Central Limit Theorem

If $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of independent RVs with $E(X_i) = \mu_i$ and $Var(X_i) = \sigma_i^2$, $i = 1, 2, \ldots$, then as $n \to \infty$, the distribution of the sum of these n random variables, namely $S_n = X_1 + X_2 + \cdots + X_n$ tends to the normal distribution with mean μ and variance σ^2, where $\mu = \sum_{i=1}^{n} \mu_i$ and $\sigma^2 = \sum_{i=1}^{n} \sigma_i^2$.

Or

If $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of independent identically distributed RVs with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$, $i = 1, 2, \ldots$, and if $S_n = X_1 + X_2 + \cdots + X_n$, then under certain general conditions, S_n follows a normal distribution with mean $n\mu$ and variance $n\sigma^2$ as $n \to \infty$.

Cor: If $\bar{X} = \frac{1}{n}(X_1 + \cdots + X_n)$, then $E(\bar{X}) = \mu$ and $Var(\bar{X}) = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}$.

$\therefore$ $\bar{X}$ follows a normal distribution with mean μ and S.D. $\frac{\sigma}{\sqrt{n}}$ i.e. $\bar{X}$ follows $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ as $n \to \infty$.

Ex An electrical firm manufactures light bulbs that have a length of life that is approximately normally distributed with mean equal to 800 hours and a standard deviation of 40 hours. Find the probability that a random sample of 16 bulbs will have an average life of less than 775 hours.

(Given $P(Z < -2.5) = 0.0062$ or $P(0 < Z < 2.5) = 0.4938$)

Sol: The sampling distribution of $\bar{X}$ will be approximately normal with $\mu_{\bar{X}} = 800$ and $\sigma_{\bar{X}} = \frac{40}{\sqrt{16}} = \frac{40}{4} = 10$.

Reqd prob $= P(\bar{X} < 775) = P(Z < -2.5)$

$\therefore$ Reqd prob $= 0.0062$ or $0.5 - P(0 < Z < 2.5)$

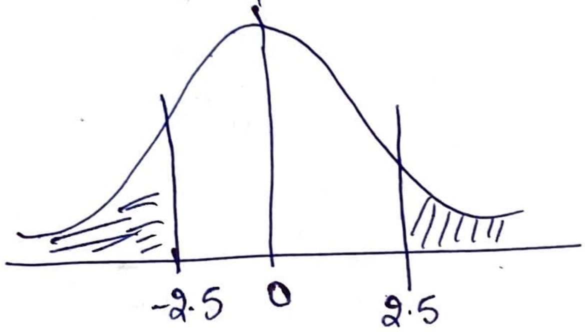
$= 0.5 - 0.4938$

$= 0.0062$

$Z = \frac{\bar{X} - \mu}{\sigma}$

$Z = \frac{775 - 800}{10}$

$= -2.5$



Ex. Travelling between two campuses of a university in a city via bus takes, on average, 28 minutes with a standard deviation of 5 minutes. In a given week, a bus transported passengers 40 times. What is the probability that the average transport time was more than 30 minutes? ($P(Z < 2.53) = 0.9943$ or $P(0 < Z < 2.53) = 0.4943$)

Sol: $\mu = 28$, $\sigma = 5$, $n = 40$

$\mu_{\bar{X}} = 28$, $\sigma_{\bar{X}} = \frac{5}{\sqrt{40}} = \frac{5}{6.3246} = 0.7906$

$P(\bar{X} > 30) = P(Z > 2.53)$

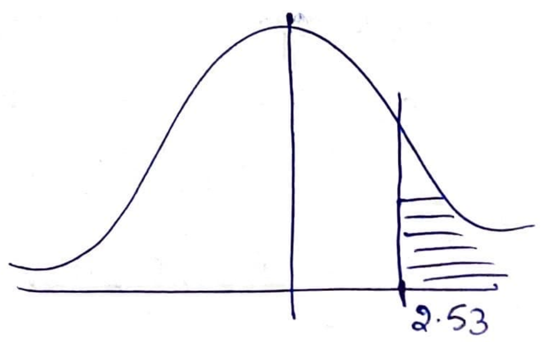
$= 1 - P(Z < 2.53)$ or $0.5 - P(0 < Z < 2.53)$

$= 1 - 0.9943$ or $0.5 - 0.4943$

$= 0.0057$ or 0.0057

$Z = \frac{\bar{X} - \mu}{\sigma} = \frac{30 - 28}{0.7906} = 2.53$

$\Rightarrow P(\bar{X} > 30) = 0.0057$.

