Methods of Estimation

1. Method of Maximum Likelihood
2. Method of Moments
3. Method of Percentile
4. Method of Minimum χ^2 or Modified Minimum χ^2
5. Method of Least squares

1. Method of Maximum Likelihood

Let X be an random variable with pmf or pdf

$$p(x) \in \{p_\theta(x) : \theta \in \Omega\}, \quad \theta = (\theta_1, \theta_2, \cdots, \theta_k)$$
$$\Omega = \text{an open subset of } \mathbb{R}^k.$$

For $X = x$, the likelihood function (LF) of θ is defined as $L_x(\theta) = p_\theta(x)$, $\theta \in \Omega$.

Here, $p_\theta(x)$ = a function of x for given θ.
while $L_x(\theta)$ = " " " θ " " $X = x$.

For $X = x$, we take the estimate of θ to be that value for which $L_x(\theta)$ is maximum.

Definition: For $X = x$, $\hat{\theta} = \hat{\theta}(x)$ is called the maximum likelihood estimate of θ if $L_x(\hat{\theta}) \geq L_x(\theta) \quad \forall \theta \in \Omega$ ------ (1)

Correspondingly, $\hat{\theta}(X)$ is called the maximum likelihood estimate (MLE).

Clearly, (1) means $L_x(\hat{\theta}) = \sup\limits_{\theta \in \Omega} L_x(\theta)$

This is equivalent to $\ln L_x(\hat{\theta}) = \sup\limits_{\theta \in \Omega} \ln L_x(\theta)$

We assumme,

(i) $\ln L_x(\theta)$ is differentiable with respect to θ

(ii) There is no terminal maximum of the likelihood function. ~~then the~~

Then the MLE of θ satisfies

$$\frac{\partial}{\partial \theta_i} \ln L_x(\theta) = 0 \; ; \; i = 1(1)k \text{ ---- } (2)$$

For $k=1$, (2) reduces to the single equation

$$\frac{d}{d\theta} \ln L_x(\theta) = 0$$

(2) is called the likelihood equation.

Any solution of (2) is called the likelihood equation estimate of θ.

Note: 1. MLE is not same as the likelihood equⁿ. estimate, ~~of θ~~ since the solution of (2) doesn't necessarily correspond to the maximum of likelihood function.

However we can check whether solution of (2), say $\tilde{\theta}$ correspond to the maximum of the likelihood function from the sufficient condition

$$\left(\left(\frac{\partial^2 \ln L_x(\theta)}{\partial\theta_i \partial\theta_j}\right)\right) \text{ is } -\text{'ve definite} \ldots \ldots (3)$$

For $k=1$, (2) reduces to $\frac{\partial^2 \ln L_x(\theta)}{\partial\theta^2} < 0$

A solution of (2) satisfying (3) provides a (relative) maxima of the likelihood function. MLE gives the largest value of likelihood function among the relative maximas.

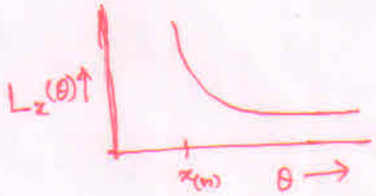
2. If the solution of (2) is unique and also satisfies (3), it will be the unique MLE.

3. If the conditions (i) and (ii) are not satisfied then MLE cannot be found by solving the likelihood equation.

Examples:

1. $x_1, \ldots, x_n$ are iid $\sim R(0,\theta)$

$$L_x(\theta) = \frac{1}{\theta^n} \text{ if } x_{(n)} \le \theta$$
$$= 0 \quad \text{otherwise}$$

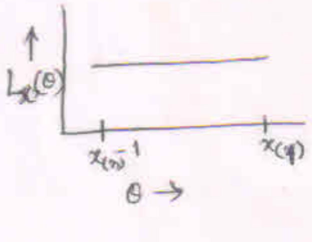


$L_x(\theta)$ is not differentiable at $\theta = x_{(n)}$.

But $\hat{\theta} = x_{(n)}$ is the MLE.

2. $x_1, \ldots, x_n$ are iid $\sim R(\theta, \theta+1)$

$$L_x(\theta) = 1 \text{ if } x_{(1)} \ge \theta,\ x_{(n)} \le \theta+1 \text{ or } x_{(n)}-1 \le \theta \le x_{(1)}$$
$$= 0 \quad \text{otherwise}$$



Any $\theta \in [x_{(n)}-1, x_{(1)}]$ is a MLE of θ.

This is an example where MLE is not unique.

3. X has the pmf $p_\theta(x) = \theta^x (1-\theta)^{1-x}$, $\theta [\frac{1}{4}, \frac{3}{4}]$

Find the MLE of θ. (H.T.)

Properties of MLE (for Small Samples)

1. Invariance Property: Let $g(\theta)$ be 1:1 function of θ.

Then, $\hat{\theta}$ = MLE of θ

$\Rightarrow g(\hat{\theta})$ = MLE of $g(\theta)$.

Proof: Let $\theta^* = g(\theta)$

Ω^* = set of all possible values of θ^*

Since $g(\theta)$ is a 1:1 function of θ

$\exists$ an inverse function $\theta = g^{-1}(\theta^*)$

Likelihood function of $\theta^* = M_x(\theta^*) = L_x(g^{-1}(\theta^*))$

Clearly, $\sup\limits_{\theta\in\Omega} L_x(\theta) = \sup\limits_{\theta^*\in\Omega^*} L_x(g^{-1}(\theta^*)) = \sup\limits_{\theta^*\in\Omega^*} M_x(\theta^*)$

Let, $\hat{\theta}$ = MLE of θ and $g(\hat{\theta}) = \hat{\theta}^*$

Then, $L_x(\hat{\theta}) = L_x(g^{-1}(\hat{\theta}^*)) = M_x(\hat{\theta}^*) \leq \sup\limits_{\theta^*\in\Omega^*} M_x(\theta^*)$

$= \sup\limits_{\theta\in\Omega} L_x(\theta) = L_x(\hat{\theta})$

$\Rightarrow M_x(\hat{\theta}^*) = \sup\limits_{\theta^*\in\Omega^*} M_x(\theta^*)$

i.e. $\hat{\theta}^* = g(\hat{\theta})$ is the MLE of $g(\theta)$.

Note: The result is also true when $g(\theta)$ is not a 1:1 function of θ.

Proof: $\theta^* = g(\theta)$, Ω^* = set of all possible values of θ^*.

Let us define, for every $\theta^* \in \Omega^*$,

$$G(\theta^*) = \{\theta \mid \theta\in\Omega \text{ and } g(\theta) = \theta^*\}$$

[If $g(\theta)$ is a 1:1 function of θ, then

$G(\theta^*) = \{\theta \in g^{-1}(\theta^*)\}$]

Define the likelihood function of θ^* as

$$M_x(\theta^*) = \sup_{\theta\in G(\theta^*)} L_x(\theta)$$

Then $\sup\limits_{\theta\in\Omega} L_x(\theta) = \sup\limits_{\theta^*\in\Omega^*} \sup\limits_{\theta\in G(\theta^*)} L_x(\theta)$

$= \sup\limits_{\theta^*\in\Omega^*} M_x(\theta^*)$

Let, $\hat{\theta}$ = MLE of θ

$g(\hat{\theta}) = \hat{\theta}^*$

$L_x(\hat{\theta}) \leq \sup\limits_{\theta\in G(\theta^*)} L_x(\theta) = M_x(\theta^*) \leq \sup\limits_{\theta^*\in\Omega^*} M_x(\theta^*)$

$= \sup\limits_{\theta\in\Omega} L_x(\theta) = L_x(\hat{\theta})$

$\Rightarrow M_x(\hat{\theta}^*) = \sup\limits_{\theta^*\in\Omega^*} M_x(\theta^*)$

i.e. $\hat{\theta}^*$ is the MLE of $\theta^* = g(\theta)$.

2. MLE and sufficiency

If a sufficient statistic T exists for $\{p_\theta(x) : \theta \in \Omega\}$, the MLE of θ must be a function of T.

