$X+Y=Z$

By uniqueness theorem we can say $\boxed{X+Y \sim B(n_1+n_2, p)}$

ex $X \sim N(\mu_1, \sigma_1^2)$ & $Y \sim N(\mu_2, \sigma_2^2)$

Both are independent then show that

$\rightarrow$ $X+Y \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

Inequalities :-

1) **Jensen's Inequality** :- Let X be a RV and $f: I \to R$ be a convex function where $I \subseteq R$ in an Interval. Then

$$\boxed{E[f(x)] \geq f[E(x)]}$$

provided expectations exist.

If f is concave then :-

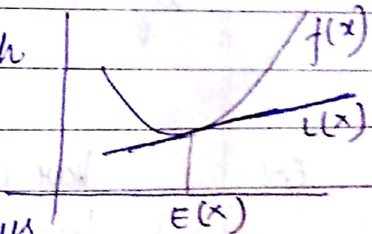
$$\boxed{E[f(x)] \leq f[E(x)]}$$

Proof of this inequality $\rightarrow$

Assumption :- function f is differentiable

Proof : Let $f: I \to R$ be differentiable

Let $E(x) = a+bx$ be a line which is tangent to $f(x)$ at $E(x)$.



Due to differentiability $L(x)$ always lie below the curve

$$\cdot\ E[f(x)] \geq E[L(x)]$$

because this line always lie below the curve

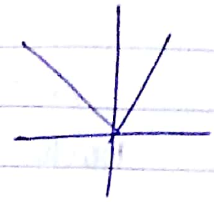
$$E[a+bx] = a+bE(x)$$
$$= L(E(x)) = f(E(x))$$

$$\boxed{E(f(x)) \geq f(E(x))}$$

Scanned by CamScanner

Date: 3/4/2018
Page No.

Ex! $f(x) = |x|$
function is defined on $I \to R$

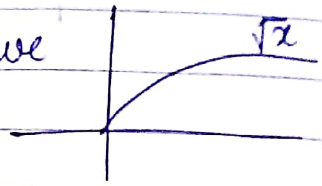


→ By the graph it is clear that graph is con~~cave~~vex.

$$\boxed{E[|X|] \geq |E[X]|}$$ for all Real Values.

Ex: $f(x) = \sqrt{x}$; $x > 0$

concave



$$\boxed{E[\sqrt{x}] \leq \sqrt{E[x]}}$$

Ex! $|E(X)| \leq E[|X|] \leq \sqrt{E(X^2)}$

X^2 is always positive, from previous example we have $E[\sqrt{x}] \leq \sqrt{E[X]}$

→ $E[\sqrt{X^2}] \leq \sqrt{E(X^2)}$ True for all $X \in R$

$X^2 > 0 \quad \forall x \in R$

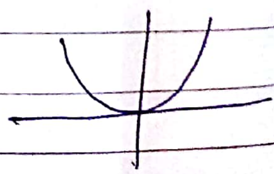
→ $E[|X|] \leq \sqrt{E(X^2)}$

by above example
$$|E(X)| \leq E[|X|] \leq \sqrt{E(X^2)}$$

Ex! $Var(x) \geq 0$
$Var(X) = E(X^2) - [E(X)]^2$

→ consider $f(x) = x^2 \quad \forall x \in R$

convex



→ $E(X^2) \geq (E(X))^2$
→ $E(X^2) - [E(X)]^2 \geq 0$

$$\boxed{Var(x) \geq 0}$$

Moments of a Random Variable :-

Let r be a positive real number and X be a RV. Then r-th moment of X about a point a is defined by : $E[(X-a)^r]$

- If we set $a=0$; then $E[X^r]$ is the r^{th} Moment of RV X about the origin.
- If we set $a = \mu$ (expectation of X) Then

$E[(X - E(X))^r]$ gives us r^{th} moment of X about its mean.

→ Skewness is obtained with the help of moment.

→ variance (spread), skewness, curtosis.

→ we know that $E[X^r]$ exists and is finite if $E[|X^r|]$ exists and is less than ∞

$$\boxed{E[X^r] = \sum_x x^r p_X(x)}$$

$$\sum_x x^r p_X(x) \leq \left| \sum_x x^r p_X(x) \right|$$

$$\leq \sum_x |x^r p_X(x)|$$

PMF always take positive value

$$= \sum_x |x^r| \, |p_X(x)| = E[|X|^r]$$

$$\sum_x x^r p_X(x) \leq \sum_x |x^r| \, |p_X(x)|$$

Date:
Page No.

Ex: If the Moment (about origin) of order $q>0$ exists for a random variable X, then show that Moments of order p exists where $0<p<q$

Sol :- $f:(0,\infty)\to R$ $\quad f(x)=x^r$ where $r>1$

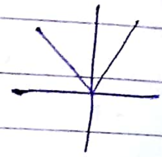
$f''(x)= r(r-1)\,x^{r-2}$

convex $\forall\, x\in(0,\infty)$

$\Rightarrow$ $E[X^r] \geq [E(X)]^r$

$[E(X)]^r \leq E[X^r] \leq E[|X|^r]$

$\Rightarrow$ $[E(X)] \leq (E[|X|^r])^{1/r}$ — ①



$0<p<q$; then $q/p>1$

Replace r by q/p in ① we get

$[E(X)] \leq (E(|X|^{q/p}))^{p/q}$ — ②

set X by X^p we get

$E[X^p] \leq (E[X]^q)^{p/q}$

p^{th} Moment exist if q^{th} Moment exist

If high Moment exist low Moment also exists

Ex: Let X be a RV with $E(X)=10$ show that $E[\ln\sqrt{x}] \leq \frac{1}{2}\ln 10$

Sol $f(x)=\ln\sqrt{x} = \frac{1}{2}\ln x$

$f'(x)=\frac{1}{2x}$ $\quad f''(x)=-\frac{1}{2x^2}<0$ concave

$E[\ln\sqrt{x}] \leq \ln\sqrt{10} = \frac{1}{2}\ln 10$

Scanned by CamScanner

Date: 4/4/18
Page No.

RV → X, we don't know its exact distribution
find out $P(X \geq a)$

- We can get some bound for this probability

→ **Markov's Inequality**: If X is RV with E(x) then

$$\boxed{P(X \geq a) \leq \frac{E(x)}{a}} \quad ; a > 0$$

Proof:

$$P(X \geq a) = \int_a^{\infty} f_X(x)\, dx$$

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx$$

$$= \int_0^{\infty} x f_X(x)\, dx$$

$$= \int_0^a x f_X(x)dx + \int_a^{\infty} x f_X(x)dx$$

$$E(X) \geq \int_a^{\infty} x f_X(x)dx \geq a \int_a^{\infty} f_X(x)dx$$

$$\boxed{P(X \geq a) \leq \frac{E(X)}{a}}$$

Generalize Result:

$$\boxed{P(X \geq a) \leq \frac{E(X^n)}{a^n} \; ; \; a > 0}$$

→ Chebyshev uses 2 constraints (info) so it gives us more clear bounds.

→ **Chebyshev Inequality**:- Let X be a Random variable with mean $\mu = E(X)$ and variance σ^2 Then for every $a > 0$.

$$\boxed{P\left[\,|X-\mu| \geq a\right] \leq \frac{\sigma^2}{a^2}}$$

Date:
Page No.

Proof :

$$P[|X-\mu| \geq a] = P[\sqrt{(X-\mu)^2} \geq a]$$

$$= P[(X-\mu)^2 \geq a^2]$$ Applying Markov's Inequality

$$\leq \frac{E[(X-\mu)^2]}{a^2}$$

$$\leq \frac{\sigma^2}{a^2}$$ Hence Proved

Ex : Weekly production of a factory $E(X) = \mu = 500$
i.e. weekly average production.
$\sigma^2 = 100$ variance in weekly production.

$P(X \geq 1000) = ?$

Applying Markov's Inequality

$$P(X \geq 1000) \leq \frac{E(X)}{a} = \frac{500}{1000} = \frac{1}{2}$$

Chebyshev Inequality :

$$P(X-500 \geq 500) \leq P[|X-500| \geq 500]$$

$$\leq \frac{100}{(500)^2}$$

$$P(X \geq 1000) \leq 0.0004$$

Chebyshev Inequality give better bounds because we are using two information here.

Ex : $X \sim B(n, p)$

estimate $P(X \geq \alpha n)$ where $p < \alpha < 1$

$E(X) = np$

$Var(X) = np(1-p)$

$R_X = 1, 2, 3 \ldots n$

$$P(X \geq \alpha n) = \frac{np}{\alpha n} = \frac{p}{\alpha}$$

By Markov's inequality

6|4|18

$\uparrow E(x)$

$P(X \geq \alpha n) = P(X - np \geq \alpha n - np)$

$\leq P(|X - np| \geq \alpha n - np)$

Now chebyshev inequality →

$P(X \geq \alpha n) \leq \frac{[np(1-p)]}{n^2(\alpha-p)^2} = \frac{p(1-p)}{n(\alpha-p)^2}$

consider $p = \frac{1}{2}$, $\alpha = \frac{3}{4}$

$\boxed{P(X \geq \frac{3}{4}n) \leq 2/3}$ Markov's

$\boxed{P(X \geq \frac{3}{4}n) \leq \frac{4}{n}}$ chebyshev inequality

we have two estimates here

→ If value of n is large ; chebyshev will give us better result. For $n > 6$ chebyshev gives us sharp bounds as compared with Markonikov Inequality.

→ 6|4|18 Law of Large Numbers (LLN) :
Central limit theorem (CLT) :

→ LLN basically describe how the average of randomly selected sample from a large population is likely to be close to the average of the whole population.

→ LLN states that the average of large number of Independent Identically Distributed Random variables converge to the expected value.

→ Independent Identically Distributed RV's (IID)

$S = \{s_1, s_2\}$

$P(s_1) = \frac{1}{3}$ $P(s_2) = \frac{2}{3}$

$X(s_1) = 1$ $\qquad$ $Y(s_1) = 2$

$X(s_2) = 0$ $\qquad$ $Y(s_2) = 1$

$P(X=1) = \frac{1}{3}$ $\qquad$ $P(Y=2) = \frac{1}{3}$

$P(X=0) = \frac{2}{3}$ $\qquad$ $P(Y=1) = \frac{2}{3}$

$X = \begin{cases} 1 \\ 0 \end{cases}$ $\qquad$ $Y = \begin{cases} 2 \\ 1 \end{cases}$

→ But they have same Prob. Mass function, their distribution is same.

→ Weak law of large Numbers :- (WLLN)

Let $X_1, X_2 \ldots$ be a sequence of IID Random variables, each having finite mean μ and variance σ^2 then for each +ve ε (>0)

$$\lim_{n\to\infty} \left[\left| \frac{X_1 + X_2 + X_3 \ldots X_n}{n} - \mu \right| \geq \varepsilon \right] = 0$$

~~and~~ or equivalently

$$\lim_{n\to\infty} P\left[\left| \frac{X_1 + X_2 \ldots X_n}{n} - \mu \right| < \varepsilon \right] = 1$$

\# Average of large number of IID Random Variables converges to expected value.

→ Fair coin → $P(H) = \frac{1}{2} = P(T)$

Tossing of fair coin is a sequence of Random variables which are IID.

→ Proof $E\left[\frac{X_1 + X_2 \ldots X_n}{n}\right] = \underline{\underline{\mu}}$.

for all X = IID RV

Date: ASHOKA
Page No.

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

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

$$= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} =$$

$$\rightarrow P\left[\left|\frac{x_1+x_2+ --- x_n}{n} - \mu\right| \geq \varepsilon\right] = \frac{\sigma^2}{n\epsilon^2}$$

Solved by chebyshev inequality

$$\longrightarrow 0 \quad as \quad n \rightarrow \infty$$

$$S_n = x_1+x_2+ --- x_n$$

$$\lim_{n\rightarrow\infty} P\left[\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right] = 0$$

Ex-1 consider an experiment with event A s.t. $P(A) = p$. Repeat the experiment n times

$$\text{let } X_i = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } \bar{A} \text{ occurs} \end{cases}$$

i.e. each X_i is Bernoullian and IID. Then

$\frac{S_n}{n} = \frac{X_1+X_2 --- X_n}{n}$ is the relative frequency of the occurrence of event A.

By WLLN, we have $\lim_{n\rightarrow\infty} P\left[\left|\frac{S_n}{n} - \mu\right| < \varepsilon\right] = 1$

i.e. probability that Relative frequency of occurrence of the event A is equal to $\mu = E(X_i) = p = P(A)$, is 1 for large n.

95

Date:
Page No.

Ex Let $X_1, X_2 \ldots$ be IID random variable with $E(X_i) = 0$ and $Var(X_i) = 1 \; \forall i$. Let $S_n = X_1 + X_2 \ldots X_n$. Then for any $x > 0$, compute →

$$\lim_{n \to \infty} P(-nx < S_n < nx)$$

→ $P(-nx < S_n < nx) = P\left[\left|\frac{S_n}{n} - 0\right| < x\right]$

$\to 1$ by WLLN

→ $\lim_{n \to \infty} P\left[\left|\frac{S_n}{n} - \mu\right| < \varepsilon\right] = 1$ $\quad S_n = X_1 + X_2 \ldots X_n$

$s \in S$

then $\lim_{n \to \infty} \frac{S_n(x)}{n} = \mu$.

convergence is probability

Strong law of Large Numbers :-

Let $X_1, X_2 \ldots X_n$ be a sequence of IID RV's, each having finite ~~number~~ mean and variance, σ^2. Then for any $\varepsilon > 0$, we have

$$P\left[\lim_{n \to \infty} \frac{X_1 + X_2 \ldots X_n}{n} = \mu\right] = 1$$

9/4/18 Properties of Expectation :-

$X \to$ RV

$E(X) = \sum_{x \in X} x \, p(x)$, X is discrete

$E(X) = \int_{-\infty}^{\infty} x \, p(x)$, X is continous

$a \leq X \leq b \quad ; \quad a \leq E(X) \leq b$

X is a discrete Random variable

$P[a \leq X \leq b] = 1$

$$E(X) = \sum_x x\,p(x) \geq \sum_x a\,p(x)$$
$$\geq a \sum_x p(x)$$
$$\geq a$$

$$E(X) \leq \sum_x b\,p(x) \leq b \sum_x p(x)$$
$$\leq b$$

expectation lies b/w a to b.

Expectation of a function →

suppose X and Y are two RV and g is two variable function

If X and Y are discrete →

$$E[g(X,Y)] = \sum_y \sum_x g(x,y)\,p(x,y)$$

where $p(x,y)$ is joint PMF

If X and Y are continous →

$$E[g(X,Y)] = \int_y \int_x g(x,y)\,f(x,y)\,dx\,dy$$

Q. An accident occurs at a pt. X i.e. uniformly distributed on a road of length L. At the time of the accident an ambulance is at the location Y i.e. also uniformly distributed on the road assuming that X and Y are independent. Find the expected distance b/w the ambulance and the pt. of the accident.

- X → position of accident
- Y → position of ambulance

ASHOKA
Date:
Page No.

$E(|X-Y|) = \iint |X-Y| \, f(x,y) \, dx \, dy$

$X \quad f_X = \frac{1}{L} - \left[\frac{1}{\beta-\alpha}\right], \quad 0 < x < L$

$f_Y = \frac{1}{L} \cdot \quad 0 < x < L$

$f(x,y) = \frac{1}{L} \times \frac{1}{L} = \frac{1}{L^2} \quad \begin{matrix} 0 < x < L \\ 0 < y < L \end{matrix}$

$\int_0^L \int_0^L |X-Y|^2 \times \frac{1}{L^2} \, dx \, dy$

$= \int_{y=0}^{L} |X-Y|^2 \, dy$

$= \int_{y=0}^{x} (x-y) \, dy + \int_{y=x}^{L} (y-x) \, dy$

$\quad y < x \qquad\qquad y > x$

$\Rightarrow \left[xy - \frac{y^2}{2}\right]_0^{y=x} + \left[\frac{y^2}{2} - xy\right]_{y=x}^{y=L}$

$\Rightarrow \left[x^2 - \frac{x^2}{2}\right] - [0] + \left[\frac{L^2}{2} - xL\right] - \left[\frac{x^2}{2} - x^2\right]$

$\Rightarrow \frac{x^2}{2} + \frac{x^2}{2} + \frac{L^2}{2} - xL$

$\Rightarrow \underline{x^2 + \frac{L^2}{2} - xL}$

$\rightarrow \frac{1}{L^2} \int_{x=0}^{L} \left(x^2 + \frac{L^2}{2} - xL\right) dx$

$\rightarrow \frac{1}{L^2} \left[\frac{x^3}{3} + \frac{L^2}{2}x - \frac{x^2}{2}L\right]_0^L$

$\Rightarrow \frac{1}{L^2} \left[\frac{L^3}{3} + \frac{L^3}{2} - \frac{L^3}{2}\right]$

$\Rightarrow \frac{L^3}{3L^2} \Rightarrow \frac{L}{3}$ Ans

Date: ASHOKA
Page No.

→ X, Y ; $E(X)$ and $E(Y)$ are finite

$E(X+Y) = E(X) + E(Y)$

$$E(X+Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (X+Y) f(x,y)\, dy\, dx$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f(x,y)\, dy\, dx + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\, f(x,y)\, dy\, dx$$

$$\Rightarrow \int_{-\infty}^{\infty} x \left(\int_{-\infty}^{\infty} f(x,y)\, dy \right) dx + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f(x,y)\, dx\, dy$$

$$\Rightarrow \int_{-\infty}^{\infty} x\, f_X(x)\, dx + \int_{-\infty}^{\infty} y\, f_Y(y)\, dy$$

$$\Rightarrow E(X) + E(Y)$$

Hence Proved

→ $E(X-Y) = E(X) - E(Y)$

$X = X_1 + X_2 + \text{------} X_n$

$E(X_i) = \text{finite}$

$E(X) = E(X_1) + E(X_2) \text{---} E(X_n)$

→ $X \geq Y$ $\quad E(X) \geq E(Y)$

$X - Y \geq 0$ $\quad E(X-Y) \geq 0$

$E(X) - E(Y) \geq 0$

$E(X) \geq E(Y)$

Boole's Inequality :-

Suppose $A_1, A_2, \text{--} A_n$ are events

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i)$$

ASHOKA

Date:

Page No.

→ Probability that atleast one event occurs; Only events are given; no Relationship given.

Define → Indicator variable for each Ai

$$X_i = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & 0; \text{ otherwise} \end{cases}$$

X → denotes no. of Ai occurs.

$$X = \sum_{i=1}^{n} X_i \rightarrow E(X) = \sum_{i=1}^{n} E(X_i)$$

clearly X denotes the no. of events Ai occur.

$$Y = \begin{cases} 1 & \text{if } X \geq 1, \quad \bigcup_{i=1}^{n} A_i \text{ occurs} \\ 0 & \text{if; otherwise} \end{cases}$$

$$\boxed{X \geq Y}$$

$$E(X) \geq E(Y) \quad — ①$$

$$E(X_i) = 1\, P(A_i) + 0\, [P(\bar{A_i})]$$
$$= P(A_i)$$

$$E(X) = \sum_{i=1}^{n} E(X_i) \quad — ②$$
$$= \sum_{i=1}^{n} P(A_i) \quad — ③$$

$$E(Y) = P\left[\bigcup_{i=1}^{n} A_i\right] \quad — ④$$

Using ③ and ④

$$P\left[\bigcup_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} P(A_i)$$

Central Limit Theorem :-

Let X_1, X_2, X_3 be a sequence of independent and identically uniformly distributed RV each having finite mean (μ) and non zero finite ex Q85, pg 405

Variance σ^2, Define

$S_n = X_1 + X_2 \ldots X_n$

$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{X_1 + X_2 \ldots X_n - n\mu}{\sigma\sqrt{n}}$

Then

$$\lim_{n\to\infty} [Z_n \le n] = N(n) = \phi(x)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2}\, dt$$

↓ standard normal distribution.

→ For this Result $X_1, X_2 \to$ Independent and $\sigma, \mu \to$ finited

Ques no. of enrollments $\geq 120 \to$ two sections

$\leq 120 \to$ one section

P(Professor will teach two sections) = $\sum_{i=120}^{\infty} e^{-100} \frac{(100)^i}{i}$

→ Poisson RV with mean 100 can be written as sum of 100 Poisson RV with Mean 1.

$X \to$ denotes the no. of the student enrolled in 1 ss.

$P[X \geq 120]$

$P[X \geq 119.5]$ by continuity correction; when the discrete ~~di~~ distribution approximated by continous distribution. Add/sub = 0.5.

$$P[X \geq 119.5] = P\left[\frac{X - 100(1)}{1\sqrt{100}} \geq \frac{119.5 - 100(1)}{1\sqrt{100}}\right]$$

$$= P\left[\frac{X - 100}{10} \geq 1.95\right]$$

$$1 - P\left[\frac{X-100}{10} \le 1.95\right]$$

$$1 - \phi(11.9.5) \qquad = 0.256$$

10/4/18 Statistical Inference :- It is the collection of all Methods that deal with drawing conclusions from the data that are sensitive to Random Variation.

Ex1 Exit poll Problem

Random sample, Randomness arises due to sampling.

Timing of exit poll is also a Random Variation.

Ex2 communication system → Suppose Transmitter transmits signal X and Receiver Receives signal Y. Signal is transmitted in terms of Bits. Y is the noisy version of X. From Y we want Idea about X.

Randomness is arrising due to Noise.

Radom sample / Data = Y.

General setup of Statistical Inference Problem

There is an unknown quantity that we would like to estimate.

We get some data (Data will be Radom Sampling)

From the data we estimate the desired quantity. We need to find some estimate about unknown quantity.

→ There are two major Approaches of statistical Inference problem.

1) classical → It deals with such unknown quantities that are deterministic in nature.

Date:
Page No.

Does not have probabilistic nature

1) **Classical Inference** → In this approach the unknown quantity is assumed to be a fixed quantity, i.e. deterministic (non-Random) quantity. that is to be estimated by the observed data.

θ: percentage of people who will cast their votes to a particular candidate say 'A'. After asking n randomly chosen voters, we might estimate θ by

$$\hat{\theta} = \frac{Y}{n}$$

where Y is the number of people (among n randomly chosen voters) who said they will vote / give their vote to 'A'

→ $\hat{\theta}$ is not deterministic / fixed → It is a Random variable, $\hat{\theta}$ depends when we are conducting exit poll, how many days before elections.

2) **Bayesian Inference** → In the Bayesian approach the unknown quantity (θ) is assumed to be a random variable and we assume that we have some Initial guess about the distribution of θ. After observing the data, we update the distribution of θ using Baye's Rule.

eg suppose a Transmitter transmitted '1' with prob 'p' and transmitted '0' with prob $(1-p)$ 'θ' is random in nature and we have some idea about the distribution of 'θ'

$\boxed{\theta \sim \text{Bernoulli}(p)}$

Random Sampling :-

Consider our aim is to get the distribution of height of people in a well defined population. (eg population between 25 and 55 in a certain country).

To do this, we define random variables $X_1, X_2 \dots X_n$ as follows:-

1) We choose a random sample of size n with replacement from the population and let X_i is the length of i^{th} chosen person.

Sampling with Replacement ?

→ when n is very large Prob is very low
→ sampling without Replacement

choosing 1st object $P = 1/n$
choosing 2nd object $P = \frac{1}{n-1}$ so choosing 2nd object is affected by 1st object.

→ When we do sampling with Replacement then Prob. of 2 things not dependent.

When $n \to$ large sampling without Replacement because prob is very low.

\# The Big advantage of sampling with replacement is that X_i's will be independent and this make analysis much simpler.

→ Now we would like to estimate the average height in the population.

Date:
Page No.

$$\hat{\theta} = \frac{X_1 + X_2 - - - X_n}{n}$$

X_i's are Independent and Identically distributed

→ The collection of RV's $X_1, X_2 - X_n$ is said to be a Random sample of size n if they are Independent and Identically distributed.
All RV's are having same distribution function

Camlin Page
Date 11 / 4 / 18

- The collection of R.V's $X_1, X_2 \ldots X_n$ is said to Random ~~var~~ sample of size n if they are Independently and Identically distributed (IID)

1. θ Average height of the population.

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

↓

point estimator of unknown quantity θ

In general a point estimator is a function of the random sample i.e.

$$\hat{\theta} = h(X_1, X_2 \cdots X_n)$$

point estimator is used to estimate the unknown quantity θ.

Properties of Random sample :-

Let $X_1, X_2 \ldots X_n$ are random sample Then

1) X_i's are Independent
2) $F_{X_1}(x) = F_{X_2}(x) = \cdots F_{X_n}(x)$
3) $E(X_i) = E(X) = \mu < \infty$
4) $Var(X_i) = Var(X) = \sigma^2 < \infty$

Sample Mean :

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

$\hat{\theta}$ and sample mean are different.

Properties of sample Mean :

1) $E(\bar{X}) = \mu$

2) $Var(\bar{X}) = \frac{\sigma^2}{n}$

3) WLLN

$$\lim_{n \to \infty} P(|\bar{X} - \mu| \geq \epsilon) = 0$$

4) CLT = $Z_n = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + X_2 \ldots X_n - n\mu}{\sigma\sqrt{n}}$

$\lim_{n \to \infty} P(Z_n \leq x) = \phi(x)$ where $\phi(x)$ is the standard normal RV.

Order statistics : Let $X_1, X_2, \ldots X_n$ be random sample from a continous distribution of f^n $F_X(x)$. Let us order X_i's from samples s.t. from the smallest to the largest and denote the resulting sequence of RV by :-

$X_{(1)}, X_{(2)} \ldots X_{(n)}$

This sequence is said to be order statistics.

Theorem : Let $X_1, X_2 \ldots X_n$ be a random sample from a continous distribution with CDF $F_X(x)$ and pdf $f_X(x)$. Let $X_{(1)}, X_{(2)} \ldots X_{(n)}$ be the ordered statistics of $X_1, X_2 \ldots X_n$. Then pdf, cdf is given by :

$$\rightarrow \quad f_{X_{(i)}}(x) = \frac{n!}{(i-1)!\,(n-i)!} \cdot f_X(x) \left[F_X(x)\right]^{i-1} \cdot \left[1 - F_X(x)\right]^{n-i}$$

$$\rightarrow \quad F_{X(i)}(x) = \sum_{k=i}^{n} \left({}^{n}C_{k}\right) \left[F_X(x)\right]^{k} \cdot \left[1 - F_X(x)\right]^{n-k}$$

Joint PDF of $X_{(1)}, X_{(2)} \ldots X_{(n)}$ is given by $\rightarrow$

$$f_{X_{(1)} X_{(2)} \ldots X_{(n)}}(x) = \begin{cases} n!\, f_X(x_1)\, f_X(x_2) \ldots f_X(x_n) & x_1 \le x_2 \le \ldots x_n \\ 0 & \text{otherwise} \end{cases}$$

$X_{(1)} = \min(X_1, X_2 \ldots X_n)$

$X_{(n)} = \max(X_1, X_2 \ldots X_n)$

$S = \{s_1, s_2, s_3\}$

X_1 : $X_1(s_1) = 1$ $\quad X_1(s_2) = 2$

$X_1(s_3) = 3$

X_2 : $X_2(s_1) = 3$, $\quad X_2(s_2) = 1$

$X_2(s_3) = 2$

X_3 : $X_3(s_1) = 2$ $\quad X_3(s_2) = 3$

$X_3(s_3) = 1$

$X_{(1)}(s_1) = X_1(s_1) = 1 = 1$

$X_{(1)}(s_2) = X_2(s_2) = 1$

$X_{(1)}(s_3) = X_3(s_3) = 1$

Ex: Let X_1, X_2, X_3, X_4 be uniformly distributed over $(0,1)$

$X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}$ is the order statistics of X_1, X_2, X_3, X_4. Find pdf of $X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}$

Sol

$$f_X(x) = 1 \quad \forall\, x \in [0,1]$$

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}$$

$$\rightarrow f_{X_{(1)}}(x) = \frac{4!}{0!\,(4-1)!} \cdot 1 \cdot x^{1-1} (1-x)^{4-1}$$

$$\rightarrow 4(1-x)^3$$

13/4/18

θ: Unknown parameter which needs to be estimated.

It is fixed

→ We need some data

$X_1, X_2 \ldots X_n$ random sample

$$\hat{\theta} = h(X_1, X_2 \ldots X_n)$$

↓

point estimator

θ: Average height of well defined population $X_1, X_2 \ldots X_n$.

$$\hat{\theta}_1 = \frac{X_1 + X_2 \ldots X_n}{n}$$

$\hat{\theta}_2 = X_1$ (only one point from sample space)

$E(\hat{\theta}_1) = E(X)$

Because all x_i's are Independent and distributed in their own way (Identically)

$E(\hat{\theta}_2) = E(X)$

→ How we can ensure that a point estimator is a good estimator for θ?

→ Three desired properties should satisfy by a point estimator which are :-

(1) <u>Bias</u> :- The bias of a point estimator $\hat{\theta}$ is denoted by $B(\hat{\theta})$ and defined by

$$\boxed{B(\hat{\theta}) = E(\hat{\theta}) - \theta}$$

fixed

<u>$E(\theta) = \theta$</u> on average how $E(\hat{\theta})$ is far away from $E(\theta)$?

<u>Ex</u>: $B(\hat{\theta}_1) = E(\hat{\theta}) - \theta$

$$B(\hat{\theta}_1) = E(\hat{\theta}_1) - \theta$$
$$= E(X) - \theta$$
$$= \theta - \theta = 0$$

θ: Av. Height $= E(X)$

$$B(\hat{\theta}_2) = E(\hat{\theta}_2) - \theta$$
$$= \theta - \theta = 0$$

Those estimator for which $B(\hat{\theta}) = 0$ are called unbiased estimator.

(Q.) Mean square Error :- For a point estimator $\hat{\theta}$ is denoted by $MSE(\hat{\theta})$ and defined by →

$$\boxed{MSE(\hat{\theta}) = E\left[(\hat{\theta}-\theta)^2\right]}$$

Those estimators are good for which mean square error is less.

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

$$E(X^2) = [E(X)]^2 + Var(X)$$

$$MSE(\hat{\theta}) = [E(\hat{\theta}-\theta)]^2 + Var(\hat{\theta}-\theta)$$

$$= Var(\hat{\theta}) + [E(\hat{\theta})-\theta]^2$$

$$= Var(\hat{\theta}) + [B(\hat{\theta})]^2$$

As θ is fixed.

ex MSE $(\hat{\theta_1})$ and MSE $(\hat{\theta_2})$ and show that $MSE(\theta_1) > MSE(\theta_2)$

$\theta_1 = X_1$

$\theta_2 = \frac{X_1 + X_2 \cdots X_n}{n}$

Sol : $MSE(\hat{\theta_2}) = E\left[\left(\frac{X_1 + X_2 \cdots X_n}{n} - \right.\right.$

- $Exp(X_1) = Exp(X_2) = \cdots\cdots = E(X)$
- $Var(X_1) = Var(X_2) = \cdots\cdots Var(X_n) = Var(X)$

$$MSE(\hat{\theta_1}) = E\left(\left[X_1 - \theta\right]^2\right)$$

$$\Rightarrow E\left(\left[X_1 - E(X)\right]^2\right)$$

$$\Rightarrow E\left[X_1^2\right] \quad E(X) = E(X_1)$$

$$\Rightarrow E\left[\left(X_1 - E(X_1)\right)^2\right]$$

$$\Rightarrow Var(X_1) = Var(X) = \sigma^2$$

$$MSE(\hat{\theta}_2) = E\left([\hat{\theta}_2 - \theta]^2\right)$$

$$\Rightarrow E\left[\frac{x_1 + x_2 \cdots x_n}{n} - E(x_2)\right]^2$$

$$\Rightarrow \frac{1}{n^2} E\left[x_2 - E(x_2)\right]^2$$

$$\Rightarrow \frac{Var(X_2)}{n^2} = \frac{\sigma^2}{n^2}$$

$$\boxed{\sigma^2 \geq \frac{\sigma^2}{n}}$$

or $\quad Var\left(\frac{x_1 + x_2 \cdots x_n}{n}\right) + \left[B(\hat{\theta}_2)\right]^2$

$$\Rightarrow \frac{\sigma^2}{n^2} + 0 =$$

3) <u>consistency of Estimator :-</u>

$\hat{\theta}_n \qquad \theta$

Let : $\hat{\theta}_1, \hat{\theta}_2 \cdots \hat{\theta}_n$ be a sequence of point estimator for θ then $\hat{\theta}_n$ is said to be consistent estimator of θ if $\rightarrow$

$$\lim_{n \to \infty} P\left[|\hat{\theta} - \theta| \geq \varepsilon\right] = 0 \quad \forall\ \varepsilon > 0$$

Ex : If $\hat{\theta}_n = \bar{X} = \dfrac{x_1 + x_2 \cdots X_n}{n}$ then show that it is a consistent estimator of θ

$\theta = E(X) = E(X_i)$.

1) If we compare it with weak law of large numbers

$$P\left[\left| \frac{x_1+x_2 \cdots x_n}{n} - E(x) \right| \geq \epsilon \right]$$

by chebyshev Inequality

$$P\left[\left| \underbrace{\frac{x_1+x_2 \cdots x_n}{n}}_{\bar{X}} - E(x) \right| \geq \epsilon \right] \leq \frac{Var(\bar{X})}{\epsilon^2}$$ by chebyshev

$$Var(\bar{X}) = \sigma^2$$

$$\leq \frac{\sigma^2}{n\epsilon^2} \to 0 \text{ as } n \to \infty$$

<u>Theorem:</u> Let $\hat{\theta}_1, \hat{\theta}_2, \ldots \hat{\theta}_n$ be a sequence of point estimator for θ. If

$$MSE(\hat{\theta}_n) \longrightarrow 0 \text{ as } n \to \infty$$

Then $\hat{\theta}_n$ is a consistent estimator of θ.

<u>Proof:</u> $P[|\hat{\theta}_n - \theta| \geq \epsilon] = P[(\hat{\theta}_n - \theta)^2 \geq \epsilon^2]$

$$\leq \frac{E[\hat{\theta}_n - \theta]^2}{\epsilon^2}$$ By Markovnikov Inequality

$$= \frac{MSE(\hat{\theta}_n)}{\epsilon^2} \to 0 \text{ as } n \to \infty$$

16/4/18

$X_1, X_2 \ldots X_n$

$E(X) = \mu$

$\bar{X} = \frac{X_1 + X_2 \cdots X_n}{n}$ is a good estimator of μ.

$$\sigma^2 = E\Big[\underbrace{(x-\mu)^2}_{Y}\Big]$$

Good point estimator for σ^2?

$x-\mu$ is a RV $\rightarrow$ $(x-\mu)^2$ = R.V.

$x-\mu = Y$ (also a random variable)

$\bar{Y} = \dfrac{Y_1 + Y_2 + \cdots Y_n}{n}$ is a good point estimator for $E(Y)$ i.e. for σ^2, where $Y_k = (X_k-\mu)^2$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{k=1}^{n} Y_k = \frac{1}{n}\sum_{k=1}^{n}(X_k-\mu)^2$$

$$\boxed{\hat{\sigma}^2 = \frac{1}{n}\sum_{k=1}^{n}(X_k-\mu)^2}$$

$\hat{\sigma}^2$ unbiased estimator with the help of WLLN, one can see $\hat{\sigma}^2$ is consistent.

In practice, we do not know the value of $\mu = E(x)$.

That's why we replace μ by $\bar{X}$

$$\boxed{\begin{aligned}\hat{S}^2 &= \frac{1}{n}\sum_{k=1}^{n}(X_k-\bar{X})^2 \\ &= \frac{1}{n}\left[\sum_{k=1}^{n} x_k^2 - n\bar{X}^2\right]\end{aligned}}\ ?$$

Camlin Page
Date / /

Theorem :- Let $X_1, X_2, \ldots X_n$ be a Random sample with Mean $E(X_i) = \mu$ & variance $Var(X_i) = \sigma^2$ suppose that we use

$$\bar{S}^2 = \frac{1}{n}\sum_{k=1}^{n}(X_k - \bar{X})^2 = \frac{1}{n}\sum_{k=1}^{n}(X_k^2 - n\bar{X}^2)$$

to estimate σ^2. Find the bias of $\bar{S}^2$.

Proof :- $B(\bar{S}^2) = E(\bar{S}^2) - \sigma^2$

$$E(\bar{S}^2) = \frac{1}{n}\sum_{k=1}^{n}\left[E(X_k^2) - nE(\bar{X}^2)\right] \quad -(1)$$

$$E(X_k^2) = Var(X_k) + [E(X_k)]^2$$

$\left[\begin{array}{l} Var(Y) = E(Y^2) - [E(Y)]^2 \\ E(Y^2) = Var(Y) + [E(Y)]^2 \end{array} \right]$

$$E(X_k^2) = \sigma^2 + \mu^2$$

$$E(\bar{X}^2) = Var(\bar{X}) + [E(\bar{X})]^2$$
$$= \frac{\sigma^2}{n} + \mu^2$$

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

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

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

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

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

For large n, $\bar{s}^2$ is an unbiased estimator of σ^2, But in general it is a biased estimator.

→ On multiplying $\bar{s}^2$ with $\left(\frac{n}{n-1}\right)$ we get

$$E(\bar{s}^2) = \sigma^2$$

and Then $\boxed{B(\bar{s}^2) = 0}$

$$E(\bar{s}^2) = \frac{1}{(n-1)}\left[\sum_{k=1}^{n} \left\{E(x_k^2) - nE(\bar{x}^2)\right]\right.$$

$$\boxed{E(s^2) = \sigma^2}$$

$$s^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \bar{x})^2 = \frac{1}{n-1}\left[\sum_{k=1}^{n}\left\{x_k^2 - n\bar{x}^2\right]\right.$$

⇒ By this s^2 is an unbiased point estimator for σ^2, s^2 is known as sample variance of σ^2.

$s = \sqrt{s^2}$ is a point estimator for S.D σ and is known as sample standard deviation.

→ S is a Random variable

$$Var(S) > 0$$

$$0 < Var(S) = E(S^2) - [E(S)]^2$$
$$= \sigma^2 - [E(S)]^2$$

i.e. $\sigma^2 > [E(S)]^2$ $= \sigma^2 > E(S)$

$B(S) = E(S) - \sigma < 0$ → S is biased estimator for σ.

Ques Let T be the time that is needed for a specific task in a factory to be completed In order to estimate the mean and variance of T, we observe a Random sample $T_1, T_2, T_3, T_4, T_5, T_6$

18, 21, 17, 16, 24, 20

Based on this find the values of sample Mean, sample variance and the sample standard deviation for this sample.

Sol Sample Mean $= \frac{18+21+17+16+24+20}{6}$

$= \frac{116}{6}$

$E(T) = 19.33$

estimator of expectation of T.

$S^2 = \frac{1}{6-1} \sum_{R=1}^{6} (T_R - 19.33)^2$

$= 8.67$

↓

sample variance

sample SD = 2.94

Camlin Page
Date 17/4/18

Prob X: height of Randomly selected person from a population. Estimate $E(X)$, $Var(X)$?

Random sample	x_1	x_2	x_3	x_4	x_5	x_6	x_7
	166.8	171.4	169.1	178.5	168.0	157.9	170.1

sample mean → $$\bar{X} = \frac{x_1 + x_2 \cdots x_7}{7}$$

$$= 168.8$$

$$S^2 = 37.7 \quad \text{(sample Variance)}$$

Ex2: Prove the following :-

(a) If $\hat{\theta}_1$ is an unbiased estimator for θ and W is a zero mean R.V. then
$$\hat{\theta}_2 = \hat{\theta}_1 + W$$
is also an unbiased estimator of θ.

Sol :-
$$E(\hat{\theta}_2) = E(\hat{\theta}_1) + E(W)$$
$$= E(\hat{\theta}_1) + 0$$

$E(\hat{\theta}_1) = \theta$ (Because $\hat{\theta}_1$ is an unbiased point estimator)

$$E(\hat{\theta}_2) = \theta$$

so $\hat{\theta}_2$ also an unbiased estimator.

(b) If $\hat{\theta}_1$ is an estimator for θ s.t. $E(\hat{\theta}_1) = a\theta + b$, $a \neq 0$ then

$\hat{\theta}_2 = \dfrac{\hat{\theta}_1 - b}{a}$ is an unbiased estimator.

with the help of one Bco unknown estimator 2nd estimator was proved unbiased

Sol we don't have any idea about the bias of $\hat{\theta}_1$ ~~~~ whether it is Biased or not!

$$E(\hat{\theta}_2) = \frac{1}{a}\left[E(\hat{\theta}) - E(b)\right]$$

$$\Rightarrow \frac{1}{a}\left[a\theta + \cancel{b} - \cancel{b}\right]$$

$$\boxed{E(\hat{\theta}_2) \Rightarrow \theta}$$ so it is unbiased.

Prob: Let X_1, X_2 be a random sample from a uniform $(0,\theta)$ distribution when θ is unknown. Define the estimator →

$$\hat{\theta}_n = \text{Max}\left[X_1, X_2 \ldots X_n\right]$$

(a) Find $B(\hat{\theta}_n)$

(b) Find $MSE(\hat{\theta}_n)$

(c) Is $\hat{\theta}_n$ consistent ?

$X_1, X_2 \ldots$ are Independent and Identically distributed $X_{(1)}, X_{(2)} \ldots X_{(n)}$

$$X_{(1)} = \text{Max}\left[x_1, x_2 \ldots x_n\right]$$

$$\vdots$$

$$X_{(n)} = \text{Max}\left[X_1, X_2 \ldots x_n\right]$$

$\hat{\theta}_n$ is nothing but $X_{(n)}$

- $$f_{X_{(i)}}(x) = \frac{n!}{(i-1)!(n-i)!} f_X(x)\left[F_X(x)\right]^{i-1}\left[1 - F_X(x)\right]^{n-i}$$

$$f_{\hat{\theta}_n}(x) = f_{X_{(n)}}(x) = \frac{n!}{(n-1)!} f_X(x) \left[F_X(x)\right]^{n-1}$$

$$\boxed{f_{X_{(n)}}(x) = n f_X(x) \left[F_X(x)\right]^{n-1}}$$

$$f_X(x) = \begin{cases} 1/\theta & 0 < x < \theta \\ 0 & \text{otherwise} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & x \leq 0 \\ x/\theta & 0 < x < \theta \\ 1 & x \geq \theta \end{cases}$$

$$\Rightarrow f_{\hat{\theta}_n}(x) = \begin{cases} n \times \frac{1}{\theta} \left(\frac{x}{\theta}\right)^{n-1} & 0 < x < \theta \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{n x^{n-1}}{\theta^n}$$

(a) $B(\hat{\theta}_n) = E(\hat{\theta}_n) - \theta$

$$E(\hat{\theta}_n) = \int_0^\theta x \cdot \frac{n x^{n-1}}{\theta^n} \, dx$$

$$\Rightarrow \frac{n}{\theta^n} \left[\frac{x^{n+1}}{n+1}\right]_0^\theta$$

$$\Rightarrow \frac{n}{n+1} \frac{\theta^{n+1}}{\theta^n} = \frac{n\theta}{n+1}$$

$$\rightarrow B(\hat{\theta}_n) = \frac{n\theta}{n+1} - \theta = \frac{-\theta}{n+1}$$

Bias of this estimator when sample large, Bias $\rightarrow 0$ so ~~sample~~ estimator is unbias estimator

(b) Mean square error $\Rightarrow \operatorname{Var}(\hat{\theta}_n) + [B(\hat{\theta}_n)]^2$

$$\operatorname{Var}(\hat{\theta}_n) = E(\hat{\theta}_n^2) - (E[\hat{\theta}_n])^2$$

$$E(\hat{\theta}_n) = \int_0^\theta x^2 \frac{n x^{n-1}}{\theta^n}\, dx$$

$$\Rightarrow \frac{n}{n+2}\theta^2$$

$$\Rightarrow \operatorname{Var}(\hat{\theta}_n) = \frac{n}{n+2}\theta^2 - \frac{n^2}{(n+1)^2}\theta^2$$

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

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

$$\Rightarrow \frac{n\theta^2}{(n+1)^2(n+2)}$$

$$\text{MSE}(\hat{\theta}_n) = \left(\frac{n}{n+2}\right)\left(\frac{\theta}{n+1}\right)^2 + \frac{\theta^2}{(n+1)^2}$$

$$\Rightarrow \frac{\theta^2}{(n+1)^2}\left[\frac{n}{n+2} + 1\right]$$

$$\Rightarrow \frac{2\theta^2}{(n+1)(n+2)}$$

(c) consistency ?

$\text{MSE}(\hat{\theta}_n) \to 0$ as $n \to \infty$

$\hat{\theta}_n$ is consistent estimator of θ.

Likelihood Function :- Let $X_1, X_2 \dots X_n$ be a random sample from a population with pdf or pmf $f(x|\theta)$. The likelihood function is defined by :-

$$L(\theta | x_1, x_2 \dots x_n) = f(x_1|\theta)\, f(x_2|\theta) \dots f(x_n|\theta)$$

where $x_1, x_2 \dots x_n$ are the realized value of Random sample $X_1, X_2 \dots X_n$.

Domain of $L(\theta, x_1, x_2 \dots x_n)$ is the set of all admissible value of the parameter θ.

Ques Let X_1, X_2, X_3, X_4 be a random sample from a Bernoulli population with parameter p. For the sample value $(1,0,1,1)$. Find likelihood function.

$$L(p | (1,0,1,1)) = p \cdot (1-p) \cdot p \cdot p$$
$$= p^3 (1-p) \quad p \in [0,1]$$

$$p_X(x) = \begin{cases} p & x = 1 \\ 1-p & x = 0 \end{cases}$$

$(1,0,0,1)$

$$L(p | (1,0,0,1)) = p^2 (1-p)^2$$

Ques Let X_1, X_2, X_3, X_4 be a random sample from exponential (λ) population. For the sample value $(1.23, 3.32, 1.98, 2.12)$. Find the likelihood estimator →

Sol $L(\lambda \mid (1.23, 3.32, 1.98, 2.12))$

$$\Rightarrow \lambda e^{-1.23\lambda} \cdot \lambda e^{-\lambda 3.32} \cdot \lambda e^{-\lambda 1.98} \cdot \lambda e^{-\lambda 2.12}$$

$$\Rightarrow \lambda^4 e^{-86.5\lambda} \qquad \lambda \in [0, \infty)$$

Maximum likelihood estimator :-

For each sample point $x = (x_1, x_2 \dots x_n)$ let $\hat{\theta}(x)$ be the parameter value at which $L(\theta \mid x)$ attains its Max. as a f^n of θ with x held fixed. A Maximum likelihood estimator (MLE) of the parameter θ based on random sample $X_1, X_2 \dots X_n$ is $\hat{\theta}(X)$.

- If $L(\theta \mid x)$ is differentiable then

$$\frac{dL}{d\theta} = 0$$

Solve it for θ. find the pt. of global Maxima

$$L(p \mid (1,0,1,1)) = p^3(1-p) \qquad p \in [0,1]$$

Ques For which value of p is the probability of the observed sample value $(1,0,1,1)$ is largest ?

$$\frac{dL}{dp} = 3p^2(1-p) - p^3$$
$$\Rightarrow 3p^2 - 4p^3$$

$$\frac{dL}{dp} = 0$$
$p = 0, \frac{3}{4} \rightarrow$ critical point

$\frac{3}{4}$ is the only critical point

$L\left(\frac{3}{4}\right) > 0 \Rightarrow$ pt. of global Maxima

$\Rightarrow p = \frac{3}{4}$ is MLE.

Ques Let X_1, X_2, X_3, X_4 be a Random sample from exp (λ) distribution. Find the MLE for the sample value (1.23, 3.32, 1.98, 2.12)

$$L(\lambda \mid 1.23, 3.32, 1.98, 2.12) = \lambda^4 e^{-8.65\lambda}$$
$$\lambda \in (0, \infty)$$

$$F(\lambda) = \log L = 4 \log \lambda - 8.65\lambda$$

$$\frac{dF}{d\lambda} = 0 \Rightarrow \frac{4}{\lambda} - 8.65 = 0$$

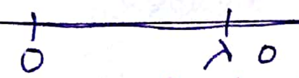
$$\boxed{\lambda = \frac{4}{8.65}}$$

$$\lambda = 0.4624$$

$$\rightarrow \frac{d^2F}{d\lambda^2} \Rightarrow -\frac{4}{\lambda^2} < 0$$

λ_0 is a point of local Maxima

$(0, \lambda_0)$ $f'(\lambda) > 0$
(λ_0, ∞) $f'(\lambda) < 0$



Ques Let $x_1, x_2 \ldots x_n$ be a Random sample from a geometric (p) distribution Find the MLE of p.

$$pmf: f(x) = p(1-p)^{x-1} \quad x = 1, 2, 3 \ldots$$

$$L(p \mid x_1, x_2 \ldots x_n) = \left[p(1-p)^{x_1-1}\right]\left[p(1-p)^{x_2-1}\right] \cdots \left[p(1-p)^{x_n-1}\right]$$

$$\Rightarrow p^n (1-p)^{x_1-1+x_2-1 \cdots x_n-1}$$

$$\Rightarrow p^n (1-p)^{\sum_{i=1}^{n} x_i - n} \quad p \in (0,1)$$

power

20/4/18

Ex1 Let $X_1, X_2 \ldots X_n$ be a random sample from geometric (p) population Find MLE of p.