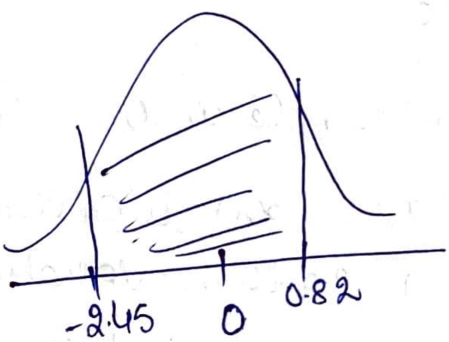


Ex: If $X_1, X_2, \ldots, X_n$ are Poisson variates with parameter $\lambda = 2$, use CLT to estimate $P(120 \le S_n \le 160)$, with $S_n = X_1 + X_2 + \cdots + X_n$ and $n = 75$. (Given $P(Z < -2.45) = 0.0071$, $P(Z < 0.82) = \cancel{0.7967}\ 0.7939$, or $P(0 < Z < 0.82) = 0.2939$, $P(0 < Z < 2.45) = 0.4929$)

Sol: $E(X_i) = 2$, $Var(X_i) = 2$

By CLT, S_n follows normal distribution with mean $n\mu = 75 \times 2 = 150$ and S.D. $\sqrt{n}\sigma = \sqrt{75} \times \sqrt{2} = \sqrt{150}$.

$$P(120 \le S_n \le 160) = P\left(\frac{120-150}{\sqrt{150}} \le Z \le \frac{160-150}{\sqrt{150}}\right)$$



$$= P(-2.45 \le Z \le 0.82)$$
$$= P(Z < 0.82) - P(Z < -2.45)$$
$$= \cancel{0.7967}\ 0.7939 - 0.0071$$
$$= \cancel{0.7896}\ 0.7868$$

or $P(0 < Z < 0.82) + P(0 < Z < 2.45)$

$= 0.7868$.

Ex: The lifetime of a certain brand of an electric bulb may be considered a RV with mean 1200 h and standard deviation 250 h. Find the probability that the average lifetime of 60 bulbs exceeds 1250 h. (Given $P(Z < 1.55) = 0.9394$ or $P(0 < Z < 1.55) = 0.4394$)

Sol: $E(X_i) = 1200$, $Var(X_i) = 250^2 \Rightarrow \sigma(X_i) = 250$

$$\mu_{\bar{X}} = 1200, \quad \sigma_{\bar{X}} = \frac{250}{\sqrt{60}} = 32.2749$$

$$P(\bar{X} > 1250) = P(Z > 1.55)$$
$$= 1 - P(Z < 1.55) \quad \text{or} \quad 0.5 - P(0 < Z < 1.55)$$
$$= \cancel{0.600}\ 0.0606$$

Estimators

Parameter: A numerical value summarizing all the data of an entire population. Eg. Population mean (μ), Population variance (σ^2) etc.

Statistic: A numerical value summarizing the sample data.

E.g. Sample mean ($\bar{x}$), sample variance (s^2) etc.

Estimator: Any function of the random sample $x_1, x_2, \ldots, x_n$ that are being observed, say, $T_n(x_1, x_2, \ldots, x_n)$ is called a statistic. Clearly, a statistic is a random variable. If it is used to estimate an unknown parameter θ of the distribution, it is called an estimator. A particular value of the estimator, say, $T_n(x_1, x_2, \ldots, x_n)$ is called an estimate of θ.

Characteristics of Estimators

1. Unbiasedness
2. Consistency
3. Efficiency
4. Sufficiency

Unbiased Estimators

Def: A statistic $\hat{\theta}$ is said to be an unbiased estimator of the parameter θ if

$$\mu_{\hat{\theta}} = E(\hat{\theta}) = \theta.$$

Remark If $E(\hat{\theta}) > \theta \Rightarrow \hat{\theta}$ is positively biased and if $E(\hat{\theta}) < \theta \Rightarrow$ Negatively biased.

Ex If $x_1, x_2, \ldots, x_n$ is a random sample from a normal population $N(\mu, 1)$. Show that $t = \frac{1}{n}\sum_{i=1}^{n} x_i^2$ is an unbiased estimator of μ^2+1.

Sol: $E(x_i) = \mu$, $Var(x_i) = 1 \ \forall\ i = 1, 2, \ldots, n.$

We know that $Var(x_i) = E(x_i^2) - [E(x_i)]^2$

$$\Rightarrow 1 = E(x_i^2) - \mu^2$$

$$\Rightarrow E(x_i^2) = \mu^2 + 1$$

Now,
$$E(t) = E\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) = \frac{1}{n}\sum_{i=1}^{n} E(x_i^2)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \mu^2+1$$

$$= \frac{1}{n}(\mu^2+1)n$$

$$= \mu^2+1.$$

$\Rightarrow E(t) = \mu^2+1$

$\Rightarrow$ t is an unbiased estimator of μ^2+1.