<u>Proof</u>: Since $\exists$ a sufficient statistic T for $\{p_\theta(x) : \theta \in \Omega\}$.

We can write,

$$L_x(\theta) = p_\theta(x) = g_\theta(t(x)) \cdot h(x)$$

Maximizing $L_x(\theta)$ w.r.t. θ

$\Leftrightarrow$ " $g_\theta(t(x))$ " θ

$\Rightarrow$ MLE of θ is a function of $t(x)$.

3. MLE and Unbiasedness

MLE is not necessarily unbiased.

Eg: 1. $x_1, \dots, x_n$ are iid $\sim R(0, \theta)$

$x_{(n)}$ is the MLE of θ.

But, $E_\theta(x_{(n)}) = \frac{n}{n+1}\theta$

$\Rightarrow$ MLE is a biased estimator.

2. $x_1, \dots, x_n$ are iid $N(\mu, \sigma^2)$.

MLE of $\mu = \bar{x}$

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

$E(\hat{\sigma}^2) \neq \sigma^2$

$\Rightarrow$ MLE of σ^2 is biased estimator.

MLE for Exponential family

I. Case of a single parameter:-

Let $p_\theta(x) = e^{c(\theta) + Q(\theta)t(x) + h(x)}$; θ is real-valued.

<u>Assumption</u>: (1) The first two derivatives of $Q(\theta)$ and $c(\theta)$ exists and are constant.

(2) $I(\theta) = E_\theta\left[\frac{\partial \ln L}{\partial \theta}\right]^2$ exists and is +ve.

The likelihood equation is

$$\frac{\partial}{\partial \theta} \ln L_x(\theta) = c'(\theta) + Q'(\theta)t(x) = 0$$

$$\text{or, } t(x) = -\frac{c'(\theta)}{Q'(\theta)}$$

<u>Result 1</u>: There exists a solution of the Likelihood Equation iff $t(x)$ and $-\frac{c'(\theta)}{Q'(\theta)}$ have the same range of values.

Proof: <u>Only if part</u>

Let $\exists$ a solution of the Likelihood Equation and let $t(x) \in (\alpha, \beta)$.

Then for any $t_0 \in (\alpha, \beta)$ $\exists$ a $\hat{\theta}$ $\ni$ $t_0 = -\frac{c'(\hat{\theta})}{Q'(\hat{\theta})}$

$\Rightarrow$ any $t_0 \in (\alpha, \beta)$ is a possible value of $-\frac{c'(\theta)}{Q'(\theta)}$ ---- (*)

Now $E_\theta\left[\frac{\partial \ln L}{\partial \theta}\right] = 0$

$\Rightarrow c'(\theta) + Q'(\theta) E_\theta[t(x)] = 0$

$\Rightarrow E_\theta[t(x)] = -\frac{c'(\theta)}{Q'(\theta)}$.

$E_\theta[t(x)] \in (\alpha, \beta)$

$\Rightarrow$ any possible value of $-\frac{c'(\theta)}{Q'(\theta)} \in (\alpha, \beta)$ ---- (**)

(*) and (**) $\Rightarrow$ Range of $-\frac{c'(\theta)}{Q'(\theta)}$ is (α, β).

<u>If part</u>

Let the range of $-\frac{c'(\theta)}{Q'(\theta)}$ = range of $t(x)$.

Then, given any value t_0 of $t(x)$ $\exists$ a value $\hat{\theta}$ $\ni$ $t_0 = -\frac{c'(\theta)}{Q'(\theta)}$.

$\Rightarrow$ The Likelihood function admits a solution

Note: The above necessary and sufficient condition is generally satisfied.

Example: $x_1, x_2, \cdots, x_n$ iid $\sim N(\theta, 1)$

$p_\theta(x) = \text{const.}\, e^{-\frac{1}{2}\sum(x_i-\theta)^2}$

$t(x) = \bar{x}$, $c(\theta) = -\frac{n\theta^2}{2}$, $Q(\theta) = n\theta$ (check)

$\therefore -\frac{c'(\theta)}{Q'(\theta)} = \theta \in (-\infty, \infty)$

$t(x) \in (-\infty, \infty)$

Result 2: (i) Any solution of the likelihood ~~function~~ equation provides a maximum of the likelihood function

(ii) The solution of the likelihood ~~function~~ equation, if it exists, is unique.

(i) & (ii) $\Rightarrow$ The solution of the likelihood is the unique MLE.

Proof: Let $\hat{\theta}$ be a solution of the likelihood equation, Then

$$c'(\hat{\theta}) + Q'(\hat{\theta})\, t(x) = 0$$

$$\Rightarrow -\frac{c'(\hat{\theta})}{Q'(\hat{\theta})} = t(x)$$

(i) It is sufficient to show that

$$\left.\frac{\partial^2 \ln L}{\partial \theta^2}\right|_{\hat{\theta}} < 0.$$

$$\left.\frac{\partial^2 \ln L}{\partial\theta^2}\right|_{\tilde\theta} = c''(\tilde\theta) + Q''(\tilde\theta)\, t(x)$$
$$= c''(\tilde\theta) - Q''(\tilde\theta)\cdot\frac{c'(\tilde\theta)}{Q'(\tilde\theta)} \quad \text{- - - - (a)}$$

$$E_\theta\left[\frac{\partial \ln L}{\partial\theta}\right] = 0.$$
$$\Rightarrow E_\theta[t(x)] = -\frac{c'(\theta)}{Q'(\theta)}$$

$$-I(\theta) = E_\theta\left[\frac{\partial^2 \ln L}{\partial\theta^2}\right]$$
$$= c''(\theta) + Q''(\theta)\, E_\theta[t(x)]$$
$$= c''(\theta) - Q''(\theta)\cdot\frac{c'(\theta)}{Q'(\theta)} \quad \text{- - - - - - (b)}$$

Hence from (a) & (b),

$$-I(\tilde\theta) = \left.\frac{\partial^2 \ln L}{\partial\theta^2}\right|_{\tilde\theta} < 0 \quad \text{Since } I(\theta) > 0 \ \ \forall\,\theta.$$

(ii) If possible, ~~the~~ let there exist more than one solution to the likelihood equation.

Consider any two consecutive solutions, say, $\tilde\theta$ and $\tilde{\tilde\theta}$.
By (i), both $\tilde\theta$ and $\tilde{\tilde\theta}$ provide maximum of the likelihood function.
So, there must exist a solution of the likelihood equation between $\tilde\theta$ and $\tilde{\tilde\theta}$. But these are consecutive solutions of the likelihood equation. Hence solution to the likelihood equation, if it exists, is unique.

Notes 1. The MLE is the unique solution of $t(x) = -\frac{c'(\theta)}{Q'(\theta)}$.
This implies that the MLE and complete sufficient statistic $t(x)$ are in 1:1 relation.
Hence, the MVUE can be obtained from the MLE just by correcting for bias.

2. We get $\left.\frac{\partial^2 \ln L}{\partial\theta^2}\right|_{\tilde\theta} = -I(\tilde\theta)$

$$\Rightarrow I(\theta) = -\left[\frac{\partial^2 \ln L}{\partial\theta^2}\right]_{\tilde\theta=\theta} \quad \text{- - - - - - (*)}$$

(*) can be used to estimate $I(\theta)$ without taking expectation.

<u>Example:</u> $x_1, x_2, \ldots, x_n$ iid $N(0,\theta)$

$$\ln L = \text{constant} - \frac{n}{2}\ln\theta - \frac{\sum x_i^2}{2\theta}$$

$$\therefore \frac{\partial \ln L}{\partial\theta} = 0 \Rightarrow -\frac{n}{2\theta} + \frac{\sum x_i^2}{2\theta^2} = 0$$
$$\Rightarrow \tilde\theta = \frac{1}{n}\sum x_i^2$$

$$\frac{\partial^2 \ln L}{\partial\theta^2} = \frac{n}{2\theta^2} - \frac{\sum x_i^2}{\theta^3} = \frac{n}{\theta^3}\left[\frac{\theta}{2} - \frac{\sum x_i^2}{n}\right] = \frac{n}{\theta^3}\left[\frac{\theta}{2} - \tilde\theta\right]$$

