Date:

Mo	Tu	We	Th	Fr	Sa	Su

1) Let $X_1, \dots X_n \overset{iid}{\sim} f(x_i) = \theta \gamma^{\theta} x_i^{-(\theta+1)} I(\gamma < x_i < \infty)$

where $\theta > 0 \cdot \gamma > 0$

a) $L(\theta) = \prod_{i=1}^{n} f(x_i)$

$= \prod_{i=1}^{n} \theta \gamma^{\theta} x_i^{-(\theta+1)} I(\gamma < x_i < \infty)$

$= \theta^{n} \gamma^{n\theta} \prod_{i=1}^{n} x_i^{-(\theta+1)} \prod_{i=1}^{n} I(\gamma < x_i < \infty)$

for $\hat{\gamma}$ $\because$ $x_i \in (\gamma, \infty)$

$\Rightarrow$ $\gamma < x_i$

using ordered data knowledge, i.e,
$X_{(1)} < X_{(2)} < \dots < X_{(n)}$

$\Rightarrow$ $\boxed{\hat{\gamma}_{MLE} = X_{(1)}}$

for $\hat{\theta}$, Taking $\hat{\gamma} = X_{(1)}$

Date: Mo Tu We Th Fr Sa Su

Notes

$$L(\theta) = \theta^n (X_{(1)})^{n\theta} \left(\prod_{i=1}^{n} x_i\right)^{-(\theta+1)}$$

$$l(\theta) = n\ln\theta + n\theta\ln(X_{(1)}) - (\theta+1)\ln\left(\prod_{i=1}^{n} x_i\right)$$

$$\Rightarrow l(\theta) = n\ln\theta + n\theta\ln(X_{(1)}) - (\theta+1)\sum_{i=1}^{n}\ln x_i$$

$$\frac{\partial l(\theta)}{\partial\theta} = \frac{n}{\theta} + n\ln(X_{(1)}) - \sum_{i=1}^{n}\ln x_i = 0$$

$$\Rightarrow \frac{n}{\theta} = \sum_{i=1}^{n}\ln x_i - n\ln(X_{(1)})$$

$$\Rightarrow \frac{n}{\theta} = \ln\left(\prod_{i=1}^{n} x_i\right) - \ln(X_{(1)})^n$$

$$\Rightarrow \frac{n}{\theta} = \ln\left(\prod_{i=1}^{n} x_i / (X_{(1)})^n\right)$$

$$\Rightarrow \boxed{\hat{\theta}_{MLE} = \frac{n}{\ln\left(\prod_{i=1}^{n} x_i / (X_{(1)})^n\right)}}$$

Date:

Mo	Tu	We	Th	Fr	Sa	Su

For verification $\frac{\partial^2 l(\theta)}{\partial \theta^2} < 0$

$$\frac{\partial^2 l(\theta)}{\partial \theta^2} = \frac{-n}{\hat{\theta}^2} + 0 - 0$$

$$\Rightarrow \frac{-n}{\hat{\theta}^2} < 0 \quad (\because \hat{\theta}^2 > 0)$$

$\Rightarrow$ maxima

$\Rightarrow \quad \hat{\gamma}_{MLE} = X_{(1)}$ and

$$\hat{\theta}_{MLE} = \frac{n}{\ln\left(\prod_{i=1}^{n} x_i / (X_{(1)})^n\right)}$$

(b) equal tailed α level LRT for

$H_0: \theta = 1$ v/s $H_1: \theta \neq 1$

LRT : $\lambda(x) = \dfