$^nC_2 + {}^nC_3 \dots$

Date / /
Page

→ if you take $n=2$ ⇒ get 91 : get much better estimation as $n \uparrow$

02-04-18 Law of Large Numbers

↓ Weak law of large Numbers

↓ Strong law of large Numbers

Q When do you say that a countable collection of r.v. $\{X_1, X_2, \dots X_n\}$ is independent ?

Ans Def^n : we say r.v. X & Y are independent if

$$P(\{X \in A\} \cap \{Y \in B\}) = P(X \in A) \cdot P(Y \in B)$$

for every Borel subsets $A, B \subseteq \mathbb{R}$

Def^n : we say that r.v. $X_1, X_2, \dots X_n$ are independent $(n \geq 2)$ if :

$$P(X_{i_1} \in A_1, X_{i_2} \in A_2, \dots, X_{i_m} \in A_m) = P(X_{i_1} \in A_1) \cdot P(X_{i_2} \in A_2) \dots P(X_{i_m} \in A_m)$$

For every subcollection $\{i_1, i_2, \dots, i_m\} \subset \{1, 2, \dots, n\}$ & $\forall A_1, A_2, \dots, A_m \subseteq \mathbb{R}$

Definition :- We say that $\{X_1, X_2, \dots, X_n\}$ is independent if every finite subcollection of $\{X_1, X_2, \dots, X_n\}$ is independent (same as in multivariable)

① Weak Law of Large Numbers (Expec^n should exist)

Let $X_1, X_2, \dots$ be a sequence of independent & identically distributed r.v., each having finite mean μ. Then, $\forall \delta > 0$,

$$\lim_{n \to \infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \delta\right) = 0 \qquad \text{--- ①}$$

∞

Scanned by CamScanner

(for same parameters) $\rightarrow$ weak law

or equivalently,

$$\lim_{n \to \infty} P\left(\left| \frac{S_n}{n} - \mu \right| < \delta \right) = 1 \quad \text{— ②}$$

where $S_n = X_1 + X_2 + \dots + X_n$

↳ ① $\longrightarrow$ seq. $\left| \frac{S_n}{n} - \mu \right| \geq \delta$ is converging to 0. (M-I)

$\Rightarrow$ $\forall\ \epsilon > 0, \delta > 0$: $\exists\ n_0(\delta, \epsilon) \in \mathbb{N}$ s.t.

$$P\left(\left| \frac{S_n}{n} - \mu \right| \geq \delta \right) \leq \epsilon \quad \forall\ n \geq n_0$$

→ threshold

(no mod required since Prob. is always +ve)

- Interpretation (weak law)

let : δ : Accuracy level $\quad\quad \epsilon$: Confidence

② $\Rightarrow$ $\forall\ \epsilon > 0, \delta > 0$: $\exists\ n_0 \in \mathbb{N}$ s.t.

$$\left| P\left(\left| \frac{S_n}{n} - \mu \right| < \delta \right) - 1 \right| < \epsilon \quad \forall\ n \geq n_0$$

↳ -ve ($P(\ ) \leq 1$)

$\Rightarrow$ $P\left(\left| \frac{S_n}{n} - \mu \right| < \delta \right) > 1 - \epsilon = 0.9 \rightarrow$ 90% confident that value of $\frac{S_n}{n}$ is in nbd : $\frac{S_n}{n} \in [\mu - 0.01, \mu + 0.01]$

$\delta = 0.01$

$S_n \rightarrow$ also a random variable (sum of r.v.)

$\frac{S_n}{n} \longrightarrow \mu \quad : \delta \quad$ with ϵ

value $\quad$ close to this value $\quad$ with this much accuracy $\quad$ this much accuracy

Let $\epsilon = 0.1$

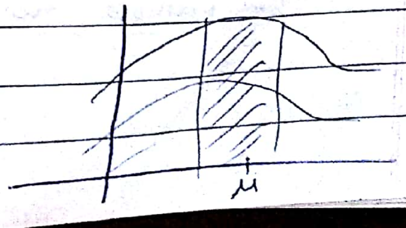
$1 - \epsilon = 0.9$

↳ suppose S_n is also pdf & $X_1, X_2, \dots X_n$ are all pdfs

① $\Rightarrow$ $\frac{S_n}{n}$ approaches to μ as $n \uparrow$

(concentrated to μ)

how near to μ : derived from δ



* Impirical Science : applies weak law of large No.

not required in case of weak law.

Example : let $X_1, X_2, \ldots$ be iid r.v. with $EX_1 = 0$ and $Var(X_1) = 1$

let $S_n = X_1 + X_2 + \ldots + X_n$. Then, for any $x > 0$, find

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

can be written as

sol^n $\mu = 0$

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| < x\right)$$ with $\mu = 0$

↳ satisfies all parameters of weak law.

$$\Rightarrow \lim_{n\to\infty} P(\quad) = 1$$

→ Random Sample : eg: find μ of heights of all students of clg

Impirical science

μ = true mean (getting mean from all original data)
(finding μ of all students) → Not practical

height : ___ distributed ___

random sample $\boxed{X_1, X_2, \ldots, X_n}$ → height of 'n' students (choose N random students & find mean)

independent r.v.

$$S_n = X_1 + X_2 + \ldots + X_n$$

$\frac{S_n}{n} \Rightarrow$ sample mean

weak law $\Rightarrow$ as $n \uparrow$, $\frac{S_n}{n}$ (sample mean) will be ~~cot~~ close to μ (original mean)

⇓

$\frac{S_n}{n}$ is a good approximation to true mean μ for large value of n. (For large accuracy & confidence : may we to wait for higher value of n)

$\delta = 0.1$ $\quad \epsilon = 0.1$ $\quad n_1$ (90% confident)

$\delta = 0.01$ $\quad \epsilon = 0.05$ $\quad n_2 \geq n_1$ (95% ")

$$\text{sample mean} \xrightarrow{\text{Converges}} \text{True mean}$$

Syllabus : multiple r.v. to 3 Inequalities

Date / / Page

Example: let $X_1, X_2, \ldots$ be independent Bernoulli trials with probability of head in each coin toss 'p'. $EX_1 = p$

$$\lim_{n\to\infty} \left(\left| \frac{S_n}{n} - p \right| \geq \epsilon \right) = 0$$

$S_n = X_1 + X_2 + \cdots + X_n \Rightarrow$ Binomal r.v. $\sim B(n,p)$

(But can't say how fast it is converging to 0)

can use Chebyshev

$\Rightarrow P\left(|S_n - np| \geq n\delta \right)$ ($n\delta \to \epsilon$)

(np ↳ Expec^n in $B(n,p)$)

$$\Rightarrow P\left(\quad \right) \leq \frac{\sigma^2}{\epsilon^2} = \frac{\cancel{n}p(1-p)}{n^{\cancel{2}}\delta^2} \leq \frac{1}{4n\delta}$$

$\Rightarrow f(p) = p(1-p)$, $p \in [0,1]$

$\Rightarrow f(p)$: max at $p = \frac{1}{2}$

3-4-18

Probability and Relative Frequency:

↳ Interpretation of Prob. of event

→ $P(H) = 0.25$ in 1 coin toss $\Rightarrow$ If you toss coin large no. of times, ~~you'll get H_n 1/4 th times~~. approximately

Consider an event A defined in some random experiment. Let $p = P(A)$. Consider 'n' independent repetitions of the random experiment & let M_n be the relative frequency of event A in these n trials

Define : $X_i(\omega) = \begin{cases} 1 & : \omega \in A \quad \text{(if event A occurs)} \\ 0 & : \omega \notin A \end{cases}$

i^{th} repetition of experiment

$$M_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$ → total no. of instances when A is occuring in n trials.

independent repetition $\Rightarrow X_i$ are also independent.

Scanned by CamScanner

Date / /
Page

X_i : discrete $\qquad P(X_i=1) = P(A) = p$
$\qquad\qquad\qquad P(X_i=0) = P(A^c) = 1-p$ → pmf don't depend on ~~Pos~~ i ⇒ identically distributed

$EX_i = 1(p) + 0(1-p) = p$ → has finite mean

can apply weak law of large no.

$$\boxed{\Rightarrow \lim_{n\to\infty} P\left(\left|\frac{M_n}{n} - p\right| < \delta\right) = 1 \quad \forall \quad \delta > 0}$$ Probability as relative frequency

* for sufficiently large value of n, relative frequency M_n will be very close to probability p.

Example continued....

Tossing a coin infinitely many times

$\Omega = \{\omega : \omega = (\omega_1, \omega_2, \omega_3, \ldots\ldots), \text{ where } \omega_i \text{ is either H or T}\}$

Uncountably Infinite Sample Space

defined by

$X_1 \to 1^{st}$ toss

$X_2 \to 2^{nd}$ toss

$$X_1(\omega) = \begin{cases} 1 & \omega_1 = H \\ 0 & \omega_1 = T \end{cases} \quad \to \text{like Bernoulli r.v.}$$

completely decided by 1^{st} coin toss

$$X_2(\omega) = \begin{cases} 1 & \omega_2 = H \\ 0 & \omega_2 = T \end{cases}$$

$X_1, X_2, \ldots$ → are independent Bernoulli r.v. with parameter p (Just like prev. section)

Applying Weak law of large no, $\forall\ \delta > 0$

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - p\right| < \delta\right) = 1$$: don't tell how fast/slow S_n approaches p.