Ex If T is an unbiased estimator of θ, show that T^2 is ~~an unbi~~ a biased estimator for θ^2.

Sol: Since T is an ~~tto~~ unbiased estimator of θ.

$\Rightarrow E(T) = \theta$.

Also, $Var(T) = E(T^2) - (E(T))^2$

$= E(T^2) - \theta^2$

$\Rightarrow E(T^2) = \theta^2 + Var(T)$, $(Var(T) > 0)$

Since $E(T^2) \neq \theta^2$

$\Rightarrow T^2$ is not an unbiased estimator i.e. T^2 is a biased estimator of θ.

Ex: If X is a binomial random variable, show that

(a) $\hat{P} = \frac{X}{n}$ is an unbiased estimator of p,

(b) $P' = \dfrac{X + \frac{\sqrt{n}}{2}}{n + \sqrt{n}}$ is a biased estimator of p.

(c) P' becomes unbiased as $n \to \infty$.

Sol: (a) $E(\hat{P}) = E\left(\frac{X}{n}\right) = \frac{1}{n} E(X) = \frac{1}{n}(np) = \frac{\cancel{n}p}{\cancel{n}} = p$

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

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

(b) $E(P') = E\left(\dfrac{X + \frac{\sqrt{n}}{2}}{n + \sqrt{n}}\right) = \dfrac{1}{n+\sqrt{n}}\left(E(X) + \dfrac{\sqrt{n}}{2}\right)$

$= \dfrac{1}{n+\sqrt{n}}\left(np + \dfrac{\sqrt{n}}{2}\right) \neq p$

$\Rightarrow P'$ is a biased estimator of p.

(C) $\lt_{n\to\infty} \frac{np+\sqrt{n}/2}{n+\sqrt{n}} = \lt_{n\to\infty} \frac{p+1/2\sqrt{n}}{1+1/\sqrt{n}} = p.$

$\Rightarrow$ p' is an unbiased estimator of p as $n\to\infty$.

Ex Show that $\frac{[\Sigma x_i(\Sigma x_i - 1)]}{n(n-1)}$ is an unbiased estimator of θ^2, for the sample $x_1, x_2, ---, x_n$ drawn on X which takes the values 1 or 0 with respective probabilities θ and $1-\theta$.

Sol: Since $x_1, x_2, ---, x_n$ is a random sample from Bernoulli population with parameter θ,

$T = \sum_{i=1}^{n} x_i \sim B(n, \theta).$

$\Rightarrow E(T) = n\theta, \ Var(T) = n\theta(1-\theta)$

$$\therefore E\left\{\frac{\Sigma x_i(\Sigma x_i - 1)}{n(n-1)}\right\} = E\left\{\frac{T(T-1)}{n(n-1)}\right\}$$

$$= \frac{1}{n(n-1)}\left\{E(T^2) - E(T)\right\}$$

$[Var(T) = E(T^2) - E(T)^2]$

$n\theta(1-\theta) = E(T^2) - n^2\theta^2$

$$= \frac{1}{n(n-1)}\left\{n\theta(1-\theta) + n^2\theta^2 - n\theta\right\}$$

$$= \frac{1}{n(n-1)}\left\{\cancel{n\theta} - n\theta^2 + n^2\theta^2 - \cancel{n\theta}\right\}$$

$$= \frac{1}{n(n-1)} n\theta^2(n-1)$$

$$= \theta^2.$$

Ex: Let X be distributed in the Poisson form with parameter θ. Show that only unbiased estimator of $e^{-(k+1)\theta}$, $k>0$ is $T(X) = -k^X$ so that $T(X)>0$ if x is even and $T(x)<0$ if x is odd.

Sol: Since X follows Poisson distribution with ~~mean~~ parameter θ.

Thus, $E(X) = \theta$ and $Var(X) = \theta$.

Now, $E(T) = E(-k^X)$

$$= \sum_{x=0}^{\infty} -k^x \left(\frac{e^{-\theta}\theta^x}{x!} \right)$$

$$= e^{-\theta} \sum_{x=0}^{\infty} \frac{(-\theta k)^x}{x!}$$

$$\left[\sum_{x=0}^{\infty} \frac{a^x}{x!} = 1 + a + \frac{a^2}{2!} + \frac{a^3}{3!} + \cdots = e^a \right]$$

$$= e^{-\theta} e^{-\theta k}$$

$$= e^{-k\theta - \theta} = e^{-(k+1)\theta}$$

$\Rightarrow$ $T(X) = (-k)^X$ is an unbiased estimator for $e^{-(k+1)\theta}$, $k>0$.

