda) = f(x_1, x_2, \dots x_n; \lambda) = \prod_{i=1}^{n} f(x_i; \lambda)$$

$$= \prod_{i=1}^{n} \lambda e^{-\lambda x_i}$$

$$= \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$

The log likelihood is

$$\ln L(\lambda) = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i$$

Now $\frac{d}{d\lambda} \ln L(\lambda) = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i$ $\Rightarrow \lambda = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\bar{x}}$

and upon equating this to zero, we obtain $\cancel{\lambda = \frac{1}{n}\sum_{i=1}^{n} x_i}$.

The second derivative of $\ln L(\lambda)$ is negative

$\Rightarrow$ The MLE of λ is $\hat{\lambda}_{MLE} = \cancel{\frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}} = \frac{n}{\sum_{i=1}^{n} x_i}$

$$\boxed{\hat{\lambda}_{MLE} = \frac{1}{\bar{x}}}$$

Scanned with CamScanner

(P3) Normal Distribution MLE:

Consider a random sample $x_1, \dots x_n$ from a normal distribution $N(\mu, \sigma^2)$. Find the MLE of for μ & σ^2.

Solution:

The Likelihood function for the normal distribution is

$$L(\mu, \sigma^2) = L(x_1, x_2, \dots, x_n; \mu, \sigma^2) = \prod_{i=1}^{n} f(x_i; \mu, \sigma^2).$$

$$= \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

Taking logarithms. gives

$$\ln L(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}$$

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^{n}\frac{(x_i-\mu)}{\sigma^2}$$

$$\frac{\partial \ln L}{\partial(\sigma^2)} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

Setting both derivatives equal to zero,

we obtain.

$$\sum_{i=1}^{n} x_i - n\mu = 0 \quad \& \quad n\sigma^2 = \sum_{i=1}^{n}(x_i-\mu)^2$$

Using the second derivative test for functions of two variables, we find the Maximum Likelihood Estimators of μ & σ^2 are

$$\boxed{\begin{aligned}\hat{\mu} &= \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x} \\ \hat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\end{aligned}}$$

Remark: The second derivative Test for functions of two variables $f(x,y)$. $\quad f_x = \frac{\partial f}{\partial x}$, $f_{xy} = \frac{\partial^2 f}{\partial x \partial y}$

Let (a,b) be a stationary point, so that $f_x = f_y = 0$ at (a,b)

Then.

- if $f_{xx}f_{yy} - f_{xy}^2 > 0$ at (a,b), then (a,b) is either a maximum or a minimum.
  - if $f_{xx} < 0$ and $f_{yy} < 0$ at (a,b), then (a,b) is a maximum point
  - if $f_{xx} > 0$ and $f_{yy} > 0$ at (a,b), then (a,b) is a minimum point
- If $f_{xx}f_{yy} - f_{xy}^2 < 0$ at (a,b), then (a,b) is a Saddle point.
- if $f_{xx}f_{yy} - f_{xy}^2 = 0$, then the test fails.