By Chebyshev's Inequality,

$$P\left(\left|\frac{S_n}{n} - p\right| \geq \delta\right) \leq \frac{1}{4n\delta^2} \quad \text{(derived earlier)} \quad \forall\ n \in N$$

Here,
Relative frequency of Head in n coin tosses

Date / /
Page

Let $\epsilon > 0$ be given. Then,

$$P\left(\left|\frac{S_n}{n} - p\right| \geq \delta\right) < \epsilon \quad \forall \quad n \geq n_0$$ ↳ want to determine this no.

$$\frac{1}{4n\delta^2} < \epsilon \Rightarrow n > \frac{1}{4\epsilon\delta^2}$$ ↳ choice for n_0

↳ $\boxed{\epsilon = 0.02, \ \delta = 0.05}$ (98% confidence)

→ $n_0 = \underset{0.0025}{12,500}$ → consider these much coin toss if you want this much confidence & accuracy

want higher value of confidence, have to wait this much
↳ can improve very significantly if we use better theorems

$$\Rightarrow P\left(\left|\frac{S_n}{n} - p\right| < 0.02\right) \geq 0.95$$

1.) If p is known, can predict outcome of event with what prob.
2.) If p is unknown, perform repeated times to get $\frac{S_n}{n}$ & get app. value of p :→ Monte Carlo Simulation

② Strong Law of Large Numbers

Let $X_1, X_2, \ldots$ be a sequence of iid r.v. such ~~tath~~ that $E X_i \overset{\mu}{<} \infty$. Then $\forall \ \delta > 0$:

$$P\left(\lim_{n\to\infty}\left|\frac{S_n}{n} - \mu\right| < \delta\right) = 1 \quad \text{where} \quad S_n = \sum_{i=1}^{n} X_i \qquad —③$$

$\boxed{\text{Strong} \Rightarrow \text{Weak}}$ Converse is not true.

→ sequence of func^n

→ $\frac{S_n}{n}$ → R.V. $Y_n = \left|\frac{S_n}{n} - \mu\right|$

$$\boxed{\lim_{n\to\infty} Y_n < \delta \quad \forall \ \delta > 0} \Rightarrow$$

$Y_n(\omega)$ → sequence of real no.

$$\lim_{n\to\infty} Y_n(\omega) < \delta \quad \forall \quad \delta > 0$$

Scanned by CamScanner

③ ⇒ $P\left(\lim_{n\to\infty} \frac{S_n}{n} = \mu\right)$

$\lim_{n\to\infty} \frac{S_n(\omega)}{n} = \mu$ for almost all $\omega \in \Omega$

One $\exists$ $N \subset \Omega$ s.t. $P(N) = 0$: There may be sequence in which limit d.n.e. & $\frac{S_n}{n} \neq p$.

One possibility HHHH ... H $\frac{S_n(\omega)}{n} = 1 \neq p$

But overall, $\lim_{n\to\infty} \frac{S_n}{n} = \mu$

→ It is much stronger than weak law.

Difference between WLLN and SLLN :-

Weak law says that

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \delta\right) \to 0 \text{ as } n \to \infty$$

→ probability of significant deviation of $\frac{S_n}{n}$ from $\mu \to 0$ as $n \to \infty$

Still, for any finite n, this probability can be positive

(may deviate from μ at finite no. of values of n though with small prob)

don't tell about these values (SLLN does)

Strong law says that $\forall$ $\delta > 0$

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \delta\right) \text{ infinite no. of times is } 0$$

4/4/18

Central Limit Theorem :

Let $X_1, X_2, \ldots$ be a sequence of iid r.v. with each having finite mean μ and non-zero variance $\sigma^2 > 0$. Define

$$S_n = X_1 + X_2 + \cdots + X_n$$

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$ Then,

★ $$\boxed{\lim_{n\to\infty} \underbrace{P(Z_n \leq x)}_{\text{cdf of } Z_n \text{ at } x} = \Phi(x) \quad \forall \; x \in \mathbb{R}}$$

when $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt$ ⇒ cdf of std. Normal r.v.

$\Rightarrow$ CDF of Z_n ~~abs~~ converges to CDF of $N(0,1)$ at every point on R

$\Downarrow$

point wise convergence.

$n\mu = E[S_n]$

$Var(S_n) = n(\sigma^2)$ $\quad [Var(S_n) = Var(X_1) + Var(X_2) + \ldots + Var(X_n)]$

$\sigma(S_n) = \sqrt{n}\,\sigma$

$\Rightarrow \underline{\underline{Z_n}} = \dfrac{S_n - n\mu}{\sigma\sqrt{n}} \Rightarrow \dfrac{X-\mu}{\sigma} \Rightarrow$ ~~Z_n is:~~ normalised r.v.

$E(Z_n) = \dfrac{E[S_n] - n\mu}{\sigma\sqrt{n}} = 0$

$Var(Z_n) = \dfrac{Var(S_n)}{(\sigma\sqrt{n})^2}$ $\rightarrow$ ~~depends only on coeff of r.v. S_n~~ so, only $\sigma\sqrt{n}$ is considered $(a^2\sigma^2)$

$= \underline{\underline{1}}$

Example : $X_i \sim$ Bernoulli (p) $\quad EX_i = p \quad Var(X_i) = p(1-p)$

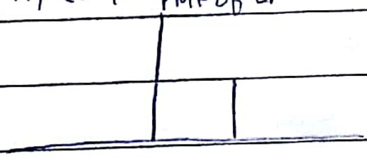
$S_n = X_1 + X_2 + \ldots + X_n$

$\Rightarrow S_n \sim B(n,p)$

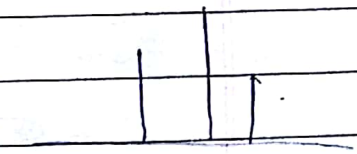
$Z_n = \dfrac{S_n - np}{\sqrt{np(1-p)}}$ $\quad$ take $p = \dfrac{1}{3}$.

PMF of Z_1

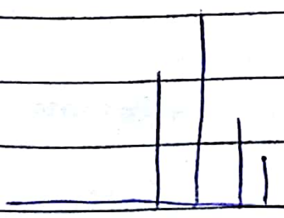
$\underline{n=1}$



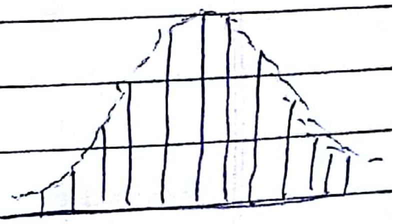
$n=2$



$n=3$



$n=30$



$\Downarrow$

takes shape of pdf of std. ~~rv~~ normal r.v.

Hence, we can conclude

CDF of $Z_{30} \simeq$ CDF of std. normal r.v.

Scanned by CamScanner

Date / /
Page

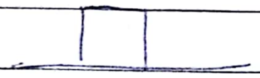
Example: $X_i \sim$ Uniform $(0,1)$

$EX_i = \frac{1}{2}$ Var $(X_i) = \frac{1}{12}$

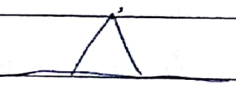
$Z_n = \dfrac{S_n - n/2}{\sqrt{n/12}}$ sum of independent r.v. ⇓ can use formula learnt earlier

So, we know Z_n has pdf.

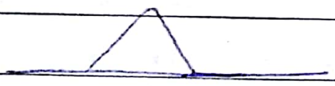
Z_1 : pdf of Z_1



Z_2 :



Z_3 :



Z_{30}

~~good~~ seems to converge to pdf of $N(0,1)$

Normal Random Variable :-

$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$ → convergent, finite ~~we numerical methods~~
↳ closed form of anti-derivative ~~for this func^n~~

↓

can use numerical method to find app. value

pdf of ~~$\Phi(x)$~~ $N(0,1)$ is symmetric ⇒

~~$\Rightarrow \Phi(x) = \Phi(-x)$~~

If $x < 0 \Rightarrow -x > 0$ $X \sim N(0,1)$

$\Phi(x) = P(X \leq x)$

$= 1 - P(X > x)$

$= 1 - P(X \leq -x)$ {symmetric graph}

~~$= 1 - \Phi(-x)$~~

Normal Approximation based on CLT :

CLT : $\lim_{n\to\infty} P(Z_n \leq x) = \Phi(x)$

⇒ For large values of n

$P(Z_n \leq x) \approx \Phi(x)$

Date / /
Page

$X_1, X_2, \ldots$ are iid r.v. & $S_n = X_1 + \cdots + X_n$

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} \approx N(0,1) \text{ if } n \to \infty$$

$\rightarrow$ ~~std N r.v.~~ $\qquad Z \sim N(0,1)$

$$\Rightarrow S_n = \sigma\sqrt{n}\, Z_n + n\mu \quad \text{new mean}, \rightarrow \text{new variance}$$

$$\hookrightarrow N(n\mu, \sigma^2 n) \sim S_n$$

$$P(S_n \le c) = P\left(\sigma\sqrt{n}\, Z_n + n\mu \le c\right)$$

$$= P\left(Z_n \le \frac{c - n\mu}{\sigma\sqrt{n}}\right) = \Phi\left(\frac{c-n\mu}{\sigma\sqrt{n}}\right)$$

$\rightarrow$ Let $S_n = X_1 + \cdots + X_n$, where X_i are iid with mean μ & var σ^2

If n is large,

$$P(S_n \le c) \simeq \Phi\left(\frac{c-n\mu}{\sigma\sqrt{n}}\right)$$

STEP 1 : Compute $n\mu$, $n\sigma^2$

STEP 2 : $P(S_n \le c) = \Phi\left(\frac{c-n\mu}{\sigma\sqrt{n}}\right)$

Example : We load on a plane 100 packages whose weights are iid ~~r.v~~ uniform [5,50] r.v. What is the probability that the total weight will exceed 3000 kg.

Sol^n : $n = 100 \quad X_i$: weight of i^{th} pkg.

$S_n = X_1 + X_2 + \cdots + X_{100}$ $\qquad \mu = \frac{55}{2} = 27.5$

$W = S_n$

~~$P(W > 3000)$~~ =

$EX_i = 27.5 \qquad Var(X_i) = 168.75$

$EW = 2750 \qquad Var(W) = 16875$

$$P(W > 3000) = 1 - P(W \le 3000)$$

$$= 1 - \Phi\left(\frac{3000 - 2750}{\sqrt{16875}}\right) = 1 - \Phi(1.92)$$

$$= 1 - [0.9726]$$

$$= 0.0274$$

→ chapter 8 of online book

6/4/18

STATISTICS (sample ≡ Random sample)

→ In Probability theory : ∃ a Prob. Model & each problem will have unique answer.

→ In statistics : ~~no~~ don't have any correct answer, may get diff. answers.

I. Deciding on the basis of testing a few samples from a shipment of a certain drug whether the quality of shipment is satisfactory.

II. Using polling data, estimate the fraction of voter population that prefers Modi over Rahul.

In probability :

let $X \sim N(100, 15)$. Find $P(X > 100)$

In statistics : we won't discuss how we will assume

Given a data, we assume that the data is coming from normal r.v. Now, we want to estimate mean & variance of this normal r.v. using the table (given data)

Statistical Inference :

We think of data as observed values of a r.v. X. The distribution of X is assumed to be in certain family of distributions. The problem is to decide on the basis of data, which member of the family could represent the distribution of X.

$X \sim N(0,1)$ $Y \sim N(2,4)$: X & Y are not identically distributed, but X & Y belong to Normal family

Classical (or Frequentist) Inference :

$\theta \longrightarrow$ fixed (not a r.v.) constant (or vector)

Bayesian Inference :

(Big θ) $\Theta \Rightarrow$ considered as r.v.

Date / /
Page

studied earlier in Weak law of large No. (same eg. : heights)

↳ Random Sampling : The r.v. $X_1, X_2, \ldots, X_n$ are called a random sample of size 'n' from the population which is ~~determined~~ having ~~common~~ pmf / pdf $f(x)$. If $X_1, X_2, \ldots, X_n$ are iid r.v. with pmf / pdf as $f(x)$ "

↳ Sampling with Replacement / without Replacement → Independence is lost in this case.
will be retained — app. same
↳ sample of heights ⇒ X_2 is not strictly independent of X_1 (won't chose X_1 again)

Practically, most of times → we use without Replacement.
But Sampling with Replacement is easy (independent)
→ We assume " " with reasoning ~~up~~ if popⁿ size is too large, ~~we~~ as r.v. will be almost independent from each other

⇒ We would like to estimate the avg. height in the ~~p~~ population. We may define an estimator

Point estimator of avg. height in popⁿ ← $\hat{\Theta} = \frac{X_1 + X_2 + \cdots + X_n}{n}$
r.v.

$\hat{\Theta} = \hat{\Theta}$ → an estimate of the avg. value.

↳ Sample Mean : Let $X_1, X_2, \ldots, X_n$ be a random sample

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

↳ Sample Variance :

$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$ → ~~Sample will be to~~ ∃ reason why we have 'n−1', not 'n'

↳ ~~Std~~ Sample Standard deviation :

$S = \sqrt{S^2}$

→ Sample variance :

$ES^2 = \sigma^2$
↳ variance of popⁿ

Date / /
Page

$\rightarrow \quad E s^2 = \frac{n-1}{n}\sigma^2$ {If you use 'n' in denominator}

Let $f(x)$ be pmf / pdf of popⁿ
and μ, σ^2 are mean & variance of $f(x)$.

$\Rightarrow \quad E X_i = \mu \quad , \quad Var(X_i) = \sigma^2 \quad \forall i = 1, 2, \ldots, n$

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

$$\rightarrow E[S^2] = \frac{1}{n-1}\sum_{i=1}^{n} E[X_i - \bar{X}]^2$$

($\bar{X}$ is r.v. here, so, can't write $(X_i - \bar{X})^2 = Var$)

$$= \frac{1}{n-1} E\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right]$$

$$= \frac{1}{n-1}\left\{E\sum_{i=1}^{n} X_i^2 - n\,E[\bar{X}^2]\right\}$$

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

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

7/4/18

Example: Let X be the height of a randomly chosen individual from a populati[on]. In order to estimate mean and variance of X, we observe a random sample $X_1, X_2, \ldots X_7$. Then, X_i is iid with distribution as r.v. X. We obtain the following values:

166.8, 171.4, 169.1, 178.5, 168.0, 157.9, 170.1

Find the sample mean, sample variance & sample S.D. for the observed sample.

Solⁿ:

$$\bar{x} = 168.8$$

$$s^2 = \frac{1}{7-1}\sum_{n=1}^{7}(x_i - \bar{x})^2 = 37.7$$

$$s = \sqrt{s^2} = 6.1$$

Scanned by CamScanner

Median : Arrange data in ascending / descending order.

$$\text{Median} = \begin{cases} \left(\frac{n+1}{2}\right)^{th} \text{ entry} & : n : \text{odd} \\ \left(\frac{\frac{n}{2} + \left(\frac{n+1}{2}\right)}{2}\right)^{th} \text{entry} & : n = \text{even} \end{cases}$$

Date / /
Page

Order Statistics :

The Order statistics of a random sample $X_1, \ldots, X_n$ are the sample values placed in ascending order. They are denoted by : $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$. Each of these are r.v. & satisfy relation :

Example : 4 workers → ₹5000 each (salary) ₹15000 → 1 supervisor

Avg. salary $= \frac{20000 + 15000}{5} = 7000$ → neither represents salary of worker / supervisor

Order statistics ← median : 5000 5000 5000 5000 15000

5000 : gives some meaningful info. ⇒ more meanful

* mean gets affected by extreme values. In that case, median gives more relevant ~~data~~ info. about data

---

$$\underset{\substack{\uparrow \\ \text{min value} \\ \text{in sample}}}{X_{(1)}} \le X_{(2)} \le \ldots \le \underset{\substack{\uparrow \\ \text{Largest} \\ \text{value in sample}}}{X_{(n)}}$$

→ It is not that $X_{(1)}$ = some fixed value, It'll be decided by values of $X_1, X_2, \ldots$ (r.v.)

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

$X_{(2)} = \text{second min} \{X_1, X_2, \ldots X_n\}$

⋮

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

Thm : Let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ denote the Order statistics of a random sample $X_1, X_2, \ldots, X_n$ from a continuous population with cdf $F_X(x)$ & pdf $f_X(x)$. Then, the pdf of $X_{(j)}$ is given by :

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

CDF of X_j :

$$F_{X_{(j)}}(x) = \sum_{k=j}^{n} \binom{n}{k} \left[F_X(x)\right]^{k} \left[1 - F_X(x)\right]^{n-k}$$

Date / /
Page

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

$$f(x_1, \ldots, x_n) = \begin{cases} n!\, f_X(x_1) \cdots f_X(x_n) & ; -\infty < x_1, x_2, \ldots x_n < \infty \\ 0 & ; \text{Otherwise} \end{cases}$$

Example: Let X_1, X_2, X_3, X_4 be a random sample from Uniform $(0,1)$ distribution. Then, find the pdf of order statistics $X_{(1)}, X_{(2)}$ & $X_{(4)}$

Soln:

$$f_X(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{Otherwise} \end{cases} \qquad F_X(x) = \begin{cases} 0 & x \le 0 \\ x & 0 \le x \le 1 \\ 1 & x \ge 1 \end{cases}$$

$$f_{X_{(1)}}(x) = \frac{4!}{0!\,3!} f_X(x) \left[F_X(x)\right]^0 \left[1 - F_X(x)\right]^3$$

↳ defined only when $0 \le x \le 1$

$\frac{4!}{1!\,3!}$

$$f_{X_{(1)}}(x) = \begin{cases} 4 \cdot 1 \cdot (1-x)^3 & 0 \le x \le 1 \\ 0 & \text{Otherwise} \end{cases}$$

Check. $\int_0^1 4(1-x)^3\, dx = \underline{1}$.

$$f_{X_{(2)}}(x) = \begin{cases} 12\, x(1-x)^2 & 0 \le x \le 1 \\ 0 & \text{Otherwise} \end{cases}$$

$$f_{X_{(4)}}(x) = \begin{cases} 4x^3 & 0 \le x \le 1 \\ 0 & \text{Otherwise} \end{cases}$$

9-4-18

Set up of Statistical Inference Point Estimation

Given a data, we assume that data is realization or observed values of a r.v. X with a pmf or pdf $f(x|\theta)$ where θ is an unknown parameter of pmf, pdf.

Draw a random sample $X_1, X_2, \ldots, X_n$: ie, $X_i \sim X \ \forall i = 1, \ldots, n$

X_i are iid.

→ formula for pmf also exists, not in textbook.
! careful b/w estimator & estimate.

Definition: A point estimator is any func^n of the random sample.

Example: Height distribution of the resident of LNMIIT campus.
→ $f(x|\theta)$ where $\theta = EX$ (avg.) → want to estimate this
X = height of a resident of LNMIIT Campus.

$\hat{\theta}$: point estimator of θ
(i) ⇒ $\hat{\theta} = \dfrac{X_1 + \dots + X_n}{n} = f(X_1, \dots, X_n)$

Estimate: realized value of an estimator (it's a no.)

→ estimator: a func^n ∴ can be anything
(ii) $\hat{\theta} = X_i \quad \forall\ i \in \{1, 2, \dots, n\}$: is also an estimator, whether it can estimate accurately or not is another thing.

→ Bias of point estimator: Let $\hat{\theta} = h(X_1, X_2, \dots, X_n)$ be a point estimator for a parameter θ. The bias of $\hat{\theta}$ is defined as:
$$B(\hat{\theta}) = E[\hat{\theta}] - \theta$$
* For good estimator, $B(\hat{\theta}) = 0$, On an avg., estimator takes same value as parameter θ itself.

We say that a point estimator $\hat{\theta}$ is an unbiased estimator of θ if ~~B(θ)~~ $B(\hat{\theta}) = 0$ for all possible values of θ. : → not fixed value
{ same as exponential : $\lambda \in [0, \infty)$ }

Example: (i) $E[\hat{\theta}] = E\left[\dfrac{X_1 + X_2 + \dots + X_n}{n}\right]$ $\{E[X_i] = \theta\}$
$= \dfrac{\theta + \theta + \dots + \theta}{n} = \theta$

⇒ $B(\hat{\theta}) = 0$ irrespective of population size

(ii) $E[\hat{\Theta}] = E[X_i] = \Theta$ → Biased : gives info about mean of point estimator

→ $B(\hat{\Theta}) = 0$

↳ unbiased estimator of population

* Both (i) and (ii) are unbiased estimator of population. But in (ii), we're just taking into account i^{th} value. → not good estimator. Whereas, (i) will give us better estimate → Just having biased is not enough.

→ **Mean square Error (MSE)** : MSE of a point estimator $\hat{\Theta}$, denoted by $MSE(\hat{\Theta})$ is defined as : (should be min^m)

$$MSE(\hat{\Theta}) = E[(\hat{\Theta}-\Theta)^2]$$: gives info. about variance of pt. estimator

$$= Var(\hat{\Theta}-\Theta) + (E(\hat{\Theta}-\Theta))^2$$ {Θ : const. here ⇒ $Var(\hat{\Theta}-\Theta) = Var(\hat{\Theta})$}

$$= Var(\hat{\Theta}) + (E(\hat{\Theta}) - \underbrace{\Theta}_{const})^2$$

$$\boxed{MSE(\hat{\Theta}) = Var(\hat{\Theta}) + (B(\hat{\Theta}))^2}$$

may/may not be 0. ⇒ better estimator when biased = 0

(i) MSE of sample mean :

$$= E[(\bar{X}-\Theta)^2]$$

$$= Var(\bar{X}) + \cancel{[B(\bar{X})]^2}^{\,0}$$

$$= Var(\bar{X})$$

$$= E[(\bar{X}-E\bar{X})^2]$$ from 1^{st} lecture

$$= E[(\bar{X}-\Theta)^2]$$

$$= Var\left(\frac{X_1+X_2+\dots+X_n}{n}\right) = \boxed{\frac{\sigma^2}{n}}$$, σ^2 : pop^n variance

(ii) MSE of $\hat{\Theta} = X_i$ is,

$$= Var(X_i) = \boxed{\sigma^2}$$

as $n \geq 2$ ⇒ $MSE_{(i)} \leq MSE_{(ii)}$ (as $n \uparrow$) ⇒

Only sample size is changing → dist. fundamental nature is not changed.

$\hat{\Theta}_1 = X_1$

$\hat{\Theta}_2 = \frac{X_1 + X_2}{2}$ ...

→ Consistency : let $\hat{\Theta}_1, \hat{\Theta}_2, \ldots$ be a sequence of point estimators of Θ, we say that $\hat{\Theta}$ is consistent estimator of Θ if

$$\lim_{n\to\infty} P\left(\left|\hat{\Theta}_n - \Theta\right| \geq \epsilon\right) = 0 \quad \forall \ \epsilon > 0$$

: as $n\uparrow$ (sample size), pt. estimator is close to true mean (Θ).

(WLLN)

(i) sample mean is a consistent estimator

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + \ldots + X_n}{n} - \underset{EX_i}{\Theta}\right| \geq \epsilon\right) = 1 \text{: By WLLN} \Rightarrow \underline{= 0}.$$

10/4/18

Th^m : let $\hat{\Theta}_1, \hat{\Theta}_2, \ldots$ be a sequence of point estimators of Θ. If $\lim_{n\to\infty} MSE(\hat{\Theta}_n) = 0$, then

$\hat{\Theta}_n$ is a consistent estimator of Θ.

Proof : $MSE[\hat{\Theta}_n] = E[(\hat{\Theta}_n - \Theta)^2]$ → Mean sq. exists

$Y = |\hat{\Theta}_n - \Theta| \geq 0$

To show :

$$\lim_{n\to\infty} P\left(|\hat{\Theta}_n - \Theta| \geq \epsilon\right) = 0 \quad \forall \ \epsilon > 0$$

⇒ can use Markov's inequality

(Mean square exists ⇒ can use $n=2$)

$$\Rightarrow P\left(|\hat{\Theta}_n - \Theta| \geq \epsilon\right) \leq \frac{E[Y^2]}{\epsilon^2} = \frac{MSE(\hat{\Theta}_n)}{\epsilon^2} \to 0$$

as $n \to \infty$.

obviously ~~Hence Proved~~

$$0 \leq \lim_{n\to\infty} P(Y \geq \epsilon) \leq 0$$

$$\Rightarrow \lim_{n\to\infty} P(Y \geq \epsilon) = 0$$

Hence Proved.

Date / /
Page

→ sample mean is an unbiased, consistent estimator of the population mean. (Same for variance)

→ Variance : Let X be a r.v. with variance $\sigma^2 > 0$. To estimate σ^2.

$$\sigma^2 = E[(X-\mu)^2] \quad \text{where} \quad \mu = EX.$$

Let $Y = (X-\mu)^2$

sample mean for r.v. Y would be an unbiased, consistent estimator.

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

estimator of sample mean of Y

↳ known quantity

(~~Pt estimator is~~ func^n of r.v. (X_i) but it can't be func^n of unknown parameter → μ has to be known)

In practice, μ is not known

→ Replace it with sample mean

Let $\bar{S}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$

↓

an estimator of population variance σ^2

→ want to check whether $\bar{S}^2$ is biased or not

$$\sum_{i=1}^{n}(X_i-\bar{X})^2 = \sum_{i=1}^{n} X_i^2 - 2X_i\bar{X} + \bar{X}^2$$

$$= (X_1^2+X_2^2+\dots+X_n^2) - 2\bar{X}\underbrace{(X_1+X_2+\dots+X_n)}_{n\bar{X}} + n\bar{X}^2$$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$= \sum X_i^2 - n\bar{X}^2$$

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

$$= \frac{1}{n}\left[\sum_{i=1}^{n}\left(var(X_i) + (EX_i)^2\right) - n\left(var(\bar{X}) + (E\bar{X})^2\right)\right]$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \rightarrow \text{Sample Variance}$$

Date / /
Page

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

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

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

{ Sample variance in ~~1st class : $\frac{1}{n-1}$ { }~~ so that ~~we get σ^2 here~~ (unbiased) }

So, $\bar{S}^2$ is a biased estimator of the population variance

→ Sample Std. Deviation : { biased estimator of population S.D. }

$S = \sqrt{S^2}$ P.T. : $E[S] \neq \sigma$

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

$$= \sigma^2 - (ES)^2 > 0$$

$\sigma^2 - (ES)^2 > 0 \Rightarrow \sigma^2 \geq (ES)^2 \rightarrow \sigma > ES$

~~$\sigma^2 - (ES)^2 < 0 \rightarrow \sigma^2 < (ES)^2$~~ $\boxed{\text{biased}}$

Example : If $\hat{\Theta}_1$ is an unbiased estimator of Θ & W is a zero mean r.v. Then show that $\hat{\Theta}_2 = \hat{\Theta}_1 + W$ is also an ~~ubi~~ unbiased estimator of Θ.

Sol^n : [Prove : $E[\hat{\Theta}_2] = \Theta$]

$$E[\hat{\Theta}_2] = E[\hat{\Theta}_1] + E[W]$$

$$= \Theta + 0 = \Theta$$

Example : If $\hat{\Theta}_1$ is an estimator of Θ s.t. $E[\hat{\Theta}_1] = a\Theta + b$ where $a \neq 0$. Then show that $\hat{\Theta}_2 = \frac{\hat{\Theta}_1 - b}{a}$ is an unbiased estimator of Θ

$$E[\hat{\Theta}_2] = \frac{1}{a} E[\hat{\Theta}_1] - \frac{b}{a}$$

$$= \frac{1}{a}[a\Theta + b] - \frac{b}{a}$$

$$= \Theta$$

Date / /
Page

Example :- Let $X_1, X_2, \ldots, X_n$ be a random sample from a uniform $(0,\theta)$ distribution where θ is unknown. Define the estimator

Last element of ordered seq. ← $\hat{\theta}_n = \max\{X_1, X_2, \ldots, X_n\}$ ↳ ~~last element of order statistics~~

(a) Find $B(\hat{\Theta}_n)$ & $MSE(\hat{\Theta}_n)$

(b) Is $\hat{\Theta}_n$ consistent estimator of Θ?

$$X_i = \begin{cases} 1/\theta & x \in (0,\theta) \\ 0 & \text{Otherwise} \end{cases}$$

PDF of $\hat{\Theta}_n$:

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

$$\Rightarrow f_X(x) = \begin{cases} 1/\theta & 0 \le x \le \theta \\ 0 & \text{Otherwise} \end{cases}$$

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

$$\Rightarrow f_n(x) = n\left(\frac{1}{\theta}\right)\left(\frac{x}{\theta}\right)^{n-1}$$

$$\Rightarrow f_n(x) = \begin{cases} (n x^{n-1})/\theta^n & 0 \le x \le \theta \\ 0 & \text{Otherwise} \end{cases}$$

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

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

$$= \frac{n}{\theta^n} \left.\frac{x^{n+1}}{n+1}\right|_0^\theta = \boxed{\frac{n\theta}{n+1}}$$

Scanned by CamScanner

Date / /
Page

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

$\therefore MSE(\hat{\Theta}) = var(\hat{\Theta}) + (B(\hat{\Theta}))^2$

$var(\hat{\Theta}_n) = E[\hat{\Theta}_n^2] - (E[\Theta_n])^2$

$= \int_0^{\Theta} x^2 \frac{n}{\Theta^n} x^{n-1} dx - \left(\frac{n\Theta}{n+1}\right)^2$

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

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

$MSE(\hat{\Theta}_n) = \frac{2\Theta^2}{(n+2)(n+1)}$

(b) $\lim_{n\to\infty} MSE(\hat{\Theta}_n) = 0$ $\Rightarrow$ It is consistent (from theorem)

11/4/18

Earlier, we had for specific (mean, var, S.D.). Now, we want estimator for generalised cases

PAGE NO.:
DATE: / /

Maximum Likelihood Estimator : find more generalised estimator.

Definition : Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with pmf / pdf $f(x|\theta)$. The likelihood func^n is defined as :

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

unknown parameter θ — given — ↳ Joint pmf / pdf

where $(x_1, x_2, \ldots, x_n)$ is an observed value of random sample $(X_1, X_2, \ldots X_n)$

Let $\bar{x} = (x_1, x_2, \ldots, x_n)$

Domain of $L(\theta | \bar{x})$ is all admissable values of unknown parameter θ.

Example : let X_1, X_2, X_3, X_4 be a random sample from Bernoulli(p) population. ↳ unknown parameter

For the sample value $(1,0,1,1)$, the likelihood func^n is :

$$L(p | 1,0,1,1) = ?$$

$$f_X(x_i) = \begin{cases} p & x_i = 1 \\ 1-p & x_i = 0 \end{cases}$$

$$L(p | 1,0,1,1) = p(1-p)p \cdot p = p^3(1-p)$$

where $p \in [0,1]$

Example : let X_1, X_2, X_3, X_4 be a random sample from exp(λ) population

For the sample values $(1.23, 3.32, 1.98, 2.12)$:

~~the~~ Likelihood func^n is :

$$L(\lambda | \bar{x}_0) =$$

estimator : func^n of r.v.
estimate : observed value of r.v.