Unbiased Estimator: An estimator $T_n = T(x_1, x_2, \ldots, x_n)$ is said to be an unbiased estimator of $\gamma(\theta)$ if $E(T_n) = \gamma(\theta)$ for all $\theta \in \Theta$, Θ is the parameter space.

Consistent Estimator

An estimator $T_n = T(x_1, x_2, \dots, x_n)$ based on a random sample of size n is said to be consistent estimator of $\gamma(\theta)$ if T_n converges to $\gamma(\theta)$ in probability i.e.

$$T_n \xrightarrow{P} \gamma(\theta) \text{ as } n \to \infty.$$

In other words, T_n is a consistent estimator of $\gamma(\theta)$ if for every $\epsilon > 0$, $\eta > 0$, there exists a positive integer $n \geq m(\epsilon, \eta)$ such that

$$P(|T_n - \gamma(\theta)| < \epsilon) \to 1 \text{ as } n \to \infty$$

$$\Rightarrow P(|T_n - \gamma(\theta)| < \epsilon) > 1 - \eta \quad \forall\ n \geq m,$$

where m is some very large value of n.

Remark :- 1. Since consistency is a property concerning the behaviour of an estimator for indefinitely large values of the sample size n, i.e., as $n \to \infty$, nothing is regarded of its behaviour for finite n.

2. If there exists a consistent estimator, say, T_n of $\gamma(\theta)$, then infinitely many such estimators can be constructed, e.g.,

$$T_n' = \left(\frac{n-a}{n-b}\right) T_n = \left(\frac{1 - a/n}{1 - b/n}\right) T_n \to T_n \text{ as } n \to \infty.$$
$$\to \gamma(\theta)$$

Hence for different value of a, b, T_n' is also consistent for $\gamma(\theta)$.

Invariance Property of Consistent Estimator

If T_n is consistent estimator of γ and f is a continuous function, then $f(T_n)$ is a consistent estimator of $f(\gamma)$.

Sufficient Conditions for Consistency

Let $\{T_n\}$ be a sequence of estimators such that

(i) $E(T_n) \to \gamma$ as $n \to \infty$

(ii) $Var(T_n) \to 0$ as $n \to \infty$,

Then, T_n is a consistent estimator of γ.

Ex: Show that in sampling from a $N(\mu, \sigma^2)$ population, the sample mean is a consistent estimator of μ.

Sol: Let $x_1, x_2, ---, x_n$ be a sample of size n drawn from a $N(\mu, \sigma^2)$ population.

Thus, $E(x_i) = \mu$, $Var(X_i) = \sigma^2$

Now, the sample mean $\overline{X}$ is defined as

$$\overline{X} = \frac{x_1 + x_2 + --- + x_n}{n}$$

$$E(\overline{X}) = \frac{1}{n}\left[E(x_1) + ---- + E(x_n)\right]$$

$$= \frac{1}{n}\left[\mu + \mu + -- + \mu\right]$$

$$= \mu$$

$E(\overline{X}) = \mu$ as $n \to \infty$.

$$Var(\overline{X}) = Var\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right)$$

$$= \frac{1}{n^2}\left[Var(x_1) + Var(x_2) + \cdots + Var(x_n)\right]$$

$$= \frac{1}{n^2}\left[\sigma^2 + \sigma^2 + \cdots + \sigma^2\right]$$

$$= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \to 0 \text{ as } n \to \infty.$$

$Var(\overline{X}) \to 0$ as $n \to \infty$.

Since $E(\overline{X}) \to \mu$ and $Var(\overline{X}) \to 0$ as $n \to \infty$, thus $\overline{X}$ is a consistent estimator of μ.

Ex: Let T_1 and T_2 be consistent estimators of γ_1 and γ_2 respectively. Prove that $aT_1 + bT_2$ is a consistent estimator of $a\gamma_1 + b\gamma_2$, where a and b are constants and independent of population.

Sol: Since T_1 and T_2 are consistent estimators of γ_1 and γ_2.

$\Rightarrow$ $E(T_1) \to \gamma_1$ and $Var(T_1) \to 0$ as $n \to \infty$

and $E(T_2) \to \gamma_2$ and $Var(T_2) \to 0$ as $n \to \infty$.

