STATISTICS - II

problems:

avg of $\mathbb{P}_1$ is property of population

i) calculate avg-height of all 18-years old boys (Jan 2025) in India.

proportion of $\mathbb{P}_2$ is property we're interested in

ii) calculate the proportion of 2 children families in India (Jan 2025).

iii) calculate the expectation of $N_{[-2,3]}(0,1)$ ← average of $\mathbb{P}_3$

$$\mathbb{E}(x) = \int x f_x(x)\, dx$$

$$f_x(x) = \frac{\phi(x)\, \mathbb{I}_{[-2,3]}(x)}{\Phi(3) - \Phi(-2)}$$

$$\int_{-2}^{3} \frac{\cancel{x}\, \phi(x)\, dx}{\Phi(3) - \Phi(-2)} = \frac{1}{\Phi(3) - \Phi(-2)} \int_{-2}^{3} x \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

Population: a collection of objects / measurements we are interested in.

i) $\mathbb{P}_1 = \{$ collection of heights of all 18-years old boys in India (Jan 2025) $\}$: continous popln ?

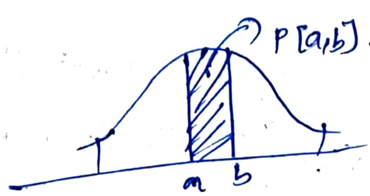
ii) $\mathbb{P}_2 = \{$ collection of no. of children in India (Jan 2025) $\}$: a discrete variable

$= \{0, 1, 0, 0, 1, 2, \dots\}$

Fix $a < b$

let $P_{(a,b)}$ be the proportion of boys in $\mathbb{P}_1$ whose heights are between $(a, b]$ → proportion is fixed in every particular interval.

iii) $\mathbb{P}_3 > N_{[-2,3]}(0,1)$



$P[a,b]$ = proportion of $\mathbb{P}_3$ taking values in $(a, b]$

$$= \int_a^b \frac{\phi(x)\, dx}{\Phi(3) - \Phi(-2)}$$

$$= F(b) - F(a)$$

Sample: subset of the populatn,

$\Downarrow$ (independence)

representative of the populatn.

$x_1, x_2, \ldots x_n$

indep indentically distributed.

$$\boxed{\mathbb{P}[X_1 \in (a,b]] = P_{(a,b]} = \int_a^b f_X(x)\,dx}$$

$\hookrightarrow$ characterizes the population

i) $\overline{X}_n$ ii) $f_n = \dfrac{\#\{2 \text{ children in } x_1, x_2, \ldots x_n\}}{n}$ iii) $\overline{X}_n$

$\mathbb{P}_2 = \{0, 0, 1, 5, 2, \ldots\}$

ii) $f_n = \dfrac{\#\{2s \text{ in } x_1, x_2, \ldots x_n\}}{n} \xrightarrow[(n\to\infty)]{} \mathbb{P}(X=2) =$ proportion of 2's in $\mathbb{P}_2$

0 as prop P_0

1 as prop P_1

$\vdots$

$$X_1 = \begin{cases} 0 & \text{with prob } P_0 \\ 1 & \text{with prob } P_1 \\ \vdots \end{cases}$$

Statistical problem: You have to take samples,

$x_1, x_2, \ldots x_n$

$\mathbb{P} = N(0,1)$ $\quad x_1 \in \mathbb{R}$

$\mathbb{P}(x_1 \in (a,b]) = P(a,b]$

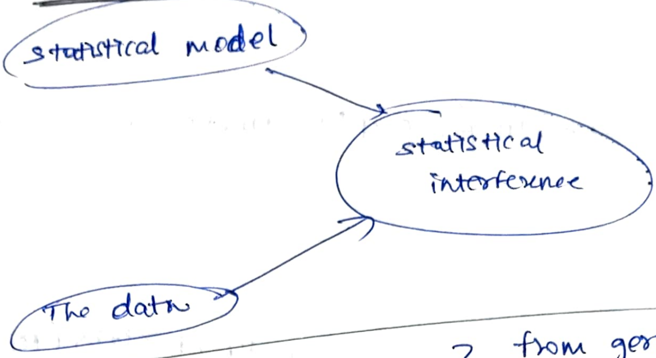
$\underbrace{x_1, x_2, \ldots x_n}_{\text{independent}}$ are identically distributed.

independent $\Rightarrow$ [one x_i doesn't effect the other one]

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

After taking the samples. we will infer about the population from the sample.

→ Statistical Interference:



---

x_i, $i = 1, 2, \dots 15$ } → from germination of seeds example.

(*) $P(x_i = 1) = \theta$
$P(x_i = 0) = 1 - \theta$ } $i = 1, 2, \dots 15$

$\boxed{x_i \overset{iid}{\sim} \text{Ber}(\theta), \quad i = 1, \dots 15}$ — (*) → statistical model.

↑ parameter

The structure which was given by us by using the Example

Ex-1)

$T = \sum_{i=1}^{15} x_i \sim \text{Bin}(15, \theta)$

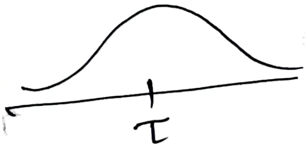
$P(T = 10) = \binom{15}{10} \theta^{10} (1-\theta)^{15-10}$

$P\left(\left(\frac{2}{3} - 0.1, \frac{2}{3} + 0.1\right) \ni \theta\right)$?

Ex-2)

$g = \frac{4\pi^2 l}{T^2}$

$\boxed{T \sim N(\tau, \sigma^2)}$ → parameters



$T_1, \dots T_{10}$

$T_i \overset{iid}{\sim} N(\tau, \sigma^2)$

$\frac{1}{n} \sum_{i=1}^{10} x_i = \bar{T}_{10} \sim N(\tau, \sigma^2/10)$

1] Model

2] Parameters [unknown]

3] By using the sample we will estimate the parameters

Parametric inference

$\sum_{i=1}^{n} T_i^2$ → statistics (↳ complete function of the given data)

Summary:

1] We are interested in some property of class (population)

2] For any $b \in \mathbb{R}$ if we know the proportion of measurements, $P(-\infty, b]$, (for any $b \in \mathbb{R}$) in the population falling below (or equal to) b, then we can ~~get~~ measure any property of the population. (*)

3] It is not possible to measure the entire population. So, we infer about the population of the sample.

→ $x_1, x_2, \ldots, x_n$, these samples are random.

→ $P(x_i \le b) = P(-\infty, b]$ (define in (*))

x_i's are identically distributed.

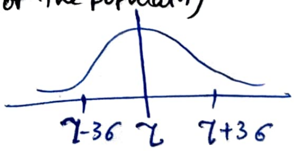
$x_1, x_2, \ldots, x_n$ are independent.

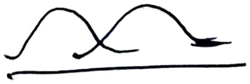
- $x_i \overset{iid}{\sim} F$, $i = 1, 2, \ldots, n$; $F(b) = P(-\infty, b]$
- $x_1, \ldots, x_n$ is a random sample from F.

4] The goal is to infer about the population from the sample.

→ Towards this, one may assume a structure of F (model), except some parameters (unknown constants, properties of the populatn)

eg: $x_i \sim N(\mu, \sigma^2)$





$\pi N(\mu_1, \sigma_1^2) + (1-\pi) N(\mu_2, \sigma_2^2)$

→ we use the samples to estimate the parameters, thereby estimate the populatn. [This procedure is called parametric inference]

What is Non-parametric inference?

↳ Not getting any structure of F main focus on the problem in hand which we want like mean, variance and we'll estimate the required.

Parametric inference:
- Point estimation (1)
- Interval estimation (3)
- Hypothesis estimation (2)

Point Estimation: $X_1, X_2, \ldots X_n$ is a random sample from $F_{\underset{\sim}{\theta}}$.

$P[-\infty, b]$ = populatn proportion in $(-\infty, b]$

populatn proportion in $(-\infty, b_1] \underset{\text{And}}{\underset{\sim}{\times}} (-\infty, b_2] = P_{(-\infty, b_1] \times (-\infty, b_2]}$

$\Rightarrow F_{X,Y}(b_1, b_2) = P(\underbrace{X_i \le b_1}_{\text{Height}}, \underbrace{Y_i \le b_2}_{\text{weight}})$

And.

$\begin{bmatrix} X_1 \\ Y_1 \end{bmatrix}, \ldots \begin{bmatrix} X_n \\ Y_n \end{bmatrix}$

$F_{X,Y}(b_1, b_2) = \int_{-\infty}^{b_1} \left\{ \int_{-\infty}^{b_2} \underbrace{f_{X,Y}(x,y)}_{\text{Joint pdf.}} \, dy \right\} dx$

$X_1, X_2, \ldots X_n$ is a random sample from $F_{\underset{\sim}{\theta}}$

we are interested in function of $\underset{\sim}{\theta}$, say $\psi = \psi(\underset{\sim}{\theta})$

Eg: $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$ $\quad \underset{\sim}{\theta} = \begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix}$ $\quad \psi : \mathbb{R}^2 \to \mathbb{R}$, $\psi(\underset{\sim}{\theta}) = \mu$.

statistic $\to T : \mathbb{R}^n \to \mathbb{R}$

A function of the samples ; $T(\underset{\sim}{X})$ is used to estimate $\psi = \psi(\underset{\sim}{\theta})$

$T(\underset{\sim}{X})$ is called an point estimation of $\psi(\underset{\sim}{\theta})$.

Notation: 1] For random variables use capital letters.

2] For realization use small letters.

3] For parameters use greek letters.

For a realization of the samples $x_1, x_2, \dots x_n$ the value of $T(\underset{\sim}{x})$ is called an estimate of $\psi(\underset{\sim}{\theta})$

which estimation is good?

Ex: $X_1, X_2, \dots X_n$ is a random sample from $Bin(m, \theta)$ ← known.

$$T_1(\underset{\sim}{X}) = \frac{\overline{X_n}}{m} \qquad \frac{E(X_i)}{m} = \theta$$

$E(Bi)\ np$

$var(Bi) = npq$

$$Var(X_i) = m\theta(1-\theta)$$

$$T_2(\underset{\sim}{X}) = \left(S_n^2 + \frac{\overline{X_n}^2}{m}\right) \Big/ m \qquad = m\theta - \frac{[E[X]]^2}{m}$$

$x_{11}, \dots x_{n1}$	$T_1(x_1)$	$T_2(x_2)$
⋮	⋮	
$x_{110k} \dots x_{n10k}$	$T_1(x_{10k})$	$T_2(x_{10k})$
	① $\overline{T_1}(\underset{\sim}{x})$	$\overline{T_2}(\underset{\sim}{x})$
	② $var(T_1(x))$	$var(T_2(x))$

$$E_\theta[T_1[\underset{\sim}{X}]] = E_\theta[\overline{X_n}/m]$$

$$= E_\theta\left[\frac{1}{mn}\sum_{i=1}^{n} X_i\right]$$

$$= \frac{1}{mn}\sum_{i=1}^{n} E_\theta(X_i)$$

$$= \frac{1}{mn}\sum_{i=1}^{n} m\theta = \underline{\underline{\theta}}$$

$$E_\theta(T_2[\underset{\sim}{X}]) = E_\theta\left[\frac{S_n^2}{m} + \frac{\overline{X_n}^2}{m^2}\right]$$

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

$$= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X_n}^2$$

$$= E_\theta\left[\frac{1}{mn}\sum_{i=1}^{n} X_i^2 - \frac{\overline{X_n}^2}{m} + \frac{\overline{X_n}^2}{m^2}\right]$$

$$= \frac{1}{mn}\sum_{i=1}^{n} E_\theta(X_i^2) - \frac{1}{m}\left[1 - 1/m\right] E_\theta(\overline{X_n}^2)$$

$$E_\theta(X_i^2) = Var_\theta(X_i) + E_\theta^2(X_i)$$

$$= m\theta(1-\theta) + m^2\theta^2$$

⊛ why it's good estimator if $E(T(\underset{\sim}{X}))$ is less?

for which property of $\psi(\theta)$

why we are taking deviation from θ

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

$$= E_\theta\left[\frac{1}{n^2}\left\{\sum_{i=1}^{n} x_i^2 + \sum_{i\neq j} x_i x_j\right\}\right]$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} E[x_i^2] + \frac{1}{n^2}\sum_{\substack{i=1 \\ i\neq j}}^{n}\sum_{j=1}^{n} E[x_i x_j]$$

verify: $E_\theta(T_2(x)) = \theta - \underbrace{\left[\frac{m-1}{mn}\right]\theta(1-\theta)}_{\text{Bias}}$.

So, $T_1(\underset{\sim}{X})$ is better estimate than $T_2(\underset{\sim}{X})$ as it's expected value is underestimating the θ.

Summary:

$X_1, X_2, \dots X_n \overset{iid}{\sim} F_{\underset{\sim}{\theta}}$ ($f_{\underset{\sim}{\theta}}$) both can be used.

Discrete distribution : $f_{\underset{\sim}{\theta}}$: pmf

continous distribution : $f_{\underset{\sim}{\theta}}$: pdf

Problem: Estimate $\psi(\underset{\sim}{\theta}) = \psi$ $\psi: \mathbb{R}^k \to \mathbb{R}$

Defn: (Estimator) A fn of $X_1, X_2, \dots X_n$, $T(\underset{\sim}{X})$, which is used to estimate ψ is called an estimator of ψ.

Q: How to choose a good estimator?

Example: $X_1, X_2, \dots X_n$ is a random sample from Bin(m,θ)

$$T_1(\underset{\sim}{X}) = \frac{\bar{X}_n}{m}, \quad T_2(\underset{\sim}{X}) = \frac{S_n^2}{m} + \frac{\bar{X}_n^2}{m^2}$$

$$E_\theta(T_1(\underset{\sim}{X})) = \theta, \quad E_\theta(T_2(\underset{\sim}{X})) = \theta - \frac{(m-1)}{mn}\theta(1-\theta)$$ → verify

Defn (Bias): Let $T(\underset{\sim}{X})$ be a estimator of $\psi(\underset{\sim}{\theta}) = \psi$ then the bias in estimation is $B_T(\psi(\theta)) = E_\theta(T(\underset{\sim}{X}) - \psi(\underset{\sim}{\theta}))$

$$B_T(\psi(\underset{\sim}{\theta})) = E_\theta[T(\underset{\sim}{X}) - \psi(\underset{\sim}{\theta})] \quad E_\theta[T(\underset{\sim}{X}) - \psi(\theta)]$$

$$B_T(\psi(\theta)) = E_\theta[T(\underset{\sim}{X}) - \psi(\theta)] \qquad = E_\theta[T(\underset{\sim}{X})] - \psi(\underset{\sim}{\theta})$$

$B_T(\psi(\underset{\sim}{\theta}))$ for T_1 is 0 $B_T(\psi(\underset{\sim}{\theta}))$ for T_2 is $-\frac{(m-1)}{mn}\theta(1-\theta)$

Remark:

[1] If $B_T(\psi(\underset{\sim}{\theta}))$ is positive then $T(\underset{\sim}{x})$ over estimates $\psi(\underset{\sim}{\theta})$

[2] If $B_T(\psi(\underset{\sim}{\theta}))$ is negative then $T(\underset{\sim}{x})$ under estimates $\psi(\underset{\sim}{\theta})$

Defn: [unbiased estimation]

$T(\underset{\sim}{x})$ is called an unbiased estimator of $\psi(\theta)$ if

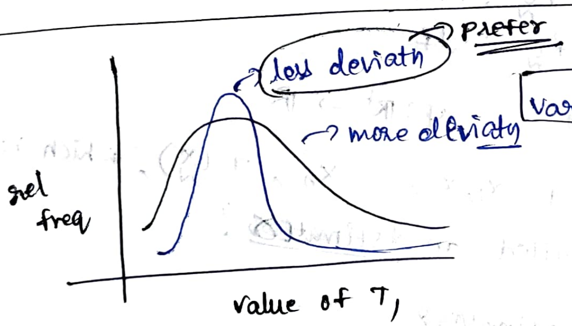
$B_T(\psi(\underset{\sim}{\theta})) = 0$, for all θ.

Q) $T_3(\underset{\sim}{x}) = \frac{x_1}{m}$ $\quad E_\theta(T_3(\underset{\sim}{x})) = \theta$

$var(T_1(\underset{\sim}{x})) < var(T_3(\underset{\sim}{x}))$ } T_1 is the better estimator than T_3.

(*) as T_1 is calculated by 'n' different samples & T_3 can be calculated by only one sample & that one sample can be bad sample

---



$$\boxed{var(T(x)) = E_\theta\left[\left[T(x) - \psi(\theta)\right]^2\right]}$$

$\hookrightarrow E_\theta(T(x))$

$$\cancel{var(T_3(x)) = E_\theta\left[(T_3(x) - \psi(\theta))^2\right]}$$

$$= \cancel{E_\theta\left[(T_3(x))^2 + \psi^2(\theta) - 2T_3(x)\psi(\theta)\right]}$$

$$= var\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right), \frac{1}{m^2}\theta(1-\theta))$$

$$var(T_3(\underset{\sim}{x})) = var\left[\frac{x_1}{m}\right]$$

$$= \frac{1}{m^2} var[x_1] = \frac{1}{m^2} m\theta(1-\theta) = \frac{1}{m}\theta(1-\theta) \to (1)$$

$$var[T_1(\underset{\sim}{x})] = var\left[\frac{\bar{x}_n}{m}\right] = \frac{1}{m^2} var(\bar{x}_n)$$

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

verify

Defn: (Relative efficiency)

let $T_1(\underset{\sim}{X})$ & $T_2(\underset{\sim}{X})$ be two unbiased estimators of $\psi(\underset{\sim}{\theta})$ then the relative efficiency of T_1 and T_2 is.

$$\boxed{RE_{T_1,T_2}(\psi(\underset{\sim}{\theta})) = \frac{Var(T_1(\underset{\sim}{X}))}{Var(T_2(\underset{\sim}{X}))}}$$

Remark: $RE_{T_1,T_2}(\psi(\underset{\sim}{\theta}))>1$ then $T_2(\underset{\sim}{X})$ is more (better) efficient than $T_1(\underset{\sim}{X})$ in estimating $\psi(\underset{\sim}{\theta})$.

Example: $RE_{T_1,T_3}(\theta) = \frac{1}{n} < 1$

↳ As it's <1, T_1 is better than T_3.

Cramer-Rao lower bound: (CRLB):

1. CRLB provides a lower bound of variance of unbiased estimators.

~~$X_1, \dots X_n$ ... f_θ~~

2. CRLB holds for certain distributions which satisfy some regularity conditions.

one of the regularity condition is: the sample space ($\mathcal{X}$) is free of the parameter (doesn't depend on parameter)
[set of values that can be realized by the samples.

Ex: $X \sim N(0,1)$

$\mathcal{X} = \mathbb{R}$ || $\mathcal{X}$ doesn't depend on θ.

If $X \sim U(0,\theta)$ || $\mathcal{X}$ depends on θ.
$\mathcal{X} = [0,\theta]$

Defn (Fisher information): let $X_1, X_2, \dots X_n$ be a random sample from f_θ. then the fisher information of θ is

$I_n(\theta) = n I(\theta)$ where, $I(\theta) = E_\theta\left[\left\{\frac{\partial}{\partial\theta}\log f_\theta(X_1)\right\}^2\right]$

Defn (CRLB): let $X_1, X_2, \dots X_n$ be a random sample from a distribution with pmf / pdf f_θ and the f_θ satisfies contain regularity conditions. Then for any estimator $T(\underset{\sim}{X})$,

$$var_\theta(T(\underset{\sim}{X})) \geq \frac{\left[\frac{\partial}{\partial\theta} E_\theta(T(\underset{\sim}{X}))\right]^2}{n\,I(\theta)}$$

Note: If $T(\underset{\sim}{X})$ is unbiased for θ. then $var_\theta(T(\underset{\sim}{X})) \geq [n(I(\theta))]^{-1}$

Ex: (continue)

$$f_\theta(x_1) = \binom{m}{x_1}\theta^{x_1}(1-\theta)^{m-x_1}$$

$$\log f_\theta(x_1) = \log\binom{m}{x_1} + x_1\log\theta + (m-x_1)\log(1-\theta)$$

$$\frac{\partial}{\partial\theta}\log f_\theta(x_1) = \frac{x_1}{\theta} + \frac{m-x_1}{1-\theta}(-1)$$

$$\frac{\partial}{\partial\theta}\log f_\theta(x_1) = \frac{x_1}{\theta} + \frac{m-x_1}{\theta-1}$$

$$I(\theta) = E_\theta\left[\frac{\partial}{\partial\theta}\log f_\theta(x_1)\right]^2$$

$$I(\theta) = E_\theta\left[\underbrace{\left(\frac{x_1}{\theta} - \frac{m-x_1}{1-\theta}\right)^2}_{\left(\frac{\partial}{\partial\theta}\log f_\theta(x_1)\right)^2}\right]$$

$$I_\theta = E_\theta\left(\frac{\partial}{\partial\theta}\log f_\theta(x_1)\right)^2$$

$$\left[\frac{\partial}{\partial\theta}(E_\theta(T(x)))\right]^2$$

$$I(\theta) = E\left[\frac{x_1^2}{\theta^2} + \left(\frac{m-x_1}{1-\theta}\right)^2 - \frac{2x_1(m-x_1)}{\theta(1-\theta)}\right]$$

$$\boxed{I(\theta) = \frac{m}{\theta(1-\theta)}}$$ verify

$E_\theta \frac{\partial}{\partial\theta}$

CRLB of $T_1(\underset{\sim}{X}) = \frac{\bar{X}_n}{m}$ is $\frac{1}{n\,I(\theta)} = \frac{1}{nm/\theta(1-\theta)}$

$$= \frac{\theta(1-\theta)}{mn}$$

$$\frac{\left[\frac{\partial}{\partial\theta}E_\theta(T(x))\right]^2}{E_\theta\left[\frac{\partial}{\partial\theta}\log f_\theta(x)\right]^2}$$

Recall: $var_\theta\left[\frac{\bar{X}_n}{m}\right] = \frac{\theta(1-\theta)}{nm}$

$$I = E_\theta\left[\frac{\partial}{\partial\theta}(\log f_\theta(x))\right]^2$$

Defn: (**Efficient estimator**): If the variance of an unbiased estimator $T(\underset{\sim}{X})$ of $\psi(\theta)$ achieves the CRLB, then it's called an efficient estimator.

Q) can a biased estimator be better than an unbiased estimator?

Sol: Yes!

Goal: $E_\theta\left\{\left[T(\underset{\sim}{X}) - \psi(\theta)\right]^2\right\}$ → minimize wrt T.

~~$MSE_T(\theta)$~~ $MSE_T(\psi(\theta))$

→ [Mean squared error] MSE

less MSE → good estimator

$MSE_T(\psi(\theta))$

$$= E_\theta\left[\left\{\underbrace{T(\underset{\sim}{x}) - E_\theta(T(\underset{\sim}{X}))}_{a} + \underbrace{E_\theta(T(\underset{\sim}{X})) - \psi(\theta)}_{b}\right\}^2\right]$$

$$= E_\theta\Big[\{T(x) - E_\theta(T(\underset{\sim}{X}))\}^2 + \{E_\theta(T(\underset{\sim}{X})) - \psi(\theta)\}^2$$

$$+ 2 E_\theta \underbrace{\{\overbrace{E_\theta(T(\underset{\sim}{X})) - \psi(\theta)}^{B_T(\psi(\theta))}\}\{T(\underset{\sim}{X}) - E_\theta(T(\underset{\sim}{X}))\}}_{=0}\Big]$$

$$= \underbrace{E_\theta\left[\{T(x) - E_\theta(T(x))\}^2\right]}_{var(T(\underset{\sim}{X}))} + B_T^2(\psi(\theta)) + 0$$

$$= var_\theta(T(\underset{\sim}{X})) + B_T^2(\psi(\theta))$$

$E_\theta[T(\underset{\sim}{X})] - \psi(\theta)$

Problems:

~~Rao Beal~~

3) $T_3 = \overline{X}_n$

$$E[\overline{X}_n] = E\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right]$$

$$= \frac{1}{n} E[X_1 + X_2 + \dots + X_n] = \frac{1}{\cancel{n}} \cancel{n} E[X_1]$$

$$\boxed{E[\overline{X}_n] = E[X_1]}$$

$$var(T(\underset{\sim}{x})) \geq \frac{\left(\cancel{\frac{\partial}{\partial\theta}} E \left(\frac{\partial}{\partial\theta} E_\theta(T(x))\right)\right)^2}{n\left[E_\theta\left[\frac{\partial}{\partial\theta}\log f_\theta(x_1)\right]^2\right]}$$

2) $T_2 = \frac{1}{8}X_1 + \frac{3}{4(n-2)}(X_2 + \dots + X_{n-1}) + \frac{1}{8}X_n$.