PAGE NO.:
DATE: / /

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{Otherwise} \end{cases}$$

all values we've taken : are +ve

$$L(\lambda | \bar{x}_0) = \lambda e^{-\lambda(1.23)} \cdot \lambda e^{-\lambda(3.32)} \cdot \lambda e^{-\lambda(1.98)} \cdot \lambda e^{-\lambda(2.12)}$$
$$= \lambda^4 e^{-\lambda(8.65)}$$

Domain of L : 0 to ∞
$\lambda \in (0, \infty)$

$$L(\theta | \bar{x}) = \prod_{i=1}^{n}$$

given random sample value : L gives probability of this random ~~variable taking values~~ : $\bar{x}$.

* If higher value of pmf / pdf $\Rightarrow$ Probability will be $\uparrow$.

4 We want to maximize this likelihood func^n wrt θ.

Definition For each sample point $\bar{x} = (x_1, x_2, \ldots, x_n)$, let $\hat{\theta}(\bar{x})$ be a parameter value at which $L(\theta|\bar{x})$ attains its (global) maximum value as a func^n of θ, with $\bar{x}$ held fixed.
$\hat{\theta}(\bar{x})$: Maximum likelihood estimate of θ given $\bar{x}$

A maximum likelihood estimator (MLE) of θ based on random sample $\bar{X} = (X_1, X_2, \ldots, X_n)$ is $\hat{\theta}(\bar{X})$

$6p - 12p^2$
$6p(1-2p^2)$

PAGE NO.:
DATE: / /

Example: Let X_1, X_2, X_3, X_4 be a random sample from Bernoulli(p) population. For which value of p is the probability of the sample values (1,0,1,1) is the largest.

Sol^n

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

polynomial func^n in closed & bounded interval for sure attains extreme values

$$f(p) = p^3 - p^4$$

$$\frac{df}{dp} = 3p^2 - 4p^3 = 0$$

$$\Rightarrow p^2(3-4p) = 0$$

$$\Rightarrow p = 0, \frac{3}{4}$$

solve just for interior point

0: Boundary point (discard)

$\frac{3}{4}$: critical point

p	0	3/4	1
$f(p)$	0	+ve	0

+ve ⇓ Global maxima

$\boxed{p = \frac{3}{4}}$ Ans. → maxm likelihood estimate of parameter p.

↳ We are interested in : maximize $L(\hat{\theta} \mid \bar{x})$

⇔ maximize $\log_e L(\theta \mid \bar{x})$ → log likelihood func^n.

always the (pmf/pdf ≠ 0)

reason : $\ln x$: ↑ func^n on $(0, \infty)$

↳ when exponential func^{n}s are involved : this is much easy

Example: Let X_1, X_2, X_3, X_4 be a random sample from exp(λ) population. Find the MLE of λ for sample values (1.23, 3.32, 1.98, 2.12)

Sol^n

$$L(\lambda \mid \bar{x}) = \lambda^4 e^{-\lambda(8.65)} \qquad \lambda \in (0, \infty)$$

→ can maximize

$f(\lambda) = \lambda^4 e^{-\lambda(8.65)}$, $\lambda \in (0, \infty)$

$\lambda^4 e^{-\lambda(8.65)}$ ↓ continuous func^n

$\lambda \in (0, \infty)$ ↓ neither closed nor bounded interval

⇒ can't say that for sure, it will have a global maxima.

~~$f'(\lambda) =$~~

Let $f(\lambda) = \log_e(\lambda^4 e^{-\lambda(8.65)})$

$= 4 \log_e \lambda - 8.65\lambda$, $\lambda \in (0, \infty)$

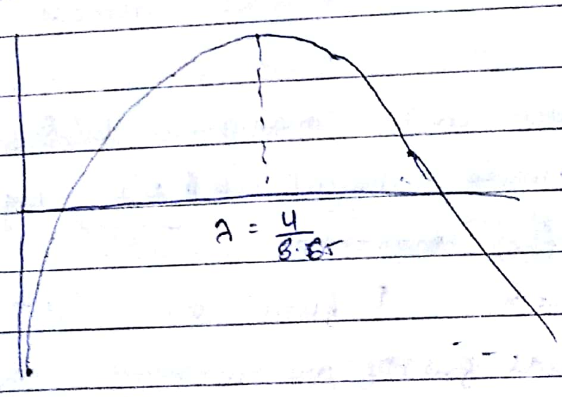
$f'(\lambda) = \dfrac{4}{\lambda} - 8.65 = 0$

⇒ $\lambda = \dfrac{4}{8.65}$: Only critical point

→ To find maxima

$\lim_{\lambda \to 0} f(\lambda) = -\infty$

$\lim_{\lambda \to \infty} f(\lambda) = -\infty$ (exponential growth dominates algebraic growth)



⇒ $\lambda = \dfrac{4}{8.65}$: point of global maxima

* have to ensure global maxima (not ~~min~~ local maxima)

↓

{ 2nd derivative tells about this only, not global maxima

PAGE NO.:
DATE: / /

OR $f'(x)$:

$$\underset{0}{\vphantom{|}} \overset{+}{\longrightarrow} \underset{\frac{40}{3.65}}{|} \overset{-}{\longrightarrow} \underset{\infty}{\vphantom{|}}$$

↳ maxima

13/04/18

Example: Let $X_1, X_2, \dots, X_n$ be a random sample from a geometric (p) distribution. Find the MLE of p.

Solution no exact value is given ⇒ have to find estimator (not estimate)

The likelihood func^n is

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

↳ pmf of geometric (p)
↓
discrete r.v.

$$f(X=k) = p(1-p)^{k-1} \quad \text{for } k = 1, 2, \dots$$

$$\Rightarrow L(p \mid x_1, x_2, \dots, x_n) = \left(p(1-p)^{x_1-1}\right)\left(p(1-p)^{x_2-1}\right)\dots\left(p(1-p)^{x_n-1}\right)$$

$$= p^n (1-p)^{\sum_{i=1}^{n} x_i - n}$$

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

$$L(p \mid \bar{x}) = p^n (1-p)^{y-n}$$

$\Rightarrow p \in (0,1)$

$f = 0$ is not permissible identically in geometric r.v.
$\Rightarrow p \neq 0, 1$

~~$\log_e(L)$~~

$$f(p) = \log_e\left(p^n (1-p)^{y-n}\right)$$

$$= n \log_e p + (y-n)\log_e(1-p)$$

$$\hookrightarrow f'(p) = \frac{n}{p} + (y-n)\left[\frac{-1}{1-p}\right] = 0$$

$$\frac{n - \cancel{np} + \cancel{np} - yp}{p(1-p)} = 0$$

$$n - yp = 0 \Rightarrow \boxed{p = n/y}$$

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

$\Rightarrow \frac{(1-p)+p}{p} = \frac{(y-n)+n}{n}$

$\Rightarrow \boxed{p = \frac{n}{y}}$: Only critical point of $f(p)$

→ To find maxima :

~~$\lim_{p \to 0} f(p)$~~

Interval $(0, n/y)$:-

↳ $0 < p < \frac{n}{y} \Rightarrow \boxed{\frac{n}{p} > y}$ & $-p > -\frac{n}{y}$

$\boxed{y > 1}$

$f'(p) = \underbrace{\frac{n}{p}}_{>y} - \frac{y-n}{1-p}$ $\qquad 1-p > 1-\frac{n}{y} = \frac{y-n}{y}$

$\Rightarrow y - \left(\frac{y-n}{1-p}\right) > 0$ or $\Rightarrow \boxed{y > \frac{y-n}{1-p}}$

$\Rightarrow f'(p) > 0$

↳ $1 > p > \frac{n}{y} \Rightarrow f'(p) < 0$

⇒ values at 0 & 1 can't be maxima

⇒ $p = \frac{n}{y}$ is global maximum of $f(p)$ over $(0,1)$

$\hat{p} = \frac{n}{\sum_{i=1}^{n} x_i}$ is the mle of p given $\bar{x}$

Hence, MLE $\hat{p}(\bar{x}) = \frac{n}{\sum_{i=1}^{n} x_i}$

PAGE NO.:
DATE: / /

know p, σ but not n

Example: Let $X_1, X_2, \ldots, X_n$ be a random sample from a binomial(k, p) population where p is known & k is unknown. Find the MLE of parameter k.