Sol :- $f(p|x) = p(1-p)^{x-1} \quad x = 1, 2, 3 \ldots \quad p \in (0,1)$

$$f(p) = \log L(p) = n \log p + (y-n) \log(1-p)$$

$$y = \sum_{i=1}^{n} x_i$$

at which pt. f^n attains max value so either find global maxima of function or

$f'(p) = 0 \Rightarrow \frac{n}{p} - \frac{(y-n)}{(1-p)} = 0$

$n(1-p) - (y-n)p = 0$

$n - \cancel{np} - yp + \cancel{np} = 0$

$n(1-y) \neq 0 \qquad n = yp$

$\frac{1-p}{p} = \frac{y-n}{n}$

$\frac{1}{p} - \cancel{1} = \frac{y}{n} - \cancel{1}$

$\boxed{p = \frac{n}{y}}$

First we need to examine $\frac{n}{y}$ lie in the interval 0 to 1 or not.

$y = \sum_{i=1}^{n} xi > n$

$\frac{n}{y} < 1$

n is positive y is positive

$\boxed{p \in (0,1)}$ $\boxed{\frac{n}{y} \in (0,1)}$

$($ 0 — p — n/y — 1 $)$

at any arbitriray pt. 'p' we need to show that $f'(p)$ is positive.

→ Let us consider $p < n/y$ and at this pt. we need to $f'(p)$.

Also $-p > -n/y$

$1-p > 1-n/y$

$1-p > \frac{y-n}{y}$

$\boxed{y > \frac{y-n}{1-p}}$

Camlin Page
Date / /

$$\Rightarrow \frac{y-n}{1-p} < y < \frac{n}{p}$$

$$\frac{n}{p} - \frac{y-n}{1-p} > 0$$

$$\Rightarrow \boxed{f'(p) > 0}$$

$f(p)$ is strictly increasing over $(0, n/y)$

In the same manner we can ~~throw~~ show by ~~decreasing~~ considering $p > \frac{n}{y}$ $f'(p) < 0 \Rightarrow$ $f(p)$ is decreasing over $(\frac{n}{y}, 1)$.

$\Rightarrow$ $p = \frac{n}{y}$ is global maxima of $f(p)$.

i.e $\hat{p} = \frac{n}{\sum_{i=1}^{n} x_i}$ is MLE of p

Ex Let $X_1, X_2, \ldots X_n$ be a Random sample from uniform $(0, \theta)$ population find MLE of θ.

Sol $f(\theta/x) = \begin{cases} 1/\theta, & 0 < x < \theta \\ 0 & \text{otherwise} \end{cases}$

$L(\theta | x_1, x_2 \ldots x_n) = \begin{cases} 1/\theta^n & 0 < x_1 \leq x_2 \ldots < x_n < \theta \\ 0 & \text{otherwise} \end{cases}$

Likelihood function is non zero

Likelihood f^n $L(\theta)$ is non zero if $\theta = \underset{1 \leq i \leq n}{\text{Max}}\, x_i$

Scanned by CamScanner

→ Also $L(\theta)$ is strictly decreasing function of θ. so $L(\theta)$ attains its Max at the smallest value of θ.

→ Smallest value of θ?

$\theta = \text{Max } x_i$

$L(\theta)$ attains its Maxima at $\theta = \max\limits_{1 \le i \le n} x_i$

smallest value of θ = Max of x_i.

$$\hat{\theta} = \text{Max}\{x_1, x_2 -- x_n\}$$

Ex Let $X_1, X_2 -- X_n$ be a Random ~~value~~ sample from Binomial (m, p) population. find MLE of p?

Sol Consider tossing of coin

$P(H) = p \qquad P(T) = 1-p$

When coin is fair.

We are tossing this coin Independently Many times. counting number of times head appears.

$$P(X = k) = {}^{m}C_{k}\, p^{k} (1-p)^{n-k}$$
$$k = 0, 1, -- n$$

We have no Idea about p!

Based on the no. of times head appears we want to get idea how ~~many~~ Many times we tossed/flipped the coin. unknow parameter $= n$.

For this we need ~~use~~ MLE Technique

$$L(p \mid x_1, x_2 \ldots x_n) = \left[{}^mC_{x_1}\, p^{x_1} (1-p)^{m-x_1} \right] \left[{}^mC_{x_2}\, (p)^{x_2} (1-p)^{m-x_2} \right] \cdots \left[{}^mC_{x_n}\, p^{x_n} (1-p)^{m-x_n} \right]$$

Likelihood function $= \prod_{i=1}^{n} \left({}^mC_{x_i} \right) p^{\sum_{i=1}^{n} x_i} (1-p)^{mn - \sum_{i=1}^{n} x_i}$

$$= \prod_{i=1}^{n} \left({}^mC_{x_i} \right) p^{y} (1-p)^{mn-y}$$

$$y = \sum_{i=1}^{n} x_i$$

$$\boxed{L(p) = c\, p^{y} (1-p)^{mn-y}}$$

$$\log L(p) = \log c + y \log p + (mn-y) \log(1-p)$$

$$\boxed{\hat{p} = \frac{1}{mn} \sum_{i=1}^{n} x_i}$$

Ques Suppose $X_1, X_2 \ldots X_n$ be a Random sample from exponential (λ) population find MLE of λ.

Sol :- $L(p \mid x_1, x_2 \ldots x_n) = \left[\lambda e^{-\lambda x_1} \right] \left[\lambda e^{-\lambda x_2} \right] \cdots \left[\lambda e^{-\lambda x_n} \right]$

$$= \lambda^n e^{-\lambda (x_1 + x_2 \ldots x_n)}$$

$$\boxed{L(p) = \lambda^n e^{-\lambda \left(\sum_{i=1}^{n} x_i \right)}}$$

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

23|4|18

Interval Estimation :-

In IE instead of finding a point $\hat{\theta}$ estimator we find two point estimator ($\hat{\theta}_l$ and $\hat{\theta}_h$)

$\hat{\theta}_l(x_1, x_2 \ldots x_n)$
$\hat{\theta}_h(x_1, x_2 \ldots x_n)$] Both are f^n of $x_1, x_2 \ldots x_n$

(Random sample)

→ This will give us good estimation.

$$\boxed{P[\hat{\theta}_l < \theta < \hat{\theta}_h] \geq 1-\alpha}$$

If $\alpha = 0.01$ Then we are 99% sure that $\hat{\theta}$ is contained in b/w

→ $\hat{\theta}_l$ and $\hat{\theta}_h$ → Higher pt. estimator (upper)
lower pt estimator

→ That Interval is known as confidence Interval.

→ We consider an interval in such a manner that it will contain that unknown quantity θ.

Interval estimation :- Let $x_1, x_2 \ldots x_n$ be Random sample from a distribution with a parameter θ that is to be estimated. An Interval estimator with confidence level $(1-\alpha)$ consists of two estimators $\hat{\theta}_l(x_1, x_2 \ldots x_n)$

and $\hat{\theta}_h(X_1, X_2 \ldots X_n)$ such that $P(\hat{\theta}_l < \theta < \hat{\theta}_h) \geq 1-\alpha$ for every possible value of θ. we say that $[\hat{\theta}_l, \hat{\theta}_h]$ is a $(1-\alpha) \times 100\%$ <u>confidence</u> <u>Interval</u> of θ.

- In classical inference unknown quantity θ is fixed.

→ Problem is to find out confidence Interval we will use concept of Cumulative Distribution Function.

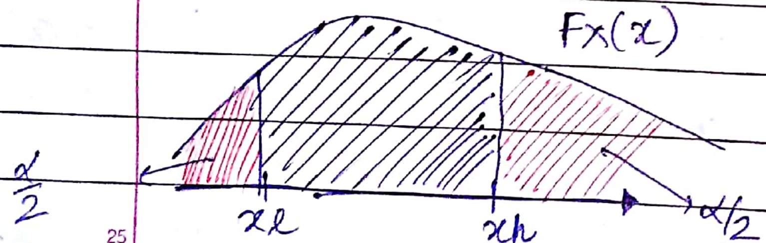
Finding Interval Estimators :-

Suppose for a Random Variable X we have CDF $F_X(x) = P(X \leq x)$

- we are looking x_l and x_h s.t.

$$P(x_l \leq x \leq x_h) = 1-\alpha$$

Because $1-\alpha$ is lower Bound of above Interval.



we are Interested in area under the curve.