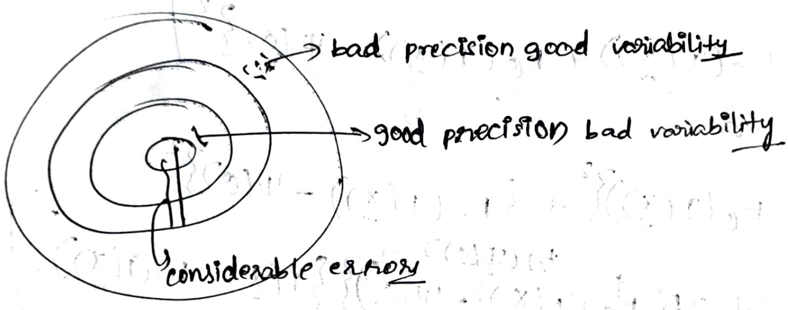
$\mathbb{E}[T_2] = \frac{1}{8}\mathbb{E}[X_1] + \frac{3}{4\cancel{(n-2)}}\cancel{(n-2)}\,\mathbb{E}[X_1] + \frac{1}{8}\mathbb{E}[X_1]$

$= \left(\frac{1}{8} + \frac{3}{4} + \frac{1}{8}\right)\mathbb{E}[X_1]$

$= \underline{\mathbb{E}[X_1]}$

(*) efficiency is reciprocal of variability (variance)

(*) precision → error should be less



---

problems:

3) $var(T_3) = \frac{\sigma^2}{n}$ $T_3 = \bar{X}_n$

$var(\bar{X}_n) = var\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] = \frac{1}{n^2}\,var\left[\sum_{i=1}^{n} X_i\right]$

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

2) $var(T_2) = \frac{(n+16)}{32(n-2)}\sigma^2$

$var\left[\frac{1}{8}X_1 + \frac{3}{4(n-2)}(X_2 + \dots + X_{n-1}) + \frac{1}{8}X_n\right]$

sample size (n) $\rightarrow$ accuracy & precision

replication / repeatition $\rightarrow$ precision

Asymptotics $\rightarrow$ ↑sing the sample size.

Problem:

$$T_1 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2$$

$\sigma^2 = E[x^2] - E[(x)]^2$

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

$\frac{1}{n-1}$

$$E[T_1] = \frac{1}{n-1} E\left[\sum_{i=1}^{n}(x_i - \bar{x}_n)^2\right]$$

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

$E[x^2] =$

$$= \frac{1}{n-1}\left[E\left[x_1^2 + x_2^2 + \dots + x_n^2\right] - n\,E[\bar{x}_n^2]\right]$$

$$= \frac{1}{n-1}\left[n\,E[x_1^2] - n\,E[\bar{x}_n^2]\right)$$

$$= \frac{n}{(n-1)}\left[E[x_1^2] - E[\bar{x}_n^2]\right)$$

$\cancel{\sigma^2 = (E[x])^2} -$

$$\sigma^2 = E[x_1^2] - (\underbrace{E[x_1]}^{\mu})^2 \Rightarrow \boxed{E[x_1^2] = \sigma^2 + \mu^2}$$

$$\frac{\sigma^2}{n^2} = E[\bar{x}_n^2] - [E[\bar{x}_n]]^2$$

$E\left[\frac{x_1 + x_2 + \dots + x_n}{n}\right]$

$\frac{1}{n} n \mu$

$$\frac{\sigma^2}{n^2} + \mu^2$$

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

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

$$\therefore \boxed{E[T_1] = \sigma^2}$$

$$\boxed{\cancel{E[T_1] = \frac{(n+1)}{n}\sigma^2}}$$

$$var(T_1) = E[T_1^2] - [E[T_1]]^2$$

$$\text{var}(T_1) = \text{var}\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2\right]$$

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

$$= \frac{1}{(n-1)^2}\left[\text{var}\left[\sum_{i=1}^{n} x_i^2\right] + \text{var}(n\bar{x}_n^2)\right.$$
$$\left. - 2\,\text{cov}\left[\sum x_i^2, n\bar{x}_n^2\right]\right]$$

$$\text{var}(T_1) = \mathbb{E}[T_1^2] - [\mathbb{E}[T_1]]^2$$

$E_\theta(f(x) - \psi(\theta)^2)$

Summary:

$B_T^2(\psi(\theta))$

$$MSE_{T_n}(\psi(\underset{\sim}{\theta})) = E_{\underset{\sim}{\theta}}\left[\{T(\underset{\sim}{x}) - \psi(\underset{\sim}{\theta})\}^2\right]$$

An estimator which minimizes $MSE_{T_n}(\psi(\underset{\sim}{\theta})) \; \forall\, \theta \in T_n$

then the estimator is the best?

$$MSE_{T_n}(\psi(\underset{\sim}{\theta})) = \underbrace{Bias_{T_n}^2(\psi(\underset{\sim}{\theta}))}_{B} + \underbrace{var_\theta(T(\underset{\sim}{X}))}_{V}$$

Example:

let $x_1, x_2, \ldots x_n$ be a random sample $N(\mu, \sigma^2)$

$$s_n^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2 \quad \text{(Biased)}$$

$$s_n^{*2} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2 \quad \text{(unbiased)}$$

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

?

$$\boxed{\begin{gathered}\mathbb{E}(x) = (n-1)\sigma^2 \\ \text{var}(x) = 2(n-1)\sigma^4\end{gathered}}$$

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

$$E(s_n^{*2}) = \sigma^2 \leftarrow UE \rightarrow \text{(unbiased estimator)}$$

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

$\frac{\sigma^4}{n^2} +$

$$\text{var}(s_n^2) = \frac{1}{n^2}\cdot 2(n-1)\sigma^4 \qquad \text{var}(s_n^{*2}) = \frac{2}{(n-1)}\sigma^4$$

$$MSE_{s_n^2}(\sigma^2) = \frac{1}{n^2}\sigma^4 + \frac{2(n-1)}{n^2}\sigma^4 = \frac{2n-1}{n^2}\sigma^4$$

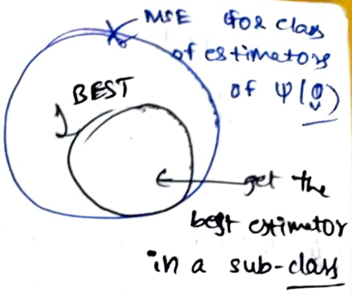
$$MSE_{s_n^{*2}}(\sigma^2) = \frac{2}{(n-1)}\sigma^4$$

Example:

consider the class of estimators of the form

$$T_{a_n}(\underset{\sim}{X}) = \frac{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}{a_n}$$

$$= \frac{X}{a_n} \longrightarrow (*)$$



$$E\left[T_{a_n}(\underset{\sim}{X})\right] = \frac{n-1}{a_n}\sigma^2$$

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

$$\operatorname{var} T_{a_n}(\sigma^2) = \frac{2(n-1)\sigma^4}{a_n^2}$$

$$MSE\, T_{a_n}(\sigma^2) = \sigma^4\frac{(n-1-a_n)^2}{a_n^2} + \sigma^4\frac{2(n-1)}{a_n^2}$$

$$= g(a_n)$$

To minimize $g(a_n)$ w.r.t a_n:

i) $\dfrac{\partial(g(a_n))}{\partial a_n} = 0 \leftarrow$ solve this equatn for a_n } First order condition.

let a_n^* be the soln.

ii) $\left.\dfrac{\partial^2(g(a_n))}{\partial a_n^2}\right|_{a_n = a_n^*} > 0$ } second order condition

$$\sigma^4\left[\frac{a_n^2\left[2(n-1-a_n)(-1)\right] - (n-1-a_n)^2(2a_n)}{a_n^4} + 2(n-1)(-2)a_n^{-3}\right] = 0$$

$$\left[a_n\left[-2n+2+2a_n\right] - 2\left[n^2+1+a_n^2+2a_n-2n-2na_n\right] - 4(n-1)\right] = 0$$

$\Rightarrow -2na_n + 2a_n + \cancel{2a_n^2} - 2n^2 - \cancel{2} - \cancel{2a_n^2} - 4a_n + \cancel{4n} + 4na_n - \cancel{4n} + \cancel{4} = 0$

$2na_n - 2a_n - 2n^2 + 2 = 0$

$a_n^*(n-1) = n^2 - 1$

$\boxed{a_n^* = n+1}$

Also check the second order condition:

$$S_n^{**2} = \frac{1}{n+1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2 ;$$ Best estimator in the choosen class.

⊛ usually, one chooses the class of unbiased estimators, say, $\mathcal{U}(\psi(\underset{\sim}{\theta}))$ and tries to minimize the MSE within the class.

If $T(\underset{\sim}{X}) \in \mathcal{U}(\psi(\underset{\sim}{\theta}))$

then $MSE_T(\psi(\underset{\sim}{\theta})) = var_\theta(T(\underset{\sim}{X}))$

⊛ Minimizing $MSE_T(\psi(\underset{\sim}{\theta}))$ within the class of unbiased estimators of $\psi(\underset{\sim}{\theta})$, i.e $\mathcal{U}(\psi(\underset{\sim}{\theta}))$ is same as minimizing the variance, $var_\theta(T(\underset{\sim}{X}))$, within the class.

⊛ The best estimator in the class $\mathcal{U}(\psi(\underset{\sim}{\theta}))$ is called uniformly minimum variance unbiased estimator (UMVUE)

Defn: (UMVUE): $T(\underset{\sim}{X})$ is called the UMVUE of $\psi(\underset{\sim}{\theta})$,

if (i) $E_{\underset{\sim}{\theta}}(T(\underset{\sim}{X})) = \psi(\underset{\sim}{\theta})$ for all $\underset{\sim}{\theta}$.

(ii) $var_{\underset{\sim}{\theta}}(T(\underset{\sim}{X})) \le var_{\underset{\sim}{\theta}}(T'(\underset{\sim}{X}))$ for all $\underset{\sim}{\theta}$,

where $T'(\underset{\sim}{X})$ is another unbiased estimator of $\psi(\underset{\sim}{\theta})$

T' also satisfies (i)

$T' \in \mathcal{U}(\psi(\underset{\sim}{\theta}))$

How do we get the UMVUE?

1) UMVUE is unique.

2) let $T(\underset{\sim}{X})$ be a "complete and sufficient" statistic.

If $E_{\underset{\sim}{\theta}}[g(T(\underset{\sim}{X}))] = \psi(\underset{\sim}{\theta}). \ \forall \underset{\sim}{\theta}$

Then $g(T(\underset{\sim}{X}))$ is the UMVUE.

3) let $X_1, X_2, \ldots X_n$ be random sample from some distribution with pdf/pmf $f_{\underset{\sim}{\theta}}$; and $f_{\underset{\sim}{\theta}}(x)$ has the following form:

$$f_{\underset{\sim}{\theta}}(x) = \exp\{a(\underset{\sim}{\theta}) + h(x) + \sum_{j=1}^{k} T_j(\underset{\sim}{X})\, b(\underset{\sim}{\theta})\}$$

Then $[T_1(\underset{\sim}{X}), \ldots T_k(\underset{\sim}{X})]$ are jointly complete sufficient. Then $f_{\underset{\sim}{\theta}}$ belongs to the exponential family.

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

$$= \exp\left\{ -\log\left((2\pi)^{n/2}\sigma^n\right)\right\} \cdot \exp\left\{ \frac{-1}{2\sigma^2} \left[\sum_{i=1}^{n} x_i^2 - 2\mu \sum_{i=1}^{n} x_i + n\mu^2 \right] \right\}$$

$$= \exp\left[-\frac{n}{2}\log(2\pi) - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2 + \frac{2\mu}{2\sigma^2}\sum_{i=1}^{n} x_i - \frac{n\mu^2}{2\sigma^2} \right]$$

$$= \exp\left[\underbrace{-\frac{n}{2}\log(2\pi)}_{h(\underset{\sim}{x})} \underbrace{- \left\{ n\log\sigma + \frac{n\mu^2}{2\sigma^2} \right\}}_{a(\underset{\sim}{\theta})} \underbrace{+ \frac{\mu}{\sigma^2} n\bar{x}_n - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2}_{T_1(\underset{\sim}{x})b_1(\underset{\sim}{\theta}) + T_2(\underset{\sim}{x})b_2(\underset{\sim}{\theta})} \right]$$

⊛ Anything depends on X, then random else not random.

⊛ Statistic is complete function of X.

$(\sum x_i, \sum x_i^2) = (T_1(\underset{\sim}{x}), T_2(\underset{\sim}{x}))$ jointly CSS $\rightarrow$ $\boxed{\text{complete sufficient statistics}}$

- poisson (λ)
- Gamma (α, β)

SUMMARY:

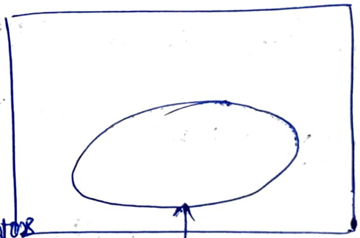
- MSE
- UMVUE

If $T^*(\underset{\sim}{x})$ is the UMVUE,

then

(i) $E_{\underset{\sim}{\theta}}(T^*(\underset{\sim}{x})) = \varphi(\underset{\sim}{\theta}) \ \forall \underset{\sim}{\theta}$

(unbiased)



class of estimators $\varphi(\underset{\sim}{\theta})$

class of unbiased estimators $\varphi(\underset{\sim}{\theta})$

(ii) $var_{\underset{\sim}{\theta}}(T^*(\underset{\sim}{x})) \le var_{\underset{\sim}{\theta}}(T(\underset{\sim}{x})) \ \forall \underset{\sim}{\theta}$,

[minimum variance] where $T(\underset{\sim}{x})$ is an unbiased estimator of $\varphi(\underset{\sim}{\theta})$

⊛ If an unbiased estimator is CRLB then it's UMVUE

$\bar{x}$: UE of μ

$\frac{1}{n}\sum(x_i - \bar{x})^2$: UE of σ^2

How to obtain the UMVUE?