$$\therefore -I(\theta) = \left.\frac{\partial^2 \ln L}{\partial\theta^2}\right|_{\tilde\theta=\theta} = \frac{n}{\theta^3}\left[\frac{\theta}{2} - \theta\right] = -\frac{n}{2\theta^2}$$

$$\therefore I(\theta) = \frac{n}{2\theta^2}$$

3. Let $\exists$ an u.e. $t(x)$ of $g(\theta)$ with variance attaining C-R lower bound. Then, $p_\theta(x)$ is of the exponential form and

$$\frac{\partial \ln L}{\partial \theta} = A(\theta)\,[t(x) - g(\theta)] \qquad \left[\because\ t(x) = g(\theta) + b(\theta)\cdot\frac{\partial \ln L}{\partial \theta}\right]$$

Hence the MLE of $g(\theta)$ is $t(x)$.

II. Case of K parameters

$$p_\theta(x) = e^{c(\theta) + \sum_{j=1}^{k} Q_j(\theta) t_j(x) + h(x)}\ ; \quad \theta = (\theta_1, \theta_2, \cdots, \theta_K)$$

Assumptions: (1) The first two derivatives of $Q_j(\theta)$; $j = 1(1)K$, and $c(\theta)$ exists and are constant.

(2) $J(\theta) = ((J_{ij}(\theta)))$ exists and is (+)ve definite, where,

$$J_{ij}(\theta) = E_\theta\left[\frac{\partial \ln L}{\partial \theta_i}\cdot\frac{\partial \ln L}{\partial \theta_j}\right]$$

The likelihood equation is $\frac{\partial \ln L}{\partial \theta_i} = 0$; $i = 1(1)K$.

i.e. $\frac{\partial c(\theta)}{\partial \theta_i} + \sum_{j=1}^{k} t_j(x)\cdot\frac{\partial}{\partial \theta_i} Q_j(\theta) = 0$; $i = 1(1)K$.

Result 1: (i) Any solution to the likelihood equation provides a maximum of the likelihood function.

(ii) A solution of the likelihood equation, if it exists, is unique.

Proof: Let $\tilde\theta = (\tilde\theta_1, \cdots, \tilde\theta_K)$ be a solution of the likelihood equation.

Then $\frac{\partial c(\theta)}{\partial \theta_i}\Big|_{\tilde\theta} + \sum_{j=1}^{k} t_j(x)\cdot\frac{\partial}{\partial \theta_i} Q_j(\theta)\Big|_{\tilde\theta} = 0$; $i = 1(1)K$ ------ (1)

(i) It is sufficient to show that $\left(\left(\frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j}\right)\right)_{\theta = \tilde\theta}$ is '−'ve definite.

$$\frac{\partial^2 \ln L}{\partial \theta^2}\Big|_{\tilde\theta} = \frac{\partial^2 c(\theta)}{\partial \theta_i \partial \theta_j} + \sum_{r=1}^{k} \frac{\partial^2 Q_r(\theta)}{\partial \theta_i \partial \theta_j}\Big|_{\tilde\theta}\cdot t_r(x) = 0\ ;\ i, j = 1(1)K \ \text{---}\ (2)$$

$E\left[\frac{\partial \ln L}{\partial \theta_i}\right] = 0$, $i = 1(1)K$

$$\Rightarrow \frac{\partial}{\partial \theta_i} c(\theta) \cancel{+ \sum_{r=1}^{k} \frac{\partial^2 Q_r(\theta)}{\partial \theta_i \partial \theta_j}\Big|_{\tilde\theta} t_r(x) = 0} + \sum_{r=1}^{k} \frac{\partial Q_r(\theta)}{\partial \theta_j} E_\theta(t_r(x)) = 0\ ;\ i = 1(1)K \ \text{--}\ (3)$$

$$-I_{ij}(\theta) = E_\theta\left[\frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j}\right] = \frac{\partial^2 c(\theta)}{\partial \theta_i \partial \theta_j} + \sum_{r=1}^{K} \frac{\partial^2 Q_r(\theta)}{\partial \theta_i \partial \theta_j} E_\theta(t_r(x)) = 0\ ;\ i, j = 1(1)K \ \text{---}\ (4)$$

(1) and (2) are same as (3) and (4), the only difference being that $t_r(x)$ is replaced by $E_\theta[t_r(x)]$ and $\tilde\theta$ by θ.