Now, $E(aT_1 + bT_2) = aE(T_1) + bE(T_2)$

$\to a\gamma_1 + b\gamma_2$ as $n \to \infty$

$\Rightarrow$ $E(aT_1 + bT_2) \to a\gamma_1 + b\gamma_2$ as $n \to \infty$.

and $Var(aT_1 + bT_2) = a^2 Var(T_1) + b^2 Var(T_2)$

$\to 0$ as $n \to \infty$

Hence, $aT_1 + bT_2$ is consistent estimator of $a\gamma_1 + b\gamma_2$.

Invariance Property of Consistent Estimator

If T_n is consistent estimator of γ and f is a continuous function, then $f(T_n)$ is a consistent estimator of $f(\gamma)$.

Sufficient Conditions for Consistency

Let $\{T_n\}$ be a sequence of estimators such that

(i) $E(T_n) \to \gamma$ as $n \to \infty$

(ii) $Var(T_n) \to 0$ as $n \to \infty$,

Then, T_n is a consistent estimator of γ.

Ex: Show that in sampling from a $N(\mu, \sigma^2)$ population, the sample mean is a consistent estimator of μ.

Sol: Let $x_1, x_2, \dots, x_n$ be a sample of size n drawn from a $N(\mu, \sigma^2)$ population.

Thus, $E(x_i) = \mu$, $Var(X_i) = \sigma^2$

Now, the sample mean $\overline{X}$ is defined as

$$\overline{X} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

$$E(\overline{X}) = \frac{1}{n}[E(x_1) + \dots + E(x_n)]$$

$$= \frac{1}{n}[\mu + \mu + \dots + \mu]$$

$$= \mu$$

$E(\overline{X}) = \mu$ as $n \to \infty$.

$$Var(\overline{X}) = Var\left(\frac{x_1 + x_2 + \text{—} + x_n}{n}\right)$$

$$= \frac{1}{n^2}\left[Var(x_1) + Var(x_2) + \text{—} + Var(x_n)\right]$$

$$= \frac{1}{n^2}\left\{\sigma^2 + \sigma^2 + \text{—} + \sigma^2\right\}$$

$$= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \to 0 \text{ as } n \to \infty.$$

$Var(\overline{X}) \to 0$ as $n \to \infty$.

Since $E(\overline{X}) \to \mu$ and $Var(\overline{X}) \to 0$ as $n \to \infty$, thus $\overline{X}$ is a consistent estimator of μ.

Ex :- Let T_1 and T_2 be consistent estimators of γ_1 and γ_2 respectively. Prove that $aT_1 + bT_2$ is a consistent estimator of $a\gamma_1 + b\gamma_2$, where a and b are constants and independent of population.

Sol :- Since T_1 and T_2 are consistent estimators of γ_1 and γ_2.

$\Rightarrow$ $E(T_1) \to \gamma_1$ and $Var(T_1) \to 0$ as $n \to \infty$

and $E(T_2) \to \gamma_2$ and $Var(T_2) \to 0$ as $n \to \infty$.

Now, $E(aT_1 + bT_2) = aE(T_1) + bE(T_2)$

$\to a\gamma_1 + b\gamma_2$ as $n \to \infty$

$\Rightarrow$ $E(aT_1 + bT_2) \to a\gamma_1 + b\gamma_2$ as $n \to \infty$.

and $Var(aT_1 + bT_2) = a^2 Var(T_1) + b^2 Var(T_2)$

$\to 0$ as $n \to \infty$

Hence, $aT_1 + bT_2$ is consistent estimator of $a\gamma_1 + b\gamma_2$.

Ex If $X_1, X_2, \ldots, X_n$ are random observations on a Bernoulli variate X taking the value 1 with probability p and the value 0 with probability $1-p$, show that $\frac{\sum x_i}{n}\left(1-\frac{\sum x_i}{n}\right)$ is a consistent estimator of $p(1-p)$.

x_i	1	0
prob	p	$1-p$

Sol: Given that $X_1, X_2, \ldots, X_n$ are random observations on Bernoulli variate with parameter p.

$E(X_i) = \sum x_i p_i = p$

$V(X_i) = E(X_i^2) - [E(X_i)]^2$
$= p - p^2 = p(1-p)$

$$T = \sum_{i=1}^{n} x_i \sim B(n,p)$$

$\Rightarrow E(T) = np, \quad Var(T) = npq$

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{T}{n} \Rightarrow E(\overline{X}) = \frac{1}{n}E(T) = p \quad \text{as } n \to \infty$$

and $Var(\overline{X}) = Var(T/n) = \frac{1}{n^2}Var(T) = \frac{1}{n^2}(npq)$

$= \frac{pq}{n} \to 0$ as $n \to \infty$.

Since, $E(\overline{X}) \to p$ and $Var(\overline{X}) \to 0$ as $n \to \infty$.