Result: If $T(\underset{\sim}{X})$ is a complete sufficient statistic (css) and $g(T(\underset{\sim}{X}))$ is an unbiased estimator of $\psi(\underset{\sim}{\theta})$, then $g(T(\underset{\sim}{X}))$ is the UMVUE.

Defn: (Exponential family)
Let $x_1, x_2, \ldots x_n$ be a random sample from $f_{\underset{\sim}{\theta}}$ be a pdf/pmf of some distribution & has the following form.

About realization

$$f_{\underset{\sim}{\theta}}(\underset{\sim}{x}) = \exp\{a(\underset{\sim}{\theta}) + h(\underset{\sim}{x}) + T_1(\underset{\sim}{x})c_1(\underset{\sim}{\theta}) + \ldots + T_k(\underset{\sim}{x})c_k(\underset{\sim}{\theta})\}$$

(*)

~~Then $\{T_1(\underset{\sim}{X}), \ldots T_k(\underset{\sim}{X})\}$~~

~~statistic, X is~~ R.V.

Then $f_{\underset{\sim}{\theta}}$ belongs to the exponential family.

Examples: Binomial, poisonn, Normal, Exponential, Gamma, Beta, Geometric, Negative Binomial.

Result: The statistic $[T_1(\underset{\sim}{X}), \ldots T_k(\underset{\sim}{X})]$ is complete sufficient.

Example: $x_1, \ldots x_n \overset{\text{IID}}{\sim} N(\mu, \sigma^2)$

$$f_{\underset{\sim}{\theta}}(\underset{\sim}{x}) = \exp\left\{-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2 + \frac{\mu}{\sigma^2}\sum x_i + \frac{n\mu^2}{2\sigma^2}\right\}$$

$$= \exp\Big\{\underbrace{-\frac{n}{2}\log(2\pi)}_{h(x)} \underbrace{-\frac{n}{2}\log\sigma^2 + \frac{n\mu^2}{2\sigma^2}}_{a(\underset{\sim}{\theta})} \underbrace{-\frac{1}{2\sigma^2}}_{c_1(\underset{\sim}{\theta})}\underbrace{\sum_{i=1}^{n} x_i^2}_{T_1(\underset{\sim}{x})} + \underbrace{\frac{\mu}{\sigma^2}}_{c_2(\underset{\sim}{\theta})}\underbrace{\sum_{i=1}^{n} x_i}_{T_2(\underset{\sim}{x})}\Big\}$$

$\left[\sum_{i=1}^{n} x_i^2, \sum_{i=1}^{n} x_i\right]$ is jointly css

Result: A one-one function of a css is also css

$\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2, \bar{x}\right]$ is also jointly css.

(*) Find UMVUE's of μ & σ^2

(i) $\bar{x}_n$ is UMVUE of μ.

(ii) Recall $E_{\underset{\sim}{\theta}}\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2\right] = \sigma^2$

$\Downarrow$

$\frac{1}{n-1}\left\{\sum_{i=1}^{n} x_i^2 - n\bar{x}_n^2\right\}$ $g\left(\binom{x}{y}\right) = \frac{1}{n-1}\{x - y_n^2\}$

is the UMVUE of σ^2

Topic ③ : How to find a "good" estimator?

$X_1, \dots X_n$ is a random sample of size n from $\beta(\alpha, \beta)$ distribution. Estimate α & β using $X_1, \dots X_n$.

$\underset{\sim}{\theta} = \binom{\alpha}{\beta}$

$$E(X_1) = \left(\frac{\alpha}{\alpha+\beta}\right)$$

$f_\theta(x) = \dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{Beta(\alpha,\beta)}$

$0 < x < 1$

$$var(X_1) = \frac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)}$$

Solve the equations:

$$\begin{cases} \bar{X}_n = E(X_1) = \dfrac{\alpha}{\alpha+\beta} \\ \dfrac{1}{n}\sum_{i=1}^{n} X_i^2 = E(X_1^2) = \dfrac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)} + \dfrac{\alpha^2}{(\alpha+\beta)^2} \end{cases}$$

1) Method of Moments (MOM)

let $\underset{\sim}{\theta} \in \mathbb{R}^k$

Sample (raw) moment of order r:

$$M_r' = \frac{1}{n}\sum_{i=1}^{n} X_i^r \; ; \; r = 0, 1, \dots$$

population (raw) moment of order r:

$$\mu_r' = E_\theta(X_1^r)$$

solve the k-equations:

$$M_r' = \mu_r' \quad \text{for } r = 1, \dots k$$

Remark: Suppose we are interested in $\psi(\underset{\sim}{\theta})$

Then first find $\hat{\theta}_{MOM}$ (i.e, the MOM estimator of $\underset{\sim}{\theta}$)

then Next use $\psi(\hat{\theta}_{MOM})$ as the MOM estimator of $\psi(\underset{\sim}{\theta})$

$X_i \overset{IID}{\sim} \text{Poisson}(\lambda)$

$$\psi(\lambda) = P(X_1 \geq 1) = 1 - P(X_1 = 0) = 1 - e^{-\lambda}$$

$$E(1 - e^{-\bar{X}_n}) \neq 1 - e^{-\lambda}$$

In the above example, $\hat{\lambda}_{MOM} = \bar{X}_n$, so that MOM estimator of $\psi(\lambda)$ is $\boxed{1 - e^{-\bar{X}_n}}$.

Example: $X_1, \dots X_n \overset{IID}{\sim} U[0,\theta]$ distribution $f_\theta(x) = \frac{1}{\theta}, \; 0 \le x \le \theta$

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

CSS

MOM: $\bar{X}_n = M_1' = E(X_1) = \frac{\theta}{2}$

$\hat{\theta}_{MOM} = 2\bar{X}_n$

HW: suppose you toss a biased coin with prob of head $p \in \{1/3, 2/3\}$. You toss the coin 5 times and get 2 heads. what is $\hat{p}$?

$X_i \overset{IID}{\sim} Bin(5, p)$ $\qquad p \in \{1/3, 2/3\}$

$f_X(x) = \binom{5}{x} p^x (1-p)^{5-x}, \quad x = 0, \dots 5.$

$f_X(2) = \binom{5}{2} p^2 (1-p)^3$

$p = 1/3$: $\frac{80}{243} = P(X=2 / p=1/3)$

$p = 2/3$: $\frac{40}{243} = P(X=2/p=2/3)$

Exercise:

1] $X_1, \dots X_n \overset{iid}{\sim} Bin(m, p)$ $\to$ show that $\frac{\bar{X}_n}{m}$ is UMVUE of p

$T = \frac{\bar{X}_n}{m}$

M-I] $E[T] = E\left[\frac{1}{m}\bar{X}_n\right] = p$

$var(T) = var\left(\frac{1}{m}\bar{X}_n\right) = \frac{p(1-p)}{nm}$

Every one-one function of CSS is also CSS

3) H.W consider a random sample of size 'n' of a geometric distribution with parameter $\theta \in (0,1)$ find UMVUE of θ.

---

7) For $N(\mu, \sigma^2)$ $\theta = (\mu, \sigma^2)$

a) Find UMVUE of μ

b) Find UMVUE of σ^2

---

4) Fo be poisson(λ) $\theta = \lambda$

a) Find UMVUE of θ

b) is it an efficient estimator.

BLUE → UE with min variance.

↳ Biased linear unbiased estimator

---

5) $X_1, X_2, \ldots X_n$ iid $U(0,\theta)$

$f_{X_1, X_2 \ldots X_n}(x_1, x_2, \ldots x_n) = f_{\underset{\sim}{X}}(x) = \left(\frac{1}{\theta}\right)^n, \quad x_i \in (0,\theta) \ \forall i$

$\equiv 0 < x_1, \ldots x_n < \theta$

$x_1 = \min\{x_1, \ldots x_n\}$

$\equiv 0 < x_1 < x_n < \theta$

$x_n = \max\{x_1, \ldots x_n\}$

↳ It doesn't belongs to exp family because range depends on the parameter.

$E(2\bar{X}_n) = 2 E\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = 2 E[x_1]$

$= 2 \cdot \frac{\theta}{2} = \theta$

$X \sim U(a,b)$

$E(X) = \frac{a+b}{2}$

$var(X) = \frac{(b-a)^2}{12}$

$\frac{4}{n^2} var(x_1 + x_2 + \cdots + x_n)$

$var(2\bar{X}_n) = 4 var(\bar{X}_n) = \frac{4}{n} var(x_1) = \frac{4}{n} \frac{(\theta)^2}{12}$

$= \frac{\theta^2}{3n}$

$P(X_n \le x) = P(\max\{X_1, \ldots X_n\} \le x)$

$= (P(X_1 \le x))^n = \left(\frac{x}{\theta}\right)^n, \quad 0 < x < \theta$

$= P(X_1 \le x, X_2 \le x, \ldots X_n \le x)$

$$f_{X_n}(x) = \frac{nx^{n-1}}{\theta^n}; \quad 0 < x < \theta$$
$$0; \quad \text{otherwise}$$

$$E[X_{(n)}] = \int_0^\theta (x) \frac{nx^{n-1}}{\theta^{n-1}} dx$$

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

$$\underline{\underline{E\left[\frac{n}{n+1} \cancel{X}\right] = \theta}}$$

$$E\Big[\underbrace{\frac{n}{n+1} X_{(n)}}_{T_1}\Big]$$

$$E[2\bar{X}_n] = \theta$$

$$var\left(\frac{n+1}{n} X_{(n)}\right) = E\left[\frac{n+1}{n} X_n\right]^2 - \theta^2$$

$$E\left[\frac{n+1}{n} X_{(n)}\right]^2 = \int_0^\theta \left(\frac{n}{n+1} x\right)^2 \frac{nx^{n-1}}{\theta^n} dx$$

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

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

$\frac{n+1}{n} X_{(n)}$ is good estimator than $2\bar{X}_n$

SUMMARY:

M_r' is to estimate μ_r'

$$E[M_r'] = E\left[\frac{1}{n} \sum_{i=1}^{r} X_i^r\right] = \frac{1}{n} \sum_{i=1}^{r} E[X_i^r]$$

$$= \mu_r'$$

$$S_n^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2, \quad var(X_i) = \sigma^2$$

Maximum likelihood estimation: (MLE)

Example-1: $x_1, x_2, \dots x_n$ from Bern(θ) ; $\theta \in \{\frac{2}{3}, \frac{1}{3}\}$

$n=5$, # heads $= 3$

$P(\text{\# heads} = 3 / \theta = 1/3) = z_1$

vs $P(\text{\# heads} = 3 / \theta = 2/3) = z_2$

As $z_2 > z_1$, $\hat{\theta}_{ML} = 2/3$

Example-2: Suppose in a shop there are 3 types of mobiles:

Type	Battery life (years)
A	2
B	3
C	6

You bought one phone from the shop which lasted for 4.5 years.

let X: be the battery life of a mobile phone with average battery life θ

$\theta \in \{2, 3, 6\}$ → generally life time is exponential distribution & sometime poiso distribution

$X \sim f_\theta$: $f_\theta(x) = \frac{1}{\theta} e^{-x/\theta}$ $E(x) = \theta$, where $\theta = \text{mean}$,

$\lambda = 1/\text{mean}$

$x = 4.5$,

How likely is to observe $x = 4.5$ when $\theta = 2$.

$f_{\theta=2}(x=4.5) = \frac{1}{2} \exp\{-4.5/2\} = 0.0527$

$f_{\theta=3}(x=4.5) = \frac{1}{3} \exp(-4.5/3) = 0.0743$

$f_{\theta=6}(x=4.5) = \frac{1}{3} \exp(-4.5/6) = 0.0787$

→ This is more likely

$X \sim f_\theta$

$f_\theta(x) = \lim_{h \to 0} \frac{P(x-h \le X < x+h)}{2h}$, $\hat{\theta}_{ML} = 6$

Example 2 (cont...)

$x_1, x_2, \dots x_5 \overset{IID}{\sim} Exp(\underset{\sim}{\theta})$

$\underbrace{x_1, x_2, \dots x_5}_{}$ Battery lives of 5 mobile phones of the same type, say 'D'

$\underset{\sim}{\theta}$ ↑ Avg battery life of type 'D' phones

~~Not taking (5)~~

$x_1 = 4.5$
$x_2 = 3$
$x_3 = 2.5$
$x_4 = 5$
$x_5 = 6$

Taking '1' sample; $\boxed{x_1 \overset{IID}{\sim} Exp(\underset{\sim}{\theta})}$

$x_1 = 4.5$

Maximize $f_\theta(x)$ evaluated at $x = 4.5$ wrt θ.