Solution:

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

$$f(X=j) = {}^{k}C_j \, p^j (1-p)^{k-j} \quad \text{for } j = 0, 1, 2, \ldots$$

$$= \left[\binom{k}{x_1} p^{x_1} (1-p)^{k-x_1}\right]\left[\binom{k}{x_2} p^{x_2} (1-p)^{k-x_2}\right] \cdots \left[\binom{k}{x_n} p^{x_n} (1-p)^{k-x_n}\right]$$

$$= \binom{k}{x_1}\binom{k}{x_2}\binom{k}{x_3} \cdots \binom{k}{x_n} p^{\sum_{i=1}^{n} x_i} (1-p)^{nk - \sum_{i=1}^{n} x_i}$$

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

$$f(k) = \binom{k}{x_1}\binom{k}{x_2} \cdots \binom{k}{x_n} p^{y} (1-p)^{nk-y}$$

admissible range of $k \in \mathbb{N}$
$\{1, 2, \ldots\}$
can't talk about derivatives here.

can treat k as continuous $[1, \infty)$. If $k \in \mathbb{Z}$ ✓
But if $k \notin \mathbb{Z}$ ✗

can guess nearest to that fraction : ~~But Cancy~~ ✗

Issues with this approach

represent no. of heads obtained in 1st toss

* If $k < \max(x_1, x_2, \ldots x_n)$, then
$L(k \mid \bar{x}) = 0$ not possible

* It is possible that all x_i, $i = 1, 2, \ldots, n = 0$ (nothing mentioned in question)

PAGE NO.:
DATE: / /

We assume that all $x_i \geq 1$
$\Rightarrow \max\{x_i\} \geq 1$

If $k \geq \max\{x_1, x_2, \ldots, x_n\} = \max_{1 \leq i \leq n}(x_i)$

Then the MLE $\hat{k}$ of $L(k \mid \bar{x}, p)$ satisfies:

$\frac{L(\hat{k})}{L(k)} \geq 1$

$$\frac{L(\hat{k} \mid \bar{x}, p)}{L(\hat{k}-1 \mid \bar{x}, p)} \geq 1 \quad \& \quad \frac{L(\hat{k} \mid x, p)}{L(\hat{k}+1 \mid x, p)} \geq 1$$

$\to$ ① $\qquad \downarrow$ ②

We have to find such a $\hat{k}$ $\qquad \frac{(k+1 \quad)}{(k \quad)} < 1$

① $$\frac{L(\hat{k} \mid \bar{x}, p)}{L(\hat{k}-1 \mid \bar{x}, p)} = \frac{\frac{k!}{x_1!(k-x_1)!} \times \frac{k!}{x_2!(k-x_2)!} \times \cdots \times \frac{k!}{x_n!(k-x_n)!} \quad (1-p)^{nk-y} \cdot \cancel{p^y}}{\frac{(k-1)!}{x_1!(k-x_1-1)!} \times \frac{(k-1)!}{x_2!(k-x_2-1)!} \cdots \quad (1-p)^{n(k-1)-y} \cdot \cancel{p^y}}$$

$$= \frac{(k)^n (1-p)^n}{(k-x_1)(k-x_2)(k-x_3)\cdots(k-x_n)} = \frac{(k(1-p))^n}{\prod_{i=1}^{n}(k-x_i)}$$

Thus, the condn for a maximum is:

$$[k(1-p)]^n \geq \prod_{i=1}^{n}(k-x_i) \qquad \text{③}$$

$$\& \quad [(k+1)(1-p)]^n < \prod_{i=1}^{n}((k+1)-x_i) \qquad \to \text{④}$$

↳ Here also, $p \neq 0, 1$ (f will become identically 0)
$\Rightarrow p \in (0,1)$

divide ③ by k^n

$$\Rightarrow (1-p)^n \geq \prod_{i=1}^{n}\left(\frac{k-x_i}{k}\right) = \prod_{i=1}^{n}\left(1-\frac{x_i}{k}\right)$$

set $\frac{1}{k} = z$

We want to solve:

$$\underbrace{(1-p)^n}_{\text{given}} = \prod_{i=1}^{n}\left(1-\underbrace{x_i}_{\text{given}}\underbrace{z}_{\text{unknown}}\right) \qquad 0 < z \leq \frac{1}{\max x_i} < 1$$

PAGE NO.:
DATE: / /

$f(z) = \overset{\to +ve}{(1-x_1 z)}(1-x_2 z) \ldots (1-x_n z)$

$f'(z) < 0 \quad \forall \; z \in \left[0, \frac{1}{\max x_i}\right)$ ↳ if $z = \frac{1}{\max x_i}$, $1 - z x_i = 0$ ⇒ $f(z) = 0$ ✗

⇒ $f(z)$ is cont. & strictly ↓ing funcⁿ.

Intermediate value Theorem:

RHS : ↓ing funcⁿ & continuous funcⁿ

LHS : value $\in [0,1)$

By IVT, ∃ $\hat{z} \in \left[0, \frac{1}{\max x_i}\right]$ s.t. eqⁿ ④ has a solⁿ

~~Define $\hat{k} = \frac{1}{\hat{z}}$~~ we don't know $1/\hat{z}$ is integer or not

Let $\hat{k} = \left\lfloor \frac{1}{\hat{z}} \right\rfloor$ → largest integer value

16/4/18

→ If $x_i = 0 \;\; \forall \; i = 1, 2, \ldots, n$ then

$L(k \mid 0,0,0,\ldots,0,p) = \prod_{i=1}^{n} \binom{k}{0} p^0 (1-p)^k$

$f(k) = (1-p)^{nk} = [(1-p)^n]^k$

$p \neq 0,1 \quad p \in (0,1)$

⇒ $1-p \in (0,1)$

⇒ $(1-p)^n \in (0,1) \quad \forall \; n$

* if $a \in (0,1)$, then $a^n \cancel{<} a \quad \forall \; n \geq 2$

Here, $k = 1$ is the maximizer

(have dealt with this case separately because won't have given correct answer using original ~~ea~~ one)

PAGE NO.:
DATE: / /

A global maximizer $\hat{k}$ of $L(k)$ must satisfy

$$\frac{L(\hat{k})}{L(k)} \geq 1 \quad \forall \; k \in [\max_i x_i, k]$$

$$\frac{L(k)}{L(\hat{k})} \leq 1 \quad \forall \quad k > \hat{k}$$

Claim, The eqn

$$(1-p)^n = \prod_{i=1}^{n} (1 - x_i z) \quad \text{—— (1)} \qquad \text{(From last class)}$$

has a unique soln in the interval $\left(0, \frac{1}{\max_i x_i}\right)$

Proof of claim:

$$f(z) = (1-x_1 z)(1-x_2 z) \ldots \ldots (1-x_n z)$$

$$f'(z) = -x_1 [(1-x_2 z) \ldots (1-x_n z)]$$
$$- x_2 [(1-x_1 z) \ldots (1-x_n z)]$$
$$\vdots$$
$$- x_n [(1-x_1 z) \ldots (1-x_{n-1} z)]$$

$$z < \frac{1}{\max_i x_i} \Rightarrow \max_i x_i z < 1$$

$$\Rightarrow 1 - x_i z < 1 \quad \forall \; i = 1, 2, \ldots, n$$

$$\Rightarrow f'(z) < 0 \quad \forall \quad z \in \left(0, \frac{1}{\max_i x_i}\right)$$

$\Rightarrow$ f is strictly $\downarrow$ing func

$f(0) = 1$ ← LHS

$f\left(\frac{1}{\max x_i}\right) = 0$ ← RHS

cont. : $\exists$ a soln
$\downarrow$ng ∴ unique soln

By IVT, $\exists$ a unique soln $\hat{z}$ in interval $\left(0, \frac{1}{\max_i x_i}\right)$

PAGE NO.:
DATE: / /

(4')

Let $\hat{k} = \left[\frac{1}{\hat{z}} \right]$ (largest integer $\le 1/\hat{z}$)

$\therefore \hat{k} \le \frac{1}{\hat{z}}$ so $\hat{z} \le \frac{1}{\hat{k}}$

$f(\hat{z}) \ge f\left(\frac{1}{\hat{k}}\right) \longrightarrow \downarrow^{ng}$ func^n

ie $\underset{\substack{\downarrow \\ \text{~~SEP~~} = f(\hat{z})}}{(1-p)^n} \ge \prod_{i=1}^{n} \left(1 - \frac{x_i}{\hat{k}}\right)$ —(2)

also, $\hat{k}+1 > \frac{1}{\hat{z}}$ so $\hat{z} > \frac{1}{\hat{k}+1}$ (from (4'))

$\Rightarrow f(\hat{z}) < f\left(\frac{1}{\hat{k}+1}\right)$

$\Rightarrow (1-p)^n < \prod_{i=1}^{n} \left(1 - \frac{x_i}{\hat{k}+1}\right)$ —(3)

Claim: $\hat{k}$ is a global maximizer of $L(k|\bar{x})$

Proof of claim:

Recall:

$$\frac{L(\hat{k})}{L(k-1)} = \frac{[k(1-p)]^n}{\prod_{i=1}^{n}(k-x_i)} \quad \text{—(5)}$$

1.) If $k < \hat{k}$, then $\frac{1}{k} > \frac{1}{\hat{k}}$

$\Rightarrow f\left(\frac{1}{k}\right) < f\left(\frac{1}{\hat{k}}\right)$

$\Rightarrow \prod_{i=1}^{n}\left(1 - \frac{x_i}{k}\right) < \prod_{i=1}^{n}\left(1 - \frac{x_i}{\hat{k}}\right) \le (1-p)^n$ Using (—(2))

$\Rightarrow \frac{[k(1-p)]^n}{\prod_{i=1}^{n}(k-x_i)} \ge 1 \Rightarrow L(\hat{k}) > L(k)$

Therefore, $L(k)$ is $\uparrow^{ng}$ in $[\max_{x_i}, \dots, k]$

Scanned by CamScanner

PAGE NO.:
DATE: / /

2.) If $k > \hat{k}$ $\Rightarrow$ $k \geq \hat{k}+1$

$\Rightarrow$ $\frac{1}{k} \leq \frac{1}{\hat{k}+1}$

$\Rightarrow$ $f\left(\frac{1}{k}\right) \geq f\left(\frac{1}{\hat{k}+1}\right) > (1-p)^n$

$\Rightarrow$ $\prod_{i=1}^{n} \left(1 - \frac{x_i}{k}\right) > (1-p)^n$

$\Rightarrow$ $\dfrac{k(1-p)^n}{\prod_{i=1}^{n}(k - x_i)} < 1$

Hence, $L(k)$ is a $\downarrow$ing func^n in $\{\hat{k}, \hat{k}+1, \hat{k}+2, \dots\}$

From 1) & 2), $\hat{k}$ is maximum (global).

- We still don't know exact value of $\hat{k}$ (IVT: $\exists$ a root, not the value of root). It just gives assurance that $\hat{k}$ exists.

Example: Let $X_1, X_2, \dots, X_n$ be a random sample from uniform $(0, \theta)$ distribution where θ is unknown. Find the MLE for θ.