$\Rightarrow \overline{X}$ is a consistent estimator of p.

Also $\frac{\sum x_i}{n}\left(1-\frac{\sum x_i}{n}\right) = \overline{X}(1-\overline{X})$, being a polynomial in $\overline{X}$, is a continuous function of $\overline{X}$.

Since, $\overline{X}$ is consistent estimator of p, by the invariance property of consistent estimators, $\overline{x}(1-\overline{x})$ is a consistent estimator of $p(1-p)$.

Maximum Likelihood Estimation

Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from a population with probability function $f(x, \theta)$, the likelihood function of the sample values $x_1, x_2, \ldots, x_n$ usually denoted by $L = L(\theta)$ is their joint probability function given by

$$L = f(x_1, \theta)\, f(x_2, \theta) \ldots f(x_n, \theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

The principle of maximum likelihood consists in finding an estimator for the unknown parameter θ, which maximises the likelihood function $L(\theta)$.

Definition

Let $x_1, x_2, \ldots, x_n$ be independent observations drawn from a probability distribution that depends on some parameter θ.

The MLE is used to maximize the likelihood function

$$L = f(x_1, x_2, \ldots, x_n \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta)$$

For maximization, $\boxed{\dfrac{dL}{d\theta} = 0; \quad \dfrac{d^2L}{d\theta^2} < 0}$

For Maximizing L, it is equivalently correct to maximize $\log L$.

In other words, the log-likelihood function is easier to work with:

$$\log L = \sum_{i=1}^{n} \log f(x_i|\theta).$$

$$\left[L = \prod_{i=1}^{n} f(x_i|\theta) \quad \log(ab) = \log a + \log b \right.$$

Ex An unfair coin is flipped 100 times, and 61 heads are observed. What is MLE?

Sol: Since, the distribution follow a Binomial distribution with parameter p.

Here $n=100$. The likelihood estimator is

$$P(H=61|p) = {}^{100}C_{61}\, p^{61}(1-p)^{39}$$

For maximization

$$\frac{d}{dp}\left(P(H=61|p)\right) = 0$$

$$\Rightarrow {}^{100}C_{61}\left[-p^{61}\,39(1-p)^{38} + 61\,p^{60}(1-p)^{39}\right] = 0$$

$$\Rightarrow p^{60}(1-p)^{38}\left(61(1-p) - p(39)\right) = 0$$

$$\Rightarrow p^{60}(1-p)^{38}(61 - 100p) = 0$$

$$\Rightarrow p = 0, \frac{61}{100}, 1.$$

Thus, likelihoods are

$$P(H=61|p=0) = 0$$

$$P\left(H=61\,\middle|\,p=\frac{61}{100}\right) = {}^{100}C_{61}\left(\frac{61}{100}\right)^{61}\left(1-\frac{61}{100}\right)^{39}$$

$$= {}^{100}C_{61}\left(\frac{61}{100}\right)^{61}\left(\frac{39}{100}\right)^{39}$$

~~P(H=61~~

$P(H=61 \mid p=1) = 0.$

$\Rightarrow$ Since $P(H=61 \mid p=\frac{61}{100})$ is maximum and hence

$\hat{p} = \frac{61}{100}$ is the MLE.

Ex Consider a Poisson distribution with probability mass function $f(x \mid \mu) = \frac{e^{-\mu}\mu^x}{x!}$, $x = 0, 1, 2, \ldots$

Suppose that a random sample $x_1, x_2, \ldots, x_n$ is taken from the distribution. What is the maximum likelihood estimator of μ?

Sol: The likelihood function is

$$L(x_1, x_2, \ldots, x_n \mid \mu) = \prod_{i=1}^{n} f(x_i \mid \mu)$$

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

Now, consider

$$\log L(x_1, x_2, \ldots, x_n \mid \mu) = -n\mu + \sum_{i=1}^{n} x_i \log\mu - \log\prod_{i=1}^{n} x_i!$$

$$\frac{\partial}{\partial\mu}\log L(x_1, x_2, \ldots, x_n \mid \mu) = -n + \sum_{i=1}^{n}\frac{x_i}{\mu}$$

$$\frac{\partial}{\partial\mu}\log L(x_1, x_2, \ldots, x_n \mid \mu) = 0$$

$$\Rightarrow n = \frac{\sum_{i=1}^{n} x_i}{\mu} \Rightarrow \hat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

$$\frac{\partial^2}{\partial \mu^2} L(x_1, --, x_n | \mu) = -\sum_{i=1}^{n} \frac{x_i}{\mu^2}$$

$\Rightarrow$ $\hat{\mu} = \bar{x}$ is the maximum likelihood estimate of μ.