Maximize $f_\theta(\underset{\sim}{x})$ evaluated at $\underset{\sim}{x} = \begin{pmatrix} 4.5 \\ 3 \\ 2.5 \\ 5 \\ 6 \end{pmatrix}$, wrt θ

$= \underbrace{f_\theta(x_1) f_\theta(x_2) f_\theta(x_3) f_\theta(x_4) f_\theta(x_5)}$

$f_\theta(x_i) = \frac{1}{\theta} e^{-x_i/\theta} \quad \forall \; i = 1, 2, \dots 5$

$f_\theta(\underset{\sim}{x}) = \frac{1}{\theta^5} e^{-\Sigma x_i/\theta}$

$= \frac{1}{\theta^5} e^{-21/\theta} \rightarrow (1)$

just to simplify the fn ? making easy to diff'ble.

$\log(f_\theta(\underset{\sim}{x})) = g_\theta(\underset{\sim}{x})$

$= -5 \log \theta - \frac{21}{\theta} \rightarrow (2)$

$\frac{\partial}{\partial \theta}(f_\theta(\underset{\sim}{x})) = (-5)\theta^{-6} e^{-21/\theta} + \frac{1}{\theta^5} e^{-21/\theta} \frac{(-21)(-1)}{\theta^2} = 0$

$= \frac{-5}{\theta^6} + \frac{21}{\theta^7} = 0$

$= \frac{21}{\theta^7} = \frac{5}{\theta^6} \Rightarrow \boxed{\theta = 21/5}$ ⊛

let $\hat{\theta}_1$ maximizes $\rightarrow (1)$ & $\hat{\theta}_2$ maximizes $\rightarrow (2)$

$\hat{\theta}_1 = \hat{\theta}_2$ } → both are equal.

↳ Apply log doesn't change (Monotone transformation)

$\boxed{\frac{\partial^2 (g(\theta))}{\partial \theta} < 0 \Rightarrow \text{concludes } \hat{\theta}_{MLE} = 4.2}$ → should express in clear

Defn: (likelihood function) → In likelihood fix sample
In joint distributn fix θ

let $X_1, X_2, \dots X_n$ be a random sample from distribution with pdf/pmf $f_\theta(\cdot)$:

The likelihood fn: $L(\theta; \underset{\sim}{x}) = f_\theta(x_1) \dots f_\theta(x_n)$

[The likelihood of observing $\underset{\sim}{x}$ when 'θ' is true]

The log likelihood fn is $l(\theta; \underset{\sim}{x}) = \log L(\theta; \underset{\sim}{x})$

Defn (MLE): The ML Estimator of 'θ' is the maximizer of $L(\theta; \underset{\sim}{x})$ wrt θ.

[MLE is also the maximizer of $l(\theta; \underset{\sim}{x})$]

$$\hat{\theta} = \arg\max L(\theta; \underset{\sim}{x})$$

is a fn (x)

Rough Analysis

$$z(y) = \underset{x \in \mathbb{R}}{\arg\max}\ f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) \quad \mathbb{R}^2 \to \mathbb{R}$$

Random → $\boxed{\hat{\theta}(\underset{\sim}{X}) = \underset{\theta}{\arg\max}\ L(\theta; \underset{\sim}{x})}$

statistic.

Example: let $X_1, X_2, \dots X_n$ be a random sample from $N(\mu, \sigma^2)$ distribution, $\underset{\sim}{\theta} = \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}$

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

find MLE of $\underset{\sim}{\theta}$.

$$L(\underset{\sim}{\theta}, \underset{\sim}{x}) = \frac{1}{(2\pi\sigma^2)^{n/2}} \left(e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}\right)$$

$$\frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$l(\underset{\sim}{\theta}, \underset{\sim}{x}) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\frac{\partial l(\underset{\sim}{\theta}, \underset{\sim}{x})}{\partial \underset{\sim}{\theta}} = \underset{\sim}{0} \Rightarrow \frac{\partial l(\theta; x)}{\partial \mu} = 0 \to \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = 0 \to ①$$

$$\frac{\partial l(\theta; x)}{\partial \sigma^2} = 0 \to -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \to ②$$

from → ① & ②

$$\boxed{\sigma^2 = \frac{1}{n}\sum(x_i - \bar{x}_n)^2}$$

SOC.

$$\frac{\partial^2 l(\underset{\sim}{\theta}; \underset{\sim}{x})}{\partial\theta\, \partial\theta} = H = \begin{bmatrix} \frac{\partial^2(\theta, x)}{\partial\mu\, \partial\mu}, & \frac{\partial^2(\theta, x)}{\partial\sigma^2 \partial\mu} \\ \frac{\partial^2(\theta, x)}{\partial\mu\, \partial\sigma^2}, & \frac{\partial^2(\theta, x)}{\partial\sigma^2 \partial\sigma^2} \end{bmatrix}$$

SUMMARY: $x_1, x_2, \ldots x_n \sim f_\theta$

likelihood $f_{\underset{\sim}{\theta}}(\underset{\sim}{x}) = f_{\underset{\sim}{\theta}}(x_1) \ldots f_{\underset{\sim}{\theta}}(x_n) = L(\underset{\sim}{\theta}; \underset{\sim}{x})$

Maximize the likelihood function $L(\underset{\sim}{\theta}; \underset{\sim}{x})$ wrt $\underset{\sim}{\theta}$.

let $L(\underset{\sim}{\theta}; \underset{\sim}{x})$ is maximized for $\hat{\theta}(\underset{\sim}{x})$, then $\hat{\theta}^{ML}(\underset{\sim}{x})$ is called the ML estimate of $\underset{\sim}{\theta}$.

In general the maximizer of $L(\underset{\sim}{\theta}, \underset{\sim}{x})$ wrt $\underset{\sim}{\theta}$, $\hat{\theta}_{ML}(\underset{\sim}{X})$ is called the ML estimator of $\underset{\sim}{\theta}$.

Example: Normal (μ, σ^2)

log likelihood.

$$l(\underset{\sim}{\theta}; \underset{\sim}{x}) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

FOC:

$$\left.\begin{aligned} \frac{\partial l(\underset{\sim}{\theta}; \underset{\sim}{x})}{\partial\mu} &= 0 \rightarrow (1) \\ \frac{\partial l(\underset{\sim}{\theta}, \underset{\sim}{x})}{\partial\sigma^2} &= 0 \rightarrow (2) \end{aligned}\right\} \quad \left.\begin{aligned} \hat{\mu} &= \overline{x_n} \\ \hat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 \end{aligned}\right\} \hat{\theta}$$

SOC:

$$\underset{(\text{Hessian})}{H}\begin{pmatrix}\mu \\ \sigma^2\end{pmatrix} = \begin{bmatrix} \frac{\partial^2 l}{\partial\mu^2}, & \frac{\partial^2 l}{\partial\sigma^2 \partial\mu} \\ \frac{\partial^2 l}{\partial\mu\,\partial\sigma^2} & \frac{\partial^2 l}{(\partial\sigma^2)^2} \end{bmatrix} = \begin{bmatrix} \frac{-n}{\sigma^2} & -\frac{\sum_{i=1}^{n}(x_i-\mu)}{\sigma^4} \\ " & \frac{n}{2\sigma^4} - \frac{\sum(x_i-\mu)^2}{\sigma^6} \end{bmatrix}$$

$$H\left(\begin{pmatrix}\hat{\mu} \\ \hat{\sigma}^2\end{pmatrix}\right) = \begin{bmatrix} \frac{-n}{s_n^2} & -\frac{\sum_{i=1}^{n}(x_i-\bar{x}_n)^2}{s_n^4} = 0 \\ 0 & \frac{n}{2s_n^4} - \frac{n s_n^2}{(s_n^2)^3} = \frac{n}{2s_n^4} - \frac{n}{s_n^4} \end{bmatrix} = \begin{bmatrix} \frac{-n}{s_n^2} & 0 \\ 0 & \frac{-n}{2s_n^4} \end{bmatrix}$$

$A^{2\times2}$ is p.d $\Leftrightarrow$ a_{11} is +ve and $|A| > 0$.
positive definite.

$A^{2\times2}$ is n.d $\Leftrightarrow$ $\boxed{a_{11} \text{ is } -ve \text{ and } |A| > 0}$
negative definite.

Here, $a_{11} = \frac{-n}{S_n^2} < 0$

& $|A| = \frac{n^2}{2 S_n^4} > 0$

$\Rightarrow \boxed{\hat{\underset{\sim}{\theta}} = \hat{\underset{\sim}{\theta}}_{MLE}}$,

Example: $X_1, X_2, \dots X_n \overset{IID}{\sim}$ Uniform $(0, \theta)$ distribution.

Likelihood function,

$$L(\theta; \underset{\sim}{x}) = f_\theta(x_1) \dots f_\theta(x_n)$$

$$= \frac{1}{\theta^n} \text{ \& } 0 \le x_i \le \theta \ \forall \ i = 1, 2, \dots n.$$

$$\Downarrow$$

$$\boxed{0 \le \min\{x_1, \dots x_n\} \le \max\{x_1, \dots x_n\} \le \theta}$$

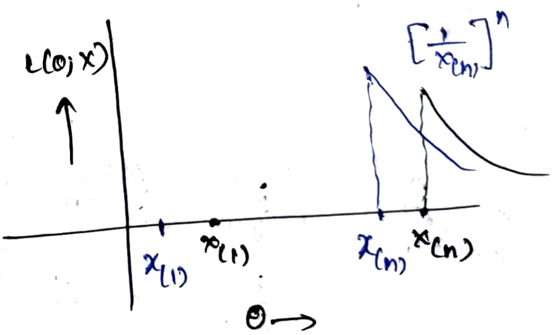
$$0 \le x_{(1)} \le x_{(n)} \le \theta$$

$$\mathbb{I}(A) = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

$$L(\theta; \underset{\sim}{x}) = \frac{1}{\theta^n} \mathbb{I}(x_{(1)} \ge 0, x_n \le \theta)$$

?

$$\hat{\theta}_{ML} = x_{(n)}$$



HW $\boxed{x_1, x_2, \dots x_n \overset{iid}{\sim} U(\alpha, \beta)}$

case(1): $\frac{x}{P_0}$ is not an integer. Then $\hat{m}_{MLE} = \left[\frac{x}{P_0}\right]$

case(2): $\frac{x}{P_0}$ is an integer. then both $\frac{x}{P_0}$ and $\frac{x}{P_0} - 1$ are MLE's of m.

Example: $X_i \sim Bern(P)$; $i = 1, 2, \cdots$

$P \in (0,1)$

$\hat{\theta}_{MLE} = \bar{X}_n$, $x_1 = 0$, $x_2 = 0$

Invariance property: If $\hat{\underset{\sim}{\theta}}_{MLE}$ is the MLE of $\underset{\sim}{\theta}$. then

$\psi(\hat{\underset{\sim}{\theta}}_{MLE})$ is the MLE of $\psi(\underset{\sim}{\theta})$

$= \hat{\psi}_{MLE}$

$X \sim Poisonn(\lambda)$ $\quad \lambda^*_{MLE} = \bar{X}_n$

$P(X=1) = e^{-\lambda}\lambda = \psi(\lambda)$, $\hat{\psi}_{MLE} = \bar{X}_n e^{-\bar{X}_n}$

Sampling distribution:

$X_1, X_2, \cdots X_n \overset{IID}{\sim} Bern(\theta)$ distribution.

$T(\underset{\sim}{X}) = \sum_{i=1}^{n} X_i$ $\quad P_\theta(T(\underset{\sim}{X}) = [n/2]) = \binom{n}{m} \theta^m (1-\theta)^{n-m}$

$[= m]$

Defn: (sampling Distribution): The distribution of a statistic $T(\underset{\sim}{X})$

- Q] How to derive sampling distribution?

I] MGF Technique:

Defn: (MGF) let $X \sim F_\theta$, then the moment generating function of X, at t is

$$M_X(t) = E_\theta[e^{tx}] \; ; \; t \in \mathbb{R}$$

Remark: 1] We say $M_X(t)$ exists for a R.V X if $E_\theta[e^{tx}]$ exists for all 't' in neighbourhood of $\underline{0}$.

2] If MGF exists for some R.V X, then all moments of X exist.

3) MGF, if exists, uniquely characterizes the distribution.