So if we eliminate $E[t_r(x)]$ from (4) using (3) and replace θ by $\tilde\theta$, we get the same result that we would get by eliminating $t_r(x)$ from (2) using (1).

This means $-J_{ij}(\tilde\theta) = \frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j}\Big|_{\tilde\theta}$

By our assumption, $J(\theta) = ((J_{ij}(\theta)))$ is (+)ve definite.

Hence, $\left(\left(\frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j}\right)\right)_{\tilde\theta}$ is '−'ve definite.

(ii) same as Result 2(ii)

Note: 1. Here also we get $-J_{ij}(\tilde{\theta}) = \frac{\partial^2 \ln L}{\partial\theta_i \partial\theta_j}\Big|_{\tilde{\theta}}$

Hence, $-J_{ij}(\theta) = \frac{\partial^2 \ln L}{\partial\theta_i\partial\theta_j}\Big|_{\tilde{\theta}=\theta}$

This means that $J(\theta) = -\left(\left(\frac{\partial^2 \ln L}{\partial\theta_i\partial\theta_j}\Big|_{\hat{\theta}=\theta}\right)\right)$

2. MLE $\tilde{\theta}$ and the complete sufficient statistics $t(x) = (t_1(x), \cdots, t_k(x))$ are in 1:1 relation if $\left(\left(\frac{\partial Q_j(\theta)}{\partial\theta_i}\right)\right)_{j=1(1)k}$ is nonsingular.

For, $Q_j(\theta) = \theta_j$, $\left(\left(\frac{\partial Q_j(\theta)}{\partial\theta_i}\right)\right) = I$.

2. Method of Moments

Suppose we have n r.v.'s $x_1, x_2, \cdots, x_n$ iid $\sim f_\theta(x)$, $\theta = (\theta_1, \theta_2, \cdots, \theta_k)$.

Let us write,

$\mu_r'(\theta) = r$th raw moment of f, $\mu_r' < \infty$, $r = 1, 2, \cdots, k$.

Let $m_r' = \frac{1}{n}\sum_{i=1}^{n} x_i^r = r$th sample raw moment, and let $\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_k$ be the roots of the equations $m_r' = \mu_r'$, $r = 1(1)k$.

Then $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_k)$ is called the moment estimator of θ.

Property:

Let there be a 1:1 correspondence between $(\theta_1, \theta_2, \cdots, \theta_k)$ and $(\mu_1', \mu_2', \cdots, \mu_k')$.

Then

$$\theta_i = f_i(\mu_1', \mu_2', \cdots, \mu_k') ;\ i = 1(1)k.$$

$$\Rightarrow \hat{\theta}_i = f_i(m_1', m_2', \cdots, m_k') ;\ i = 1(1)k.$$

$$m_r' \xrightarrow{P} \mu_r' ;\ r = 1(1)k$$

$\Rightarrow f_i(m_1', m_2', \cdots, m_k') \xrightarrow{P} f_i(\mu_1', \mu_2', \cdots, \mu_k')$ for $i = 1(1)k$, provided f_i's are continuous functions i.e. $\hat{\theta}_i \xrightarrow{P} \theta_i ;\ i = 1(1)k$.