Assumption

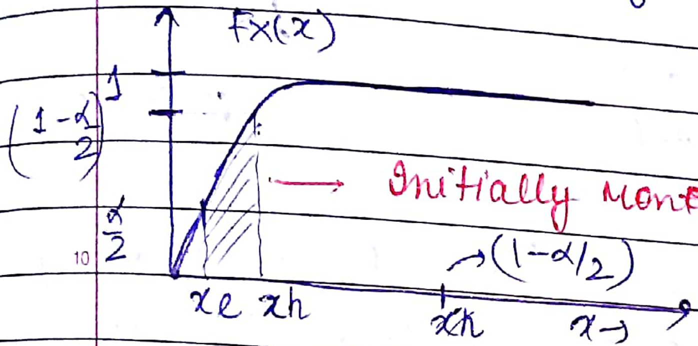
$$P(x < x_l) = \alpha/2 \text{ and } P(x > x_h) = \alpha/2$$

$$P(x < x_l) = \alpha/2 \qquad P(x < x_h) = 1 - \alpha/2$$

$F_X(x_l) = \frac{\alpha}{2}$ and $F_X(x_h) = 1-\alpha/2$

$$\boxed{x_l = F_X^{-1}(\alpha/2) \text{ and } x_h = F_X^{-1}(1-\alpha/2)}$$

Two points can be find out in this way.



→ Initially monotonically rising and then constant

Example: Let $Z \sim N(0,1)$ find x_l and x_h such that $P(x_l < Z < x_h) = 0.95$

Standard normal Random Variable

$1-\alpha = 0.95$ $\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$

$P(Z < x_l)$ $P(x < x_l) = 0.025$

$P(Z > x_h)$ $P(x < x_h) = 1 - 0.025 = 0.975$

0.975 corresponds to $\underline{1.976}$

By standard Table.

$P(Z < x_h) = 0.975$

$$\boxed{x_h = 1.976}$$

SNRV is symmetric about $\mu = 0$

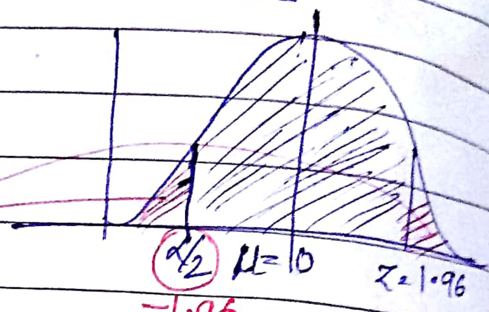
$P(Z < 1.976) = 0.976$ $\boxed{x_h = 1.976}$

Due to symmetricity of Normal curve we have:-

$\rightarrow x_l = F_X^{-1}(0.025) = -1.96$

$$\boxed{P(-1.96 < z < 1.96) = 0.95}$$

These are same



Example :- Let $X_1, X_2, \ldots X_n$ be a random sample from a normal distribution $N(\theta, 1)$. Find a 95% confidence interval for θ.

Solution : $\hat{\theta} = \bar{X} = \dfrac{X_1 + X_2 \ldots X_n}{n}$

$\bar{X} \sim N(\theta, \frac{1}{n})$

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

$\dfrac{X-\mu}{\sigma} \sim N(0,1)$

$\dfrac{\bar{X}-\theta}{\frac{1}{\sqrt{n}}} \sim N(0,1)$

$\sqrt{n}(\bar{X}-\theta) \sim N(0,1)$

$P(x_l < \sqrt{n}(\bar{X}-\theta) < x_h) = 0.95$

$P(-1.96 < \sqrt{n}(\bar{X}-\theta) < 1.96) = 0.95$

$P\left(\bar{X} - \dfrac{1.96}{\sqrt{n}} < \theta < \bar{X} + \dfrac{1.96}{\sqrt{n}}\right) = 0.95$

$$\boxed{\hat{\theta}_l = \bar{X} - \frac{1.96}{\sqrt{n}}}$$

$$\boxed{\hat{\theta}_h = \bar{X} + \frac{1.96}{\sqrt{n}}}$$

24/4/18

Pivotal Quantity :- Let $X_1, X_2 \ldots X_n$ be a random sample from a distribution with a parameter θ that is to be estimated. The Random Variable Q is said to be a pivot or a pivoted quantity. if it has ~~a~~ the following quantities →

1) It is a function of observed data $X_1, X_2 \ldots X_n$ and the unknown parameter θ but it is Independent of any other unknown parameter

$$Q = Q(X_1, X_2 \ldots X_n, \theta)$$

2) The probability distribution of Q does not depend on θ.

→ These are 3 steps in the pivotal Method for finding confidence Interval :

1) First, find a pivotal quantity $Q(X_1, X_2 \ldots X_n, \theta)$

2) Find an Interval for Q s.t.

$$P(q_l, Q, q_h) = 1-\alpha$$

3) Using some algebra, convert the above equation to the following →

Camlin Page
Date / /

$$\boxed{P(\hat{\theta}_l < \theta < \hat{\theta}_n) = 1-\alpha}$$

Ques Let $X_1, X_2 \ldots X_n$ be a Random sample from a distribution with known variance $Var(X_i) = \sigma^2$, and unknown Mean $E(X_i) = \theta$. Find a $(1-\alpha)$ confidence interval for θ. Assume n is large

Solution We will use central limit theorem (n is large)

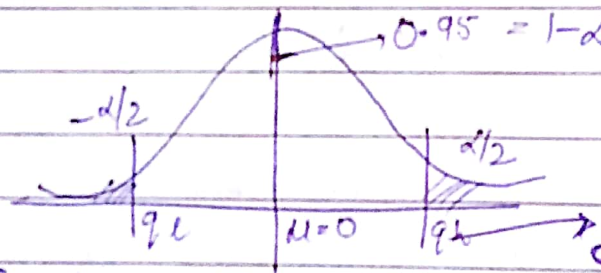
here :-

$$Q = \frac{\bar{X} - \theta}{\sigma/\sqrt{n}} \sim N(0,1)$$

Random sample $\sim N(\theta, \sigma^2)$

→ In this example distribution is not given

$$P\left(-Z_{\frac{\alpha}{2}} < Q < Z_{\frac{\alpha}{2}}\right) = 1-\alpha$$



Because of symmetricity of NRV

$$P\left(-Z_{\frac{\alpha}{2}} < \frac{\bar{X} - \theta}{\sigma/\sqrt{n}} < Z_{\frac{\alpha}{2}}\right) = 1-\alpha$$

$$P\left(\bar{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \theta < \bar{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$\left[\bar{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$$ is

$(1-\alpha) \times 100\%$ confidence interval for unknown parameter θ

Ans An Engineer is measuring a quantity θ. It is assumed that there is a Random error in its measurement, so the engineer will take N measurements and Report the average of the measu. as the estimated value of θ. Here N is assumed to be large enough so that CLT applies. If X_i is the value i.e. obtained in the i^{th} measurement, we assume that $X_i = \theta + W_i$, where W_i is the error in the i^{th} measurement. We assume that W_i's are IID with $E(W_i) = 0$ and $Var(W_i) = 4$. The Engineer reports the average of the measurements $\bar{X} = \frac{X_1 + X_2 + \cdots X_n}{n}$.

How many measurements does the engineer need to make until he is 90% that the final error is < 0.25.

Value of n s.t.

Sol:-

$$P(\theta - 0.25 < \bar{X} < \theta + 0.25) \geq 0.90$$

$$\alpha = 1 - 0.9$$

$$\alpha = 0.1$$

$$\frac{\alpha}{2} = 0.05$$

$$X_i = \theta + w_i$$

$$E(X_i) = E(\theta) + E(W_i)$$

$$E(X_i) = E(\theta)$$

$$Var(X_i) = \overset{0}{\cancel{Var(\theta)}} + Var(w_i)$$

$$\boxed{Var(X_i) = 4}$$

$\sigma^2 = 4$

$$Q = \frac{\bar{x} - \theta}{\sigma/\sqrt{n}} \sim N(0,1)$$

$\sigma = 2$

$$Q = \frac{\bar{x} - \theta}{2/\sqrt{n}} \sim N(0,1)$$

$$z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 0.25$$ (comparing with last ex)

$$\frac{z_{\alpha}}{\sqrt{n}} = 0.25$$

$$z_{\alpha/2} = 0.25 \times \sqrt{n}/2$$

$$z_{\frac{\alpha}{2}} = 1.645$$ for 90% confident

$$1.645 = \frac{0.25 \times \sqrt{n}}{2}$$

$$\boxed{n \geq 174}$$