Example 1: $X \sim N(\mu, \sigma^2)$

$$M_X(t) = \cancel{E[e^{tx}] = E\left[e^{\sigma t\left(\frac{x-\mu}{\sigma}\right) + \frac{t\mu}{\sigma}}\right]}$$

$$\cancel{= e^{t\mu/\sigma}\, E\left[e^{t\left(\frac{x-\mu}{\sigma}\right)}\right]}$$

$\frac{x-\mu}{\sigma} = Z \sim N(0,1)$

$dz = \frac{dx}{\sigma}$

$$\cancel{= e^{t\mu/\sigma} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 + t\left(\frac{x-\mu}{\sigma}\right)} \cdot dx}$$

$$\cancel{= e^{t\mu/\sigma} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2 + tz}\, dz}$$

$$\cancel{= e^{t\mu/\sigma} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z^2 - 2tz + t^2)}}_{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-t)^2}} dz \cdot e^{t^2/2}}$$

$$M_X t = E[e^{tx}] = E\left[e^{\sigma t\left(\frac{x-\mu}{\sigma}\right) + t\mu}\right]$$

$$= e^{t\mu}\, E\left[e^{\sigma t\left(\frac{x-\mu}{\sigma}\right)}\right]$$

$$E_x(g(x)) = \int_{-\infty}^{\infty} g(x) \cdot f_x(x)\, dx$$

$$= e^{t\mu} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 + \sigma t\left(\frac{x-\mu}{\sigma}\right)} \cdot dx$$

$$= e^{t\mu} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2 + \sigma t z} \cdot dz$$

$$= e^{t\mu} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z - t\sigma)^2} \cdot dz}_{=1} \cdot e^{t^2\sigma^2/2}$$

→ like pdf of Normal distrib $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$ equates to 1 even if shifts

$$= e^{t\mu + t^2\sigma^2/2}$$

Ex 1] $x_1, x_2, \dots x_n \overset{IID}{\sim} N(\mu, \sigma^2)$

what is the dist $T(\underset{\sim}{x}) = \sum_{i=1}^{n} x_i = y$

$M_y(t) = E[e^{ty}]$

$= E\left[e^{t \sum_{i=1}^{n} x_i}\right]$

$= E\left[e^{tx_1 + tx_2 + \dots + tx_n}\right]$

$= E[e^{tx_1}] \dots E[e^{tx_n}] = M_{x_1}(t) \dots M_{x_n}(t)$

$= \left[e^{t\mu + t^2\sigma^2/2}\right]^n = \underline{e^{tn\mu + t^2 n\sigma^2/2}}$

Q) $x_i \overset{IND}{\sim} N(\mu_i, \sigma_i^2)$ $\quad y = x_1 + x_2 + \dots + x_n.$

$\sum x_i \sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$

II] <u>change of variables:</u>

$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \longrightarrow \underset{\sim}{y} = T(\underset{\sim}{x})$

$\underset{\sim}{T} : R^n \to R^n$ one-one function

① <u>Jacobian:</u>

$J = J_{\underset{\sim}{x} \to \underset{\sim}{y}} = \left\| \det\left(\frac{\partial \underset{\sim}{x}}{\partial \underset{\sim}{y}}\right) \right\|$ abs det

$= \left\| \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \dots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \dots & \frac{\partial x_n}{\partial y_n} \end{bmatrix} \right\|$

let $f_x(\underset{\sim}{x})$ is the pdf of $\underset{\sim}{x}$, evaluated at $\underset{\sim}{x}$,

$f_{\underset{\sim}{y}}(\underset{\sim}{y})$: is the pdf of $\underset{\sim}{y}$, evaluated at $\underset{\sim}{y}$

② Find the range of $\underset{\sim}{y}$

$f_{\underset{\sim}{y}}(\underset{\sim}{y}) = f_{\underset{\sim}{x}}\left(\underset{\sim}{T}^{-1}(\underset{\sim}{y})\right) \cdot J_{\underset{\sim}{x} \to \underset{\sim}{y}}$ $\quad \underset{\sim}{y} \in$ appropriate Range.

Example 2: $X_i \overset{iid}{\sim}$ Gamma (α, β) ; $i = 1, 2$. Find the distribution of $Y_1 = \dfrac{X_1}{X_1 + X_2}$

Define, $Y_2 = X_1 + X_2$

Then $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \to \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$ is one-one

$\left.\begin{array}{l} X_1 = Y_1 Y_2 \\ X_2 = Y_2 - Y_1 Y_2 = Y_2(1 - Y_1) \end{array}\right\} \underset{\sim}{T}^{-1}(\underset{\sim}{y})$

$y = T(x)$
$x = T^{-1}(\underset{\sim}{y})$
$f_X(x)$
$f_Y(y) = f_X(T$

Range: $0 \le Y_1 \le 1$

$Y_2 \ge 0$

Jacobian: $J_{\underset{\sim}{x} \to \underset{\sim}{y}} = \left\| \begin{bmatrix} y_2 & y_1 \\ -y_2 & (1 - y_1) \end{bmatrix} \right\|$

$= |(y_2(1 - y_1) + y_1 y_2)|$

$= y_2$

$\dfrac{x^{\alpha - 1} e^{-x/\beta}}{\Gamma(\alpha)\, \beta^{\alpha}}$

$f_{\underset{\sim}{y}}(\underset{\sim}{y}) = f_{\underset{\sim}{x}}(\underset{\sim}{T}^{-1}(\underset{\sim}{y})) \cdot J_{\underset{\sim}{x} \to \underset{\sim}{y}}$

$\Gamma(\alpha) = \int_0^{\infty} e^{-t}\, t^{\alpha - 1}\, dt$

? $= \left[\dfrac{\beta^{2\alpha}}{(\Gamma_\alpha)^2} e^{-\beta(x_1 + x_2)} (x_1 x_2)^{\alpha - 1} \right]_{T^{-1}(y)} \cdot y_2$

$\Gamma(\alpha) = (\alpha - 1)!$

$= \left[\dfrac{\beta^{2\alpha}}{(\Gamma_\alpha)^2} e^{-\beta y_2} \left(y_2^2\, y_1 (1 - y_1) \right)^{\alpha - 1} \right] y_2$

$(\beta)^{\alpha}$

$= \dfrac{\beta^{2\alpha}}{(\Gamma_\alpha)^2} e^{-\beta y_2}\, y_2^{2\alpha - 1}\, y_1^{\alpha - 1} (1 - y_1)^{\alpha - 1}$

$$f_{Y_1}(y_1) = \left\{ \underbrace{\int_0^\infty \frac{\beta^{2\alpha}}{(\Gamma\alpha)^2} e^{-\beta y_2} \cdot y_2^{2\alpha-1} \, dy_2}_{\text{Gamma}(\beta, 2\alpha)} \right\} \cdot y_1^{\alpha-1} (1-y_1)^{\alpha-1}$$

$$= \frac{\Gamma 2\alpha}{(\Gamma\alpha)^2} \cdot 1 \cdot y_1^{\alpha-1} \cdot (1-y_1)^{\alpha-1}$$

$Y_1 \rightarrow \text{Beta}(\alpha, \alpha)$

Example 3: Let X be an abs. cont R.V with CDF F. Define $U = F(x)$. What is the dist of U?

Inverse transform: $x = F^{-1}(U)$

Range of U : $0 \leq U \leq 1$

Jaccobian : $\frac{\partial x}{\partial y} = \left(\frac{\partial y}{\partial x}\right)^{-1} = \left(\frac{\partial F(x)}{\partial x}\right)^{-1} = \frac{1}{f_X(x)} = \frac{1}{f_X(F^{-1}(u))}$

$$f_U(u) = f_X(F^{-1}(u)) \cdot \frac{1}{f_X(F^{-1}(u))} = 1$$

$U \sim \text{Uniform}(0,1)$ distributn, $0 \leq u \leq 1$.

Remark: We can use this result to simulate a random number following arbitrary cont. distribution.

Ex: Suppose we want to simulate a random sample from $\text{Exp}(\lambda)$ and $\lambda = 3.7$.

- You can simulate a random number u_0 between 0 & 1.

Treat u_0 as a realization of $F(X) = U$.

So, $F(x_0) = u_0$

$\Rightarrow 1 - e^{-\lambda x_0} = u_0$

$\Rightarrow x_0 = \frac{-\log(1-u_0)}{\lambda}$

$X \sim \text{Exp}(\lambda)$

$F(x) = \int_0^x \lambda e^{-\lambda w} dw$

$= \left[-e^{-\lambda w}\right]_0^x = 1 - e^{-\lambda x}$

Example 1 [Box-Muller Transformation]

$U_i \overset{iid}{\sim} \text{Unif}(0,1)$

$Z_1 = \sqrt{-2\log U_1}\,\cos(2\pi U_2)$

$Z_2 = \sqrt{-2\log U_1}\,\sin(2\pi U_2)$

$\sqrt{-2\log U_1} = r$ $\quad$ $Z_1 = r\cos\theta$

$2\pi U_2 = \theta$ $\quad$ $Z_2 = r\sin\theta$

$\begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \to \begin{bmatrix} r \\ \theta \end{bmatrix} \to \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$

$0 \le r < \infty \quad , \quad 0 \le \theta \le 2\pi$

$f_{\begin{bmatrix} U_1 \\ U_2 \end{bmatrix}}\left(\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right) = 1 \qquad 0 \le u_i \le 1 \quad , \quad i = 1,2.$

$r^2 = -2\log u_1$

$\Rightarrow u_1 = \exp\left\{-\frac{r^2}{2}\right\}$

$u_2 = \frac{1}{2\pi}\theta$

$J_{\underline{u} \to \binom{r}{\theta}} = \left\| \begin{matrix} e^{-r^2/2}(-r) & 0 \\ 0 & \frac{1}{2\pi} \end{matrix} \right\|$

$= \frac{1}{2\pi} r e^{-r^2/2}$

$f_{\binom{r}{\theta}}\left(\begin{pmatrix} r \\ \theta \end{pmatrix}\right) = \frac{1}{2\pi} r \cdot e^{-r^2/2} \quad ; \quad r \ge 0, \; 0 \le \theta \le 2\pi$

2nd transformation: $\begin{pmatrix} r \\ \theta \end{pmatrix} \to \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$

Range: $-\infty < z_i < \infty; \; i = 1,2,$

$z_1 = r\cos\theta$

$z_2 = r\sin\theta$

$r = \sqrt{z_1^2 + z_2^2}$

$\theta = \tan^{-1}\left(\frac{z_2}{z_1}\right)$

$J_{\binom{r}{\theta} \to \binom{z_1}{z_2}} = \frac{1}{J_{\binom{z_1}{z_2} \to \binom{r}{\theta}}}$

$= \left\| \begin{matrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{matrix} \right\|^{-1} = \frac{1}{r}$

$f_{\left(\binom{z_1}{z_2}\right)}\left(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right) = f_{\left(\binom{r}{\theta}\right)}\left(\begin{pmatrix} r \\ \theta \end{pmatrix}\right) \cdot \frac{1}{r}$ $\qquad -\infty < z_i < \infty$, $i = 1,2$

$= \frac{1}{2\pi} r \cdot e^{-r^2/2} \cdot \frac{1}{r} = \frac{1}{2\pi} e^{-(z_1^2 + z_2^2)/2} = \frac{1}{\sqrt{2\pi}} e^{-z_1^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-z_2^2/2}$

$Z_i \overset{IID}{\sim} N(0,1)$

$f_{\left(\binom{X}{Y}\right)}\left(\binom{x}{y}\right) = h(x)\cdot g(y)$ } $\to$ X & Y are independent

$\int f_{\left(\binom{X}{Y}\right)}\left(\binom{x}{y}\right) dx\,dy = \int \underbrace{\frac{h(x)}{c}}_{e^{-\lambda x}} \cdot g(y)\cdot c \cdot dx\,dy$

$= 1$

$$\boxed{\frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}}$$

III] Distribution Function Technique:

Example: $X_i \overset{IID}{\sim} U(0,\theta)$

$T(\underset{\sim}{X}) = X_{(n)}$

let $F_{T,\theta}$ be the CDF of T

$F_{T,\theta}(t) = P_\theta(T \le t) = P_\theta(X_{(n)} \le t)$

$= P_\theta[X_1 \le t, X_2 \le t, \ldots X_n \le t]$

$= P_\theta[X_1 \le t]\, P[X_2 \le t] \ldots P[X_n \le t]$

$= (P_\theta[X_1 \le t])^n$

$= \left[\int_0^t \frac{1}{\theta} dx\right]^n = \frac{t^n}{\theta^n} \qquad \underline{0 \le t \le \theta}$

let $f_{T,\theta}$ be the pdf of T

Then $f_{T,\theta}(t) = \frac{d}{dt} F_{T,\theta}(t) = \frac{n t^{n-1}}{\theta^n}$; $\underline{0 \le t \le \theta}$

let $U = X_{(1)}$

$F_{U,\theta}(u) = P_\theta(U \le u) = P_\theta(X_{(1)} \le u)$

$= 1 - P_\theta(X_{(1)} > u)$

$= 1 - P_\theta[X_1 > u, X_2 > u, \ldots X_n > u]$

$= 1 - (P_\theta(X_1 > u))^n = (1 - [1 - P(X_1 \le u)]^n)$

The technique:

let $T = T(\underset{\sim}{X})$

$$F_T(t) = P(T(\underset{\sim}{X}) \le t)$$
$$= P(T(\underset{\sim}{X}) \in (-\infty, t])$$
$$= P(\underset{\sim}{X} \in \underbrace{T^{-1}((-\infty, t])}_{A_{T,t}})$$

$$A_{T,t} = \{ \underset{\sim}{x} : T(\underset{\sim}{x}) \in (-\infty, t] \}$$

$$= \begin{cases} \int_{A_{T,t}} f_{\underset{\sim}{X}}(\underset{\sim}{x})\, dx & \text{if abs. cont} \\ \sum_{A} f_{\underset{\sim}{X}}(\underset{\sim}{x}) & \text{if discrete} \end{cases}$$

Ex: 2] $X \sim N(0,1)$. Find the PDF of $X^2 = Y$.

$$F_Y(y) = P(Y \le y) = P(X^2 \le y)$$
$$= P(-\sqrt{y} \le X \le \sqrt{y})$$
$$= \Phi(\sqrt{y}) - \Phi(-\sqrt{y})$$
$$= 2\Phi(\sqrt{y}) - 1$$

$$f_Y(y) = 2\phi(\sqrt{y}) \cdot \frac{1}{2(\sqrt{y})} \qquad (0 \le y < \infty)$$

$$= \frac{(y)^{-1/2}}{\sqrt{2\pi}} e^{-y/2}$$

$$\sim \text{Gamma}\left(\tfrac{1}{2}, \tfrac{1}{2}\right)$$

[$X_i \overset{\text{IID}}{\sim} \text{Gamma}(\alpha, \beta)$; $i = 1, 2, \ldots, n$

$\sum_{i=1}^{n} X_i \sim \text{Gamma}(n\alpha, \beta)$ distn]

[$M_{X_1}(t) = \frac{\beta^\alpha}{(\beta - t)^\alpha}$ $\quad t < \beta$

$= (1 - t/\beta)^{-\alpha}$]

Ex: 2] (cont...) $X_i \overset{\text{iid}}{\sim} N(0,1)$

What is the distn of $\underbrace{X_1^2}_{Z_1} + X_2^2 + \ldots + \underbrace{X_n^2}_{Z_n} = Y$?