Example: $x_1, x_2, \cdots, x_n$ iid $\sim f_\theta(x)$

$$f_\theta(x) = \frac{1}{\Gamma(\theta)} e^{-x} x^{\theta-1} ;\ x > 0, \theta > 0$$

To estimate θ,

we have $E_\theta(x_i) = \frac{1}{\Gamma(\theta)} \int_0^\infty e^{-x} x^\theta dx = \theta$

$\therefore$ Moment estimator of θ is $\hat{\theta} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

$$E_\theta(x_i^2) = \frac{\Gamma(\theta+2)}{\Gamma(\theta)} = \theta(\theta+1)$$

$$\Rightarrow Var(x_i) = \theta(\theta+1) - \theta^2 = \theta$$

$\therefore$ By CLT, $\sqrt{n}(\bar{x} - \theta) \xrightarrow{L} N(0, \theta)$

$\therefore$ Asymptotic variance of $\bar{x} = \frac{\theta}{n}$

C-R lower bound to the variance of u.e.

$$= \frac{1}{n E\left(-\frac{\partial^2 \ln f}{\partial \theta^2}\right)}$$

$$= \qquad (?)$$

Advantage: This method is easy to apply.

Limitation: The method cannot be applied if population moments do not exists. e.g. Cauchy distn.

3. Method of Percentile.

Let X be a random variable with d.f. $F(x,\theta)$. Let ξ_p be the quantile of order p for this distribution.

Now $F(\xi_p, \theta) = p$

$\Rightarrow \theta = \phi(\xi_p)$, say.

Now $\hat{\theta} = \phi\{x_p\}$, where x_p is the sample quantile of order p.

p is chosen such that $Var(\hat{\theta}) = Var\{\phi(x_p)\}$ is minimum.

Advantage: This method could be used for any distn., even if moments do not exist.

4. Method of Minimum χ^2 or Modified Minimum χ^2.

$x_1, x_2, \ldots, x_n$ are independent observations on $X \sim f_\theta(x)$, θ may be vector-valued. Suppose the observations are grouped into k mutually exclusive and exhaustive classes.

Let f_i = The observed frequency of the ith class, $\sum_{i=1}^{k} f_i = n$.

$\pi_i(\theta)$ = Theoretical probability of the ith class, $\sum_{i=1}^{k} \pi_i(\theta) = 1$.

$n\pi_i(\theta)$ = Expected frequency of the ith class, $i = 1(1)k$.

Then the measure of discrepency between the expected and the observed frequencies is given by

$$\chi^2(\theta) = \sum_{i=1}^{k} \frac{(f_i - n\pi_i(\theta))^2}{n\pi_i(\theta)},$$

That value of θ which minimizes $\chi^2(\theta)$ is called the minimum χ^2- estimator of θ.

Another measure of discrepency is given by

$$\chi^2(\theta) = \sum_{i=1}^{k} \frac{(f_i - n\pi_i(\theta))^2}{f_i} \text{ if } f_i \neq 0 \quad \forall\, i = 1(1)k$$

$$= \Sigma_1 \frac{(f_i - n\pi_i(\theta))^2}{f_i} + 2\Sigma_2\, n\pi_i(\theta) \text{ if } f_i = 0 \text{ for some } i$$

where Σ_1 = Summation over all i's with $f_i \neq 0$

Σ_2 = " " " " " $f_i = 0$

That value of θ which minimizes $\chi^2(\theta)$ is called the modified minimum χ^2 estimator of θ.

Properties: Under certain regularity conditions the above estimators are (i) Consistent, (ii) asymptotically normal and (iii) asymptotically efficient.

Disadvantage: The method is too cumbersome as compared to the other methods of estimation.

5. Method of Least Squares

Let $x_1, x_2, \ldots, x_n$ be n r.v.s with $E(x_i) = \tau_i(\theta)$, $i = 1(1)n$.
Then the value of θ which minimizes

$$S^2 = \sum_{i=1}^{n} (x_i - \tau_i(\theta))^2$$

is called the least squares estimator of θ.

Properties: Generally, LS estimators have no optimal properties even asymptotically. However if $\tau_i(\theta) = a_{i1}\theta_1 + a_{i2}\theta_2 + \cdots + a_{ik}\theta_k$, x_i's are uncorrelated with common variance say σ^2 and a_{ij}'s are known, then the LS estimator of any linear parametric function is MVLUE. If, further, the x_j's be normally distributed, then the estimator is also MVUE.