__Ex__ Suppose 10 rats are used in a biomedical study where they are injected with cancer cells and then given a cancer drug that is designed to increase their survival rate. The survival times, in months, are 14, 17, 27, 18, 12, 8, 22, 13, 19 and 12. Assume that the exponential distribution applies. Give a maximum likelihood estimate of the mean survival time.

__Sol__: The probability density function for the exponential random variable X is

$$f(x;\beta) = \begin{cases} \frac{1}{\beta} e^{-x/\beta}, & x > 0 \\ 0, & \text{elsewhere.} \end{cases}$$

The likelihood function for the data is

$$L(x_1, x_2, \text{---}, x_n | \beta) = \frac{1}{\beta^{10}} e^{-\sum_{i=1}^{n} x_i/\beta}$$

$$\Rightarrow \log L(x_1, x_2, \text{---}, x_n | \beta) = \log \frac{1}{\beta^{10}} + \log e^{-\sum_{i=1}^{n} x_i/\beta}$$

$$= -10 \log \beta - \frac{1}{\beta} \sum_{i=1}^{n} x_i$$

$$\Rightarrow \frac{\partial L}{\partial \beta} = \frac{-10}{\beta} + \frac{1}{\beta^2} \sum_{i=1}^{n} x_i$$

$$\frac{\partial L}{\partial \beta} = 0 \Rightarrow \frac{+10}{\beta} = \frac{1}{\beta^2} \sum_{i=1}^{n} x_i$$

$$\Rightarrow \hat{\beta} = \frac{1}{10} \sum_{i=1}^{n} x_i = \bar{x} = \frac{162}{10} = 16.2$$

$$\frac{\partial^2 L}{\partial \beta^2} = \frac{10}{\beta^2} - \frac{2}{\beta^3} \sum_{i=1}^{n} x_i$$, which will be a negative value for $\beta = 16.2$.

$\therefore$ The MLE of parameter β, the population mean is the sample average $\bar{x} = 16.2$.

Ex In random sampling from normal population $N(\mu, \sigma^2)$, find the maximum likelihood estimator for (i) μ when σ^2 is known

(ii) σ^2 when μ is known

(iii) the simultaneous estimation of μ and σ^2.

Sol. $X \sim N(\mu, \sigma^2)$

$$\therefore n(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

The likelihood function is given by

$$L(x_1, x_2, \_\_, x_n \mid \theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2}$$

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

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

(i) When σ^2 is known, the likelihood equation for estimating μ is

$$\frac{\partial}{\partial \mu}(\log L) = 0$$

$$\Rightarrow \frac{\partial}{\partial \mu}(\log L) = \frac{-1}{2\sigma^2}\sum_{i=1}^{n} 2(x_i-\mu)(-1) = 0$$

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

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

$$\Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}.$$

Hence, M.L.E. for μ is the sample mean $\bar{x}$.

(ii) When μ is known, the likelihood equation for estimating σ^2 is

$$\frac{\partial}{\partial \sigma^2}(\log L) = 0 \Rightarrow -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0$$

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

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

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

Hence, M.L.E. for σ^2 is $\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$.

(1) The likelihood equations for simultaneous estimation of μ and σ^2 are

$$\frac{\partial}{\partial \mu}(\log L)=0 \quad \text{and} \quad \frac{\partial}{\partial \sigma^2}(\log L)=0,$$

which gives $\hat{\mu}=\bar{x}$ and $\hat{\sigma}_x^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2$

$$=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$= s^2$, the sample variance.

$\Rightarrow$ The M.L.E. of μ is the sample mean $\bar{x}$ and M.L.E. of σ^2 is the sample variance s^2.

Ex: Prove that the maximum likelihood estimate of the parameter α of a population having density function $\frac{2}{\alpha^2}(\alpha-x)$, $0<x<\alpha$, for a sample of unit size is $2x$, x being the sample value. Show that the estimate is biased.

Sol: For a random sample of unit size $(n=1)$, the likelihood fun. is

$$L(\alpha)=f(x/\alpha)=\frac{2}{\alpha^2}(\alpha-x);\ 0<x<\alpha.$$

$$\log L=\log\left(\frac{2}{\alpha^2}\right)+\log(\alpha-x)=\log 2-2\log\alpha+\log(\alpha-x)$$

$$\frac{d}{d\alpha}(\log L)=0 \Rightarrow -\frac{2}{\alpha}+\frac{1}{\alpha-x}=0 \Rightarrow -2\alpha+2x+\alpha=0$$

$$\Rightarrow -\alpha+2x=0$$

$$\Rightarrow \alpha=2x$$

$$\frac{d^2}{d\alpha^2}(\log L)=\frac{2}{\alpha^2}-\frac{1}{(\alpha-x)^2},$$ which is negative at $\alpha=2x$.