As X_i's are IID

Z_i's are also IID, $Z_i \sim \text{Gamma}(1/2, 1/2)$

$\Rightarrow Y = \sum_{i=1}^{n} Z_i \sim \underbrace{\text{Gamma}\left(\frac{n}{2}, \frac{1}{2}\right)}_{\chi^2_{(n)} \text{ distributn.}}$

Defn: (χ^2- distribution)

let $X_1, X_2, \ldots X_n$ be IID $N(0,1)$ variates then $Y = X_1^2 + \ldots + X_n^2$ is said to follow a χ^2 distribution with degrees of freedom n.

Properties: 1) $E[Y / Y \sim \chi^2_{(n)}] = n$

$\text{var}[Y / Y \sim \chi^2_{(n)}] = 2n$

2) [Additive property] If $Y_i \overset{IID}{\sim} \chi^2_{(n_i)}$ distribution $i = 1, 2, \ldots$

then $Y_1 + Y_2 \sim \chi^2_{(n_1+n_2)}$ distributn

3) $X_1, X_2, \ldots X_n \overset{IID}{\sim} N(\mu, \sigma^2)$

Then $\dfrac{\sum_{i=1}^{n} (X_i - \bar{X}_n)^2}{\sigma^2} \sim \chi^2_{(n-1)}$

PROOF: let $Z_1, Z_2, \ldots Z_n \overset{IID}{\sim} N(0,1)$

$$\sum_{i=1}^{n} (Z_i - \bar{Z})^2 = \sum_{i=1}^{n} Z_i^2 - n\bar{Z}^2$$

$$\underset{\sim}{Z} = \begin{bmatrix} Z_1 \\ \vdots \\ Z_n \end{bmatrix}$$

$$= \underset{\sim}{Z}^T \underset{\sim}{Z} - \frac{\underset{\sim}{Z}' \underset{\sim}{1}\, \underset{\sim}{1}' \underset{\sim}{Z}}{n}$$

$$\bar{Z} = \frac{\underset{\sim}{Z}' \underset{\sim}{1}}{n}$$

$$= \underset{\sim}{Z}^T \left(I - \frac{\underset{\sim}{1}\, \underset{\sim}{1}'}{n} \right) \underset{\sim}{Z}$$

$$= \underset{\sim}{Z}^T H_n \underset{\sim}{Z}$$

Step-2) $Z_1, Z_2, \ldots Z_n \sim N(0,1)$

$\underset{\sim}{Y} = W \underset{\sim}{Z}$ where W is such that

$\underset{\sim}{Z} = W^{-1} \underset{\sim}{Y}$

$$W = \begin{bmatrix} 1/\sqrt{n} & 1/\sqrt{n} & \cdots & 1/\sqrt{n} \\ & W_{(1)} & & \end{bmatrix}$$

$= W^T \underset{\sim}{Y}$

$\sum Z_i^2$ in terms of $\underset{\sim}{Y}$ and $W^T W = I$

$n=3$, $\underset{\sim}{a_1} = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$ $\underset{\sim}{a_2} = \begin{pmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$ $a_3 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ -1/\sqrt{3} \end{pmatrix}$

Gram-schmidt orthogonalization.

find the distribution of $\underset{\sim}{y}$.

Defn: ($\chi^2_{(n)}$ distn) $X_1, X_2, \dots X_n \overset{iid}{\sim} N(0,1)$

$$T = \sum_{i=1}^{n} X_i^2 \sim \chi^2_n \equiv \text{Gamma}\left(\frac{n}{2}, 1/2\right)$$

Property: 1) $T \sim \chi^2_{(n)}$

$E(T) = n$; $\text{var}(T) = 2n$

2) If $X_1, X_2, \dots X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

Then $\sum_{i=1}^{n} \frac{(X_i - \bar{X}_n)^2}{\sigma^2} \sim \chi^2_{(n-1)}$

PROOF:

Step 1: let $Z_1, \dots Z_n \overset{iid}{\sim} N(0,1)$

$$ns_n^2 = \sum_{i=1}^{n} (Z_i - \bar{Z}_n)^2 = \underbrace{\sum_{i=1}^{n} Z_i^2}_{\underset{\sim}{Z}^T \underset{\sim}{Z}} - \underbrace{n\bar{Z}_n^2}_{\underset{\sim}{Z}'\left(\frac{\underset{\sim}{1}\underset{\sim}{1}'}{n}\right)\underset{\sim}{Z}} = \underset{\sim}{Z}' \underbrace{\left(I - \frac{1}{n}\underset{\sim}{1}\underset{\sim}{1}'\right)}_{H} \underset{\sim}{Z}$$

$\underset{\sim}{Z} = \begin{bmatrix} Z_1 \\ \vdots \\ Z_n \end{bmatrix}$

$\bar{Z}_n = \frac{\underset{\sim}{1}'\underset{\sim}{Z}}{n}$

$\underset{n\times 1}{\underset{\sim}{x}} \rightarrow \underset{n\times 1}{\underset{\sim}{y}}$

$\underset{\sim}{y} = \underset{n\times n}{A}\underset{\sim}{x}$

$= (H\underset{\sim}{Z})' H \underset{\sim}{Z}$

$H = H^T H = H^2$

$\text{rank}(H) = \text{tr}(H) = n-1$

Step-2: $\underset{\sim}{y} = W\underset{\sim}{Z}$ where $W = \begin{bmatrix} 1/\sqrt{n} & \dots & 1/\sqrt{n} \\ & W_{(1)} & \end{bmatrix}$ with $W^T W = I$

$\}$ othogonal Transformation

$$W = \begin{bmatrix} \frac{1}{\sqrt{n}}\underset{\sim}{1}' \\ w_2' \\ \vdots \\ w_n' \end{bmatrix}$$

$\underset{\sim}{a_1}, \underset{\sim}{a_2}, \ldots \underset{\sim}{a_n}$ is basis of $\mathbb{R}^n$.

$\underset{\sim}{b_1}, \underset{\sim}{b_2}, \ldots \underset{\sim}{b_n}$ is orthogonal basis of $\mathbb{R}^n$.

$$\begin{cases} b_i^T b_j = 0 & \forall \; i \neq j \\ \|b_j\| = 1 & \forall \; i \end{cases}$$

$$\underset{\sim}{b_1} = \frac{\underset{\sim}{a_1}}{\|\underset{\sim}{a_1}\|} \qquad \underset{\sim}{b_2} = \frac{\underset{\sim}{a_2} - f(\underset{\sim}{a_1}, \underset{\sim}{a_2})}{\|\underset{\sim}{a_2} - f(\underset{\sim}{a_1}, \underset{\sim}{a_2})\|}$$

$$J_{\underset{\sim}{z} \to \underset{\sim}{y}} = \left\| \frac{\partial \underset{\sim}{z}}{\partial \underset{\sim}{y}} \right\| \qquad \frac{\partial \underset{\sim}{y}}{\partial \underset{\sim}{z}} = \begin{vmatrix} \frac{\partial y_1}{\partial z_1} & \cdots & \frac{\partial y_1}{\partial z_n} \\ \vdots & & \\ \frac{\partial y_n}{\partial z_1} & \cdots & \frac{\partial y_n}{\partial z_n} \end{vmatrix}$$

$$= \frac{1}{\left\| \frac{\partial \underset{\sim}{y}}{\partial \underset{\sim}{z}} \right\|}$$

$$= \frac{1}{\|w\|} = 1$$

$$y_1 = w_{11} z_1 + \ldots + w_{1n} z_n$$

$$y_2 = w_{21} z_1 + \ldots + w_{2n} z_n$$

$$\vdots$$

$$y_n = w_{n1} z_1 + \ldots + w_{nn} z_n$$

$$\frac{\partial \underset{\sim}{y}}{\partial \underset{\sim}{z}} = \frac{\partial (w \underset{\sim}{z})}{\partial \underset{\sim}{z}} = w$$

Range: $y_i \in \mathbb{R} \quad \forall \; i = 1, \ldots n.$

$$f_{\underset{\sim}{y}}(\underset{\sim}{y}) = f_{\underset{\sim}{z}}(w^{-1}\underset{\sim}{y}) \cdot J$$

$$= e^{-\frac{1}{2}\sum y_i^2/2} \big/ (2\pi)^{n/2} \qquad y_i \overset{IID}{\sim} N(0,1); \quad i = 1, \ldots n.$$

$$f_{\underset{\sim}{z}}(\underset{\sim}{z}) = \frac{1}{(2\pi)^{n/2}} e^{-\sum_{i=1}^{n} z_i^2/2}$$

$$\sum_{i=1}^{n} z_i^2 = \underset{\sim}{z}' \underset{\sim}{z} = \underbrace{(w^{-1}\underset{\sim}{y})'(w^{-1}\underset{\sim}{y})}_{\underset{\sim}{y}' \underbrace{(w^{-1})' w^{-1}}_{(w^T w)^{-1}} \underset{\sim}{y}} = \underset{\sim}{y}' \underbrace{(w^T w)^{-1}}_{I} \underset{\sim}{y} = \underset{\sim}{y}^T \underset{\sim}{y} = \sum_{i=1}^{n} y_i^2$$

$$\underset{\sim}{y}^T \underset{\sim}{y} = (W\underset{\sim}{z})' W \underset{\sim}{z} = \underset{\sim}{z}' W' W \underset{\sim}{z}$$

$$W = \begin{bmatrix} \frac{1}{\sqrt{n}} \underset{\sim}{1}' \\ W_{(1)} \end{bmatrix}$$

$$= \underset{\sim}{z}' \left[\frac{1}{\sqrt{n}} \underset{\sim}{1} : W_{(1)}^T \right] \begin{bmatrix} \frac{1}{\sqrt{n}} \underset{\sim}{1}' \\ W_{(1)} \end{bmatrix} \underset{\sim}{z}$$

$$W^T = \left[\frac{1}{\sqrt{n}} \underset{\sim}{1} : W_{(1)}^T \right]$$

$$= \underset{\sim}{z}' \left[\frac{1}{n} \underset{\sim}{1} \underset{\sim}{1}' + W_{(1)}^T W_{(1)} \right] \underset{\sim}{z}$$

$$\underset{\sim}{y} = W \underset{\sim}{z}$$

$$= \frac{1}{n} \underset{\sim}{z}' \underset{\sim}{1} \underset{\sim}{1}' \underset{\sim}{z} + \underset{\sim}{z}' W_{(1)}^T W_{(1)} \underset{\sim}{z}$$

$$y_1 = \frac{1}{\sqrt{n}} \underset{\sim}{z}' \underset{\sim}{1} = \frac{\sum z_i}{\sqrt{n}}$$

$$= y_1^2 + \underset{\sim}{z}' W_{(1)}^T W_{(1)} \underset{\sim}{z}$$

$$\Rightarrow y_1^2 + y_2^2 + \dots + y_n^2 = y_1^2 + \underset{\sim}{z}' W_{(1)}^T W_{(1)} \underset{\sim}{z}$$

$$\Rightarrow \underbrace{y_2^2 + \dots + y_n^2}_{\chi^2_{(n-1)}} = \underset{\sim}{z}^T \underbrace{W_1^{(T)} W_{(1)}}_{H} \underset{\sim}{z}$$

$$W^T W = I$$

$$\Rightarrow \left[\frac{1}{n} \underset{\sim}{1} \underset{\sim}{1}' + W_{(1)}^T W_{(1)} \right] = I$$

$$\Rightarrow W_{(1)}^T W_{(1)} = I - \frac{1}{n} \underset{\sim}{1} \underset{\sim}{1}' = H$$

$$\Rightarrow y_2^2 + \dots + y_n^2 = \underbrace{\underset{\sim}{z}' H \underset{\sim}{z}}_{n s_n^2} \sim \chi^2_{(n-1)}$$

$$y_1 = \sqrt{n}\, \bar{z}_n$$

$$y_1 \perp\!\!\!\perp \{y_2, \dots, y_n\}$$

$$\Rightarrow \underset{\bar{z}_n}{\underset{\|}{g(y_1)}} \perp\!\!\!\perp \underbrace{h\left(\begin{pmatrix} y_2 \\ \vdots \\ y_n \end{pmatrix}\right)}_{s_n^2}$$