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

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

$$L(\theta | x_1, x_2, \dots x_n) = \begin{cases} \frac{1}{\theta^n} & 0 < x_i < \theta \quad i = 1, 2, \dots n \\ 0 & \text{Otherwise} \end{cases}$$

Scanned by CamScanner

PAGE NO.:

DATE: / /

all $x_i < \Theta$

$\Rightarrow \max_i x_i < \Theta$ $\qquad \frac{1}{\Theta^n}$: $\downarrow^{ng}$ func^n

$\Theta \in (0, \infty)$

$L(\Theta)$ is max^m at $\hat{\Theta} = \max_i x_i$

$MLE(\hat{\Theta}) = \max_i (x_i)$

17/04/18

Maximum Likelihood Estimation of a vector parameter

let $\bar{\Theta} = (\Theta_1, \Theta_2, \ldots, \Theta_k)$

$\bar{\Theta} \in \mathbb{R}^k$

Let $X_1, X_2, \ldots, X_n$ be a random sample from the population $f(x|\bar{\Theta})$ Define:

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

$\hat{\bar{\Theta}}$ which maximizes $L(\bar{\Theta})$ is called a MLE

can use multi-variable calculus. But in practice, it takes a lot of time. (Partial derivative must exist & be cont, ...)

can give info. about local maxima only, we need global maxim.

Example : Let $X_1, X_2, \ldots, X_n$ be iid $N(\mu, \sigma^2)$ with both μ & σ^2 being unknown. The likelihood func^n $L(\mu, \sigma^2 | \bar{x}) = ?$

$$L(\mu, \sigma^2 | \bar{x}) = \prod_{i=1}^{n} f(x_i | \mu, \sigma^2)$$

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

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

PAGE NO.:
DATE: / /

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

{ We need $L(\sigma^2) \Rightarrow$ use σ^2 }

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

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

Let $\sigma^2 = \theta$

$$f(\mu,\theta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\theta - \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\theta} \quad \text{—} ④$$

$$(\mu,\theta) \in \mathbb{R} \times (0,\infty)$$

Partial derivatives:

$$\frac{\partial f}{\partial \mu} = -\sum_{i=1}^{n}\frac{2(x_i-\mu)(-1)}{2\theta} = \sum_{i=1}^{n}\frac{(x_i-\mu)}{\theta} = 0 \quad ①$$

$$\frac{\partial f}{\partial \theta} = -\frac{n}{2}\left(\frac{1}{\theta}\right) + \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\theta^2} = 0 \quad \text{—} ②$$

From ① :

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

$$\Rightarrow \sum_{i=1}^{n} x_i = n\mu$$

$$\Rightarrow \mu = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x} \text{ (let)}$$

From ② :

$$-\frac{n}{\cancel{2}}\left(\frac{1}{\theta}\right) + \sum_{i=1}^{n}\frac{(x_i - \bar{x})}{\theta} = 0$$

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

PAGE NO.:
DATE: / /

This is the only critical point. Question is whether it is global maxima or not

---

Lemma: Let $x_1, x_2, \ldots, x_n$ be any real no & $\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}$. Then, ← (1)

$$\min_{a \in \mathbb{R}} \left\{ \sum_{i=1}^{n} (x_i - a)^2 \right\} = \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Proof:

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

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

$$= \sum_{i=1}^{n} (x_i - \bar{x})^2 + n(\bar{x} - a)^2 + 2(\bar{x} - a) \cancel{\sum_{i=1}^{n} (x_i - \bar{x})}^{0}$$ (from (1))

$$= \underbrace{\sum_{i=1}^{n} (x_i - \bar{x})^2}_{\geq 0} + \underbrace{n(\bar{x} - a)^2}_{\geq 0}$$

To find min: $a = \bar{x}$

RHS is min^m when $a = \bar{x}$

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

From (4):

maximizer of f $\equiv$ minimizer of $-f = g$

$$g(\mu, \theta) = n \log \theta + \frac{1}{\theta} \sum_{i=1}^{n} (x_i - \mu)^2$$

{ won't affect removed const values from f }

We're to minimize g here.

PAGE NO.:
DATE: / /

From the lemma, $\sum_{i=1}^{n} (x_i - \mu)^2$ will be minimized when $\mu = \bar{x}$ (regardless of value of θ)

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

$h(\theta) = n \log \theta + \frac{c^2}{\theta}$: have to minimize this (have to show it has min. value at $\theta = \frac{c^2}{n}$ computed earlier)

$h'(\theta) = \frac{n}{\theta} - \frac{c^2}{\theta^2} = 0$

$n\theta = c^2$

$\Rightarrow \theta = \frac{c^2}{n}$

if $n > \frac{c^2}{\theta}$ or $\theta > \frac{c^2}{n}$ $\Rightarrow$ h is $\uparrow^{ng}$ func^n $(h' > 0)$

$\theta < \frac{c^2}{n}$ $\Rightarrow$ $h' < 0$ $\Rightarrow$ h is $\downarrow^{ng}$ func^n

$h' = 0 \Rightarrow \theta = \frac{c^2}{n}$

$\downarrow^{ng}$ $\uparrow^{ng}$ +ve

0 $\theta = \frac{c^2}{n}$ ∞

global minima (as found earlier)

$$\hat{\mu}(\bar{x}) = \left(\sum_{i=1}^{n} x_i\right) \frac{1}{n}$$

Not sample variance $\leftarrow$ $\hat{\sigma}^2(\bar{x}) = \frac{1}{n} \sum_{i=1}^{n} (\bar{x} - x_i)^2$: biased (discussed earlier)

→ Maximum likelihood Estimator is not always unbiased

Scanned by CamScanner

* In point estimator: we had more accuracy (exact pt.). Now, we are calculating interval : more ambiguous. Then, why to study it? see example 2

PAGE NO.:
DATE: / /

Interval Estimation

An interval estimate of a real-value parameter θ is any pair of func^ns $L(x_1, x_2, \ldots, x_n)$ & $U(x_1, x_2, \ldots x_n)$ of a sample that satisfies

$$L(\bar{x}) \leq U(\bar{x}) \quad \forall \ \bar{x} \in \text{Range of random sample } \bar{X} = (X_1, X_2, \ldots, X_n)$$

If $\bar{X} = \bar{x}$ is observed, the inference $L(\bar{x}) \leq \theta \leq U(\bar{x})$ is made. The random interval $[L(\bar{X}), U(\bar{X})]$ is called an interval estimator

example 1: Let X_1, X_2, X_3, X_4 be a random sample from $N(\mu, 1)$ distribution. Then $[\bar{X}-1, \bar{X}+1]$ is an interval estimator for μ.

$P(\bar{X} = \mu) = ?$ → pt. estimator

$$\bar{X} = \frac{X_1 + X_2 + X_3 + X_4}{4} \qquad \mu: \text{unknown \& fixed qty.}$$

unbiased estimator of mean pop^n

$\frac{X_1}{4}, \frac{X_2}{4}, \frac{X_3}{4}, \frac{X_4}{4}$: also $N\left(\frac{\mu}{4}, \frac{1}{16}\right)$ independent

⇒ $\bar{X}$ is also Normal r.v. $\sim N\left(\mu, \frac{1}{4}\right)$ add^n

⇒ $P(\bar{X} = \mu) = 0$

↓

$\bar{X} \sim N$ & has pdf ⇒ $P(\bar{X} = a \text{ pt.}) = 0$

↳ But if we calculate interval, $P(\quad) \neq 0$.

$P(\mu \in [\bar{x}-1, \bar{x}+1])$

$P(\bar{x} - 1 \leq \mu \leq \bar{x} + 1)$ (make it $\frac{x-\mu}{\sigma}$)

$= P\left(-1 \leq \bar{X} - \mu \leq 1\right)$

$= P\left(-2 \leq \frac{\bar{X}-\mu}{\sqrt{1/4}} \leq 2\right)$

$= N(2) - N(-2) = N(2) - (1 - N(2)) = 2N(2) - 1$
$= 0.944$

⇒ If we are in this interval, we're 95% sure, the true value of unknown parameter will lie in this interval.

→ Although we've lost accuracy, we have gained this above reason to calculate the interval estimate.

18-4-18

Definition : We say that an interval estimator $[L(x_1, x_2, \ldots x_n), U(x_1, x_2, \ldots x_n)]$ of a parameter has the confidence level "$1-\alpha$" (α doesn't depend on θ) if

$$P\left[\theta \in [L(x_1, \ldots, x_n), U(x_1, x_2, \ldots x_n)]\right] = 1-\alpha \quad \text{— ①}$$

→ In last example,

$$\bar{x} = \text{sample mean} = \frac{x_1 + x_2 + x_3 + x_4}{4}$$

$$P\left[\mu \in [\bar{x}-1, \bar{x}+1]\right] = 0.944 = 1-\alpha$$

→ We are 94.4 % sure that true value of mean will lie in this interval. There's an uncertainity also here.

Relative Freq. Interpretation : Take a sample & calculate its mean, 95% of time mean will lie in the given interval.

→ Std. values of $1-\alpha$: 90%, 95%, 99%.

→ From ①, θ (unknown parameter) is just fixed const., not r.v. ~~Moreov~~. Moreover, the interval mentioned is random.

PAGE NO.:
DATE: / /

① → $P[\theta \le U(x_1, \dots x_n)$ and $\theta \ge L(x_1, \dots, x_n)]$

const. ↓ random ↓ random

⇕

$P[L(x_1, \dots x_n) \le \theta \le U(x_1, \dots, x_n)]$

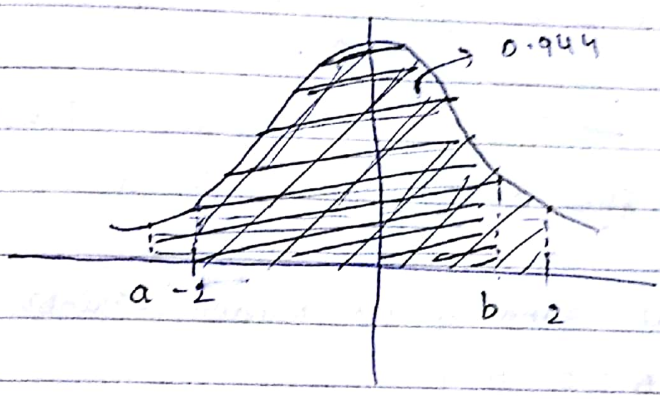
↳ The interval estimator ① $[\bar{X}-1, \bar{X}+1]$ of parameter μ has confidence level 0.944. Is this interval estimator unique?