$\Rightarrow \hat{\alpha}=2x$ is the MLE of α.

T.P. : $E(\hat{\alpha}) \neq \alpha$.

$$E(\hat{\alpha}) = E(2X) = \int_0^{\alpha} (2x)\,\frac{2}{\alpha^2}(\alpha - x)\,dx$$

$$= \frac{4}{\alpha^2}\int_0^{\alpha} (x\alpha - x^2)\,dx$$

$$= \frac{4}{\alpha^2}\left[\frac{x^2}{2}\alpha - \frac{x^3}{3}\right]_0^{\alpha}$$

$$= \frac{4}{\alpha^2}\left[\frac{\alpha^3}{2} - \frac{\alpha^3}{3}\right] = \frac{4}{\alpha^2}\,\frac{\alpha^3}{6} = \frac{4\alpha}{6} = \frac{2\alpha}{3}.$$

$E(\hat{\alpha}) = \frac{2\alpha}{3} \neq \alpha \Rightarrow \hat{\alpha} = 2x$ is a biased estimate of α.

Ex Show that s^2 is a biased estimator of σ^2, where $x_1, x_2, \ldots, x_n$ be a random sample with mean $E(x_i) = \mu$, $Var(x_i) = \sigma^2$.

Sol: T.P. : $E(s^2) \neq \sigma^2$.

Now, $s^2 = E(x_i^2) - [E(x_i)]^2$

$$= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - (\bar{x})^2$$

$$E(s^2) = \frac{1}{n}\sum_{i=1}^{n} E(x_i^2) - E(\bar{x})^2 \quad -(1)$$

Now, $E(x_i^2) = Var(x_i) + [E(x_i)]^2 = \sigma^2 + \mu^2$

and $Var(\bar{x}) = E[(\bar{x})^2] - [E(\bar{x})]^2$

Since $\bar{x} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$

$$\therefore \quad \frac{\sigma^2}{n} = E(\bar{x})^2 - \mu^2 \Rightarrow E(\bar{x})^2 = \frac{\sigma^2}{n} + \mu^2$$

Thus, (1) becomes,

$$E(s^2) = \frac{1}{n}\sum_{i=1}^{n}(\sigma^2+\mu^2) - \frac{\sigma^2}{n} - \mu^2$$

$$= \sigma^2 + \cancel{\mu^2} - \frac{\sigma^2}{n} - \cancel{\mu^2}$$

$$= \sigma^2\left(1-\frac{1}{n}\right)$$

$$\neq \sigma^2$$

$\Rightarrow E(s^2) \neq \sigma^2$

$\Rightarrow s^2$ is a biased estimator of σ^2.

Ex: S^2 is an unbiased estimator of σ^2:

Sol: $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

$= \frac{1}{n-1} \sum_{i=1}^{n} (x_i^2 + \bar{x}^2 - 2x_i \bar{x})$

$= \frac{1}{n-1} \left[\sum_{i=1}^{n} x_i^2 + \bar{x}^2 \sum_{i=1}^{n} 1 - 2\bar{x} \sum_{i=1}^{n} x_i \right]$

$= \frac{1}{n-1} \left[\sum_{i=1}^{n} x_i^2 + n\bar{x}^2 - 2\bar{x}(n\bar{x}) \right]$

$= \frac{1}{n-1} \left[\sum_{i=1}^{n} x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 \right]$

$= \frac{1}{n-1} \left[\sum_{i=1}^{n} x_i^2 - n\bar{x}^2 \right]$

Now, $E(S^2) = \frac{1}{n-1} \left[\sum_{i=1}^{n} E(x_i^2) - n\, E(\bar{x})^2 \right]$

Since, $Var(x_i) = E(x_i^2) - [E(x_i)]^2$

$\sigma^2 = E(x_i^2) - \mu^2 \Rightarrow \boxed{E(x_i^2) = \sigma^2 + \mu^2}$

Also, $\bar{x} \sim N(\mu, \frac{\sigma^2}{n}) \Rightarrow \boxed{E(\bar{x})^2 = \frac{\sigma^2}{n} + \mu^2}$

$\therefore E(S^2) = \frac{1}{n-1} \left[n(\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right) \right]$

$= \frac{1}{n-1} \left[n\sigma^2 - \sigma^2 \right] = \frac{\sigma^2(n-1)}{n-1} = \sigma^2$

$\Rightarrow E(S^2) = \sigma^2 \Rightarrow S^2$ is an unbiased estimator of σ^2.