Remark: $\bar{z}_n$ is independent of s_n^2

$\boxed{\text{Jacobian of orthogonal transformation}}$

Step-3: $x_1, \dots, x_n \sim N(\mu, \sigma^2)$

$x_i = \mu + \sigma z_i$, $i = 1, 2, \dots n$, $z_i \sim N(0,1)$

$$\sum_{i=1}^n (x_i - \bar{x}_n)^2 = \sum_{i=1}^n (\mu + \sigma z_i - \mu - \sigma \bar{z}_n)^2 = \sigma^2 \sum_{i=1}^n (z_i - \bar{z}_n)^2$$

$$\Rightarrow \frac{\sum_{i=1}^n (x_i - \bar{x}_n)^2}{\sigma^2} = \sum_{i=1}^n (z_i - \bar{z}_n)^2 \sim \chi^2_{(n-1)}$$

Corollary: $E\left[\sum_{i=1}^{n}\right.$

F-distribution:

Defn: $(F_{m,n})$ If $X \sim \chi^2_{(m)}$ distribution & $Y \sim \chi^2_{(n)}$ distribution independently, then

$$T = \frac{X/m}{Y/n} = \frac{nX}{mY} \sim F_{m,n}$$

PDF:

Step-1] Start with joint pdf of X & Y

$$\begin{bmatrix} X \\ Y \end{bmatrix}$$

consider the transformation $T = \frac{nX}{mY}$

$W = Y$

Step-2] Range $T \in [0, \infty)$
$W \in [0, \infty)$

$$J = \left\| \begin{matrix} \frac{m}{n}W & \frac{m}{n}T \\ 0 & 1 \end{matrix} \right\| = \frac{m}{n}W \qquad X = \frac{m}{n}TW$$

$$f_{\left(\begin{smallmatrix} T \\ W \end{smallmatrix}\right)}\left(\begin{pmatrix} t \\ w \end{pmatrix}\right) = f_{\left(\begin{smallmatrix} X \\ Y \end{smallmatrix}\right)}\left(\begin{pmatrix} x = m/n\, tw \\ y = w \end{pmatrix}\right) \cdot \frac{m}{n}w$$

Step-3] $$f_T(t) = \int_0^\infty f_{\left(\begin{smallmatrix} T \\ W \end{smallmatrix}\right)}\left(\begin{pmatrix} t \\ w \end{pmatrix}\right) \cdot dw$$

Properties: (1) $\frac{1}{T} \sim F_{n,m}$

(2) $E(T) = \frac{n}{n-2}$ $\qquad$ $var(T) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$

T-distribution :

Defn: let $Z \sim N(0,1)$ and $X \sim \chi^2_{(n)}$ distribution, independently,

then $T = \frac{Z}{+\sqrt{X/n}} \sim t_{(n)}$ distribution.

PROPERTIES:

(1) $T^2 \sim F_{1,n}$ distribution

(2) $t_{(1)} = $ cauchy $(0,1)$, distributn.

If $X_i \overset{iid}{\sim} N(0,1)$ then, $i=1,2$

$\frac{X_1}{|X_2|} \sim$ cauchy $(0,1)$

(3) $E(T)=0$. $n>1$

$var(T) = \frac{n}{n-2}$ $n>2$

(4) let $X_1, \dots X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

Then $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}{n-1}}} \sim t_{(n-1)}$ distributn

$$\boxed{W = \underbrace{\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}}_{N(0,1)} \cdot \frac{1}{\sqrt{\underbrace{\frac{\sum_{i=1}^{n}(X_i-\bar{X}_n)^2}{\sigma^2}}_{\chi^2_{(n-1)}/n-1} \Big/ n-1}}}$$

& $\bar{X}_n \perp\!\!\!\perp nS_n^2$

March - 11th $\longrightarrow$ QUIZ.

Mean deviation is least measured about median.

Large sample Theory:

Ex: $X_1, \ldots X_n \overset{IID}{\sim} N(\mu, \sigma^2)$

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu, \sigma^2/n)$$

Ex: $X_1, \ldots X_n \overset{IID}{\sim} N(\mu, \sigma^2)$

$$nS_n^2 \sim \sigma^2 \chi^2_{(n-1)}$$

Ex: $X_1, \ldots X_n \overset{IID}{\sim} U(0, \theta)$; $T = X_{(n)}$

$$f_T(t) = \frac{n t^{n-1}}{\theta^n}; \quad 0 < t < \theta$$

Q) How the sampling distributions behave as $n \to \infty$

(A) convergence in probability: $\{a_n\}$ $a_n \to a$ as $n \to \infty$

$\epsilon = 0.0001$

for $n \geq N_\epsilon$

$|a_n - a| < \epsilon$

$a_n = 1/n$ $a = 0$

$\epsilon = 0.01$ $N_\epsilon = 101$

$\underset{RV}{X_n} \to \underset{RV}{X}$ as $n \to \infty$

X_n: proportion of $\{H\}$ of a fair when tossed n times

$X_n \to p$

$$X_n = \begin{cases} 0 & \text{w.p } p_{n,1} \\ 1/n & \text{w.p } p_{n,2} \\ \vdots & \\ 1 & \text{w.p } p_{n,n} \end{cases}$$

$|X_n - p|$ is also random

$E|X_n - p| = e_n$; $e_n \to 0$ as $n \to \infty$

$$X_n \xrightarrow{L^1} p$$

$\sqrt{E|X_n - p|^2} = w_n$ $\quad w_n \to 0$ as $n \to \infty$, $X_n \xrightarrow{L^2} p$

Defn: (Probability convergence) Fix any $\epsilon > 0$

consider $p_n = P(|X_n - p| > \epsilon)$

If $p_n \to 0$ as $n \to \infty$, then we say that $X_n \to X$

let $\{X_n\}$ be a seq of R.V's and X be another random variable, [all in the same prob space], then $X_n \xrightarrow{P} X$ if for any $\epsilon > 0$, $\underbrace{P(|X_n - X| > \epsilon)}$ $\to 0$ as $n \to \infty$

$$P(\{\omega : |X_n(\omega) - X(\omega)| > \epsilon\})$$

$f_T(t) = \frac{nt^{n-1}}{\theta^n}, 0 < t < \theta$

$$F_T(t) = \begin{cases} 0 & \text{when } t < 0 \\ \frac{t^n}{\theta^n} & \text{when } 0 \le t \le \theta \\ 1 & \text{when } t > \theta \end{cases}$$

Example: $X_1, \dots X_n \overset{IID}{\sim} U(0, \theta)$. $T_n = X_{(n)}$

claim: $X_{(n)} = T_n \xrightarrow{P} \theta$

Proof: Fix $\epsilon > 0$, $P_n = P(|T_n - \theta| > \epsilon)$

$$= 1 - P(|T_n - \theta| \le \epsilon)$$

$$= 1 - P(-\epsilon \le T_n - \theta \le \epsilon)$$

$$= 1 - P(\theta - \epsilon \le T_n \le \theta + \epsilon)$$

$$= 1 - \{P(\theta - \epsilon \le T_n \le \theta) + \underbrace{P(\theta < T_n \le \theta + \epsilon)}_{0}\}$$

$$F(\theta + \epsilon) - F(\theta) = 1 - 1 = 0$$

$$= 1 - \{P(\theta - \epsilon \le T_n \le \theta)\}$$

$$= 1 - \{F_{T_n}(\theta) - F_{T_n}(\theta - \epsilon)\}$$

$$= \left(\frac{\theta - \epsilon}{\theta}\right)^n = \left(1 - \frac{\epsilon}{\theta}\right)^n \to 0 \quad \text{as } n \to \infty$$

$T_n \xrightarrow{P} \theta \iff T_n \xrightarrow{P} X$ when $X = \begin{cases} \theta & \text{w.p } 1 \\ 0 & \text{otherwise} \end{cases}$

$\therefore$ degenerate at θ

Defn: [consistency]

An estimator T_n of 'θ' is called consistent if conv through probability

$$T_n \xrightarrow{P} \theta$$

Ex: $X_1, X_2, \dots X_n \overset{IID}{\sim} U(0, \theta)$ Then $X_{(n)}$ is a consistent estimator of θ

Some Further Details:

Theorem [continous mapping] If $X_n \xrightarrow{P} X$ and $g: \mathbb{R} \to \mathbb{R}$ is continous then $g(X_n) \xrightarrow{P} g(X)$

Ex: $X_{(n)}^2 \xrightarrow{P} \theta^2$, in the above example.

Theorem [Markov inequality] Let X be a non-neg random variable

Then $\boxed{P(X > \epsilon) \leq \frac{E(X)}{\epsilon}}$

Proof:

$$E(x) = \int_0^\infty x f_x(x)\, dx$$

$$= \int_0^\epsilon x f_x(x)\, dx + \int_\epsilon^\infty x f_x(x)\, dx$$

$$\geq 0 + \epsilon \int_\epsilon^\infty f_x(x)\, dx = \epsilon P(X > \epsilon)$$

$$\Rightarrow P(X \leq \epsilon) \leq \frac{E[X]}{\epsilon}$$

Corollarly (chebyshev's inequality) $\mu = E(x)$

$$P(|x - \mu| > \epsilon) = P((x-\mu)^2 > \epsilon^2) \leq \frac{E((x-\mu)^2)}{\epsilon^2} = \frac{var(x)}{\epsilon^2}$$

Corollary 2: To show $\underbrace{T_n \xrightarrow{P} \theta}_{T_n \text{ is consistent for } \theta}$.

It's enough to show that $E(T_n) \to \theta$ as $n \to \infty$

and $var(T_n) \to 0$ as $n \to \infty$

$$\left.\begin{array}{l} E[T_n] \to \theta \\ \text{and } var(T_n) \to 0 \text{ as } n \to \infty \end{array}\right\} \begin{array}{l} \Rightarrow T_n \xrightarrow{P} \theta \\ \not\Leftarrow \end{array}$$

Proof: For any $\epsilon > 0$,

$$P_n = P(|T_n - \theta| > \epsilon)$$

$$= P((T_n - \theta)^2 > \epsilon^2) \leq \frac{E[(T_n - \theta)^2]}{\epsilon^2} \to (*)$$

$$T_n - \theta = T_n - E[T_n] + E[T_n] - \theta.$$

$(T_n - \theta)^2 \leq 2(T_n - E(T_n))^2 + 2(E(T_n) - \theta)^2$

Now, RHS of $(*)$ $\leq \dfrac{3\,E([T_n - E[T_n)]^2) + 3\,[E(T_n) - \theta]^2}{\varepsilon^2}$

$$= \underbrace{\frac{3 \cdot var(T_n)}{\varepsilon^2}}_{\substack{\to 0 \text{ as} \\ var(T_n) \to 0 \\ \text{when } n \to \infty}} + \underbrace{\frac{3 \cdot \{E(T_n) - \theta\}^2}{\varepsilon^2}}_{\substack{\to 0 \text{ as} \\ E[T_n] \to \theta \text{ when} \\ n \to \infty}}$$

Corollary:

[1] let $X_1, X_2, \ldots X_n$ be a random sample from $U(0,\theta)$

$E[X_{(n)}] = \frac{n}{n+1}\theta \to \theta$ as $n \to \infty$

$var[X_{(n)}] = \underbrace{\qquad\qquad}$ $\to 0$ as $n \to \infty$

[2] let $X_1, X_2, \ldots X_n$ be a random sample from some distribution with common mean μ & finite variance σ^2.

$T_n = \overline{X_n}$ $\quad E[T_n] = \mu$

$var[T_n] = \frac{\sigma^2}{n} \to 0$ as $n \to \infty$

By the sufficient condition $T_n \xrightarrow{P} \mu$

Weak law of Large numbers: (WLLN)

y let $X_1, X_2, \ldots X_n$ be IID from some distribution with mean μ & finite variance σ^2, then $\bar{X}_n \xrightarrow{P} \mu$

* let X_i be a random sample from some distribution with mean μ_i & variance σ_i^2, such that $\max \sigma_i^2 \leq M$ then if $X_1, X_2, \ldots X_n$ are independent

then $(\overline{X_n} - \overline{\mu_n}) \xrightarrow{P} 0$, $\quad \overline{\mu_n} = \frac{1}{n}\sum_{i=1}^{n} \mu_i$

as $n \to \infty$