$\boxed{\text{Not unique}}$

(No other interval has value 0.944 for same μ)

We got this 0.944 from:

$P(-2 \le Z \le 2) = 0.944$ $\qquad Z = \dfrac{\bar{X}-\mu}{\sqrt{1/4}}$



can get $[a, b]$ s.t. value = 0.944

↳ If we check

$P(-1.68 < Z < 2.7) = 0.95$ ≈ getting app. same value ②

$P(-1.68 < 2(\bar{X}-\mu) < 2.7)$

$P\left(\frac{-1.68}{2} < \bar{X}-\mu < \frac{2.7}{2}\right)$

Ques Among ① & ②, which is better? (Confidence is same in both cases)

length of ① : 2 $\qquad$ length of ② : 4.38

If length is less ⇒ more precision

* Confidence level with min. length are preferred

Scanned by CamScanner

PAGE NO.:
DATE: / /

→ The way we used to calculate Interval Estimator was very generalised. Now, we'll learn :

Pivotal Method

Definition : let $X_1, X_2, \ldots, X_n$ be a random sample from the population $f(\ |\theta)$. let $Q = f(X_1, X_2, \ldots, X_n, \theta)$. If distribution of Q doesn't depend on θ ~~and any other parameter of~~, then Q is called a pivot.
r.v.

Example : let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\mu, 9)$. Then,

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

$$= Q_1(X_1, \ldots, X_n, \mu) \quad : \text{Pivot. Prove it.}$$

→ We have to find distribution of $\bar{X} - \mu$

$\bar{X} - \mu$: also ~~rand~~ Normal Random Variable

$$\sim N\left(0, \frac{9}{n}\right)$$

$$E[\bar{X}-\mu] = 0 \qquad [= E\bar{X} - \mu]$$

$$Var(\bar{X}-\mu) = \frac{1}{n^2}\left[Var(X_1) + Var(X_2) \ldots + Var(X_n)\right]$$

$$= \frac{1}{n^2}(9 \times n) = \frac{9}{n}$$

$N\left(0, \frac{9}{n}\right)$: Nowhere μ has come into the picture, → in distribution of a
Hence, Q is a pivot.

↳ Std. Normal R.V. $\frac{X-\mu}{\sigma} \sim N(0,1)$ $\qquad \bar{X} \sim N\left(\mu, \frac{9}{n}\right)$

Here, $Q_2 = \frac{\bar{X} - \mu}{\sqrt{\frac{9}{n}}} \sim N(0,1)$

↳ another pivot (distrib^n don't depend on μ)

Scanned by CamScanner

→ Mean of exponential r.v. $= \frac{1}{\lambda}$

PAGE NO.:
DATE: / /

$$Q_3 = \frac{\bar{X}}{\mu} \sim N\left(1, \frac{9}{\mu^2 n}\right)$$

μ/μ ($E\bar{x}$) ; $\frac{1}{\mu^2}$ (var $\bar{x}$)

: distribution depends on $\mu \Rightarrow$ not a pivot

Example Suppose that we're to obtain a single observation Y from an exponential distribution with mean θ. Use Y to form a confidence interval for θ with confidence level $= 0.90$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad \Rightarrow E[X] = \frac{1}{\lambda}$$

we'll use this here

$$\left(f_X(x) = \begin{cases} \frac{1}{\lambda} e^{-x/\lambda} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad \Rightarrow E[X] = \lambda \right.$$

Let : $Q = \frac{Y}{\theta}$; [1st find cdf, diff. to get pdf]

cdf : $P\left[\frac{Y}{\theta} \leq y\right] = P[Y \leq \theta y]$

$$= \int_0^{\theta y} \frac{1}{\theta} e^{-x/\theta} \, dx$$

$$= \frac{1}{\theta}\left[e^{-x/\theta}\right] \times \frac{-\theta}{1} \Big|_0^{\theta y}$$

$$= \left[1 - e^{-\theta y/\theta}\right] = 1 - e^{-y}$$

⇓

pdf $(Q) = e^{-y}$

OR $x = u\theta$

$$= \int_0^y \frac{1}{\theta} e^{-u(\theta/\theta)} \, du \, (\theta)$$

$$= \int_0^y e^{-u} \, du \Rightarrow \frac{Y}{\theta} \sim \exp(1)$$

⇒ Q is pivot

Lecture :- 25/4/18

PAGE NO. :
DATE : / /

To find no. a & b s.t.

$$P(a \le \theta \le b) = 0.90$$

$$= \cancel{P(a\theta \le y \le b\theta) = 0.90}$$

$$= P(\theta \le b) - \cancel{P(\theta \le a) = 0.90}$$

$$= \int_0$$

$$= P(\{\theta < a\} \cup \{\theta > b\}) = 0.10$$

(let) → easy to use symmetric

$P(\theta < a) = 0.05 \Rightarrow \int_0^a e^{-u}\,du = 0.05$

$P(\theta > b) = 0.05 \Rightarrow \int_b^\infty e^{-u}\,du = 0.05$

$\Rightarrow 1 - e^{-a} = 0.05$ & $e^{-b} = 0.05$

$a = 0.051$ $b = 2.996$

$$P\left(0.051 \le \frac{Y}{\theta} \le 2.996\right) = 0.90$$

↳ we need to find for θ (Confidence Interval)

$$P\left(\frac{Y}{2.996} \le \theta \le \frac{Y}{0.51}\right)$$

$$\left[\frac{Y}{2.996}, \frac{Y}{0.51}\right] \rightarrow \text{Interval}$$

→ will be supplied

① guess pivot : verify it is pivot

② choose a & b : we've taken symmetry — smaller length is preferred.

(other possible values of a & b also)

which to choose : where length of interval is min

③ Find interval for unknown parameter

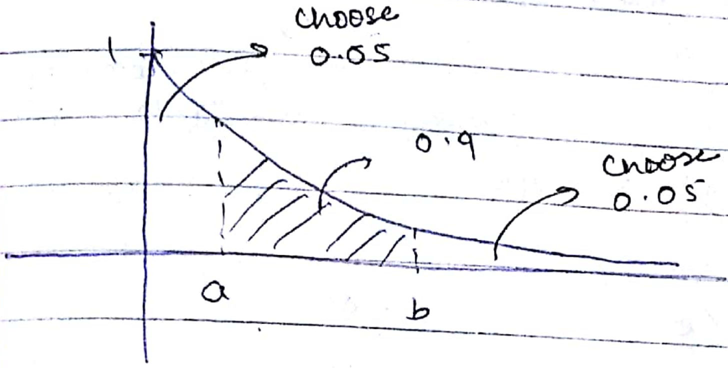
Scanned by CamScanner

27/01/18

PAGE NO.:

DATE: / /

→ $f_\Theta(u) = \begin{cases} e^{-u} & u \geq 0 \\ 0 & \text{Otherwise} \end{cases}$



{ can choose other a & b's }

If Y = 0.75 is observed, accordingly, the confidence interval can be estimated (observed)

with 90% confidence, the true value of Θ lies in

$\left[\frac{0.75}{2.996}, \frac{0.75}{0.51} \right]$

↳ $[L(X), U(X)]$ ← Confidence Interval.

- If $L(X) = -\infty$ → $(-\infty, U(X)]$
- If $U(X) = \infty$ ⇒ $[L(X), \infty)$
- both $L(X) = -\infty$ $U(X) = \infty$ ⇒ giving no. info. (Θ lies on ℝ : we already know it)

↳ We might also have $(L(X), U(X))$ as Interval.

Scanned by CamScanner

PAGE NO. :
DATE : / /

Example Suppose that we take a sample of size 1 from uniform distribution defined on the interval $[0, \theta]$ where θ is unknown. Find a 95% lower confidence bound for θ.

↓
no upper bound

Solution

$\rightarrow U(X) = \infty$

To find $L(X)$ s.t

$$P(\theta \geq L(X)) = 0.95$$

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

Take $U = \dfrac{x}{\theta}$ (pivot) $\quad \theta > 0$

$$\text{cdf} = P(U \leq y) = P(x \leq \theta y)$$

$$= \int_0^{\theta y} \frac{1}{\theta}\, dx \qquad [\theta y < \theta] \quad \boxed{0 < y < 1}$$

$$= y \rightarrow \text{pdf} = 1$$

OR

$x = \theta u$

$$= \int_0^{y} \frac{1}{\theta}\, \theta \, du$$

$\rightarrow$ pdf $(u) = 1$

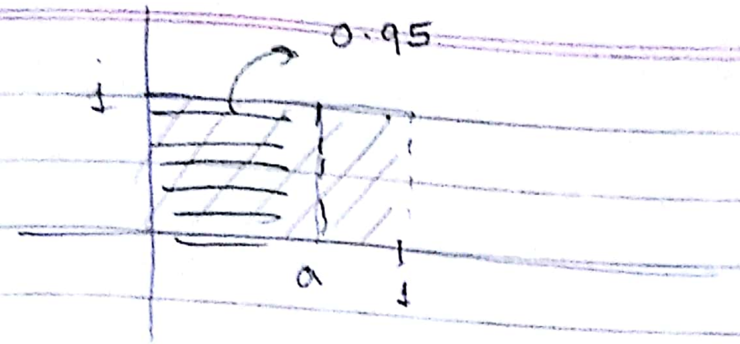
$$f_U(y) = \begin{cases} 0 & ; \ y < 0 \\ 1 & 1 > y \geq 0 \\ 0 & y \geq 1 \end{cases}$$

$\rightarrow U \sim$ uniform $(0, 1)$

$\rightarrow U$ is pivot

Scanned by CamScanner

PAGE NO.:
DATE: / /



$$\cancel{P\left(U = \frac{X}{\theta} \geq a\right)} =$$

$$P\left(\theta \geq L(X)\right) = 0.95$$

$$P\left(\overbrace{\frac{X}{\theta}}^{u} \leq a\right) = 0.95$$

$$= \int_0^a 1\,du = \underline{a = 0.95}$$

$$\Rightarrow P\left(\frac{X}{\theta} \leq 0.95\right) = 0.95$$

$$\Rightarrow P\left(\theta \geq \underbrace{\frac{X}{0.95}}_{\text{Lower bound}}\right) = 0.95$$

↳ This technique is not applicable in all situations, but applicable in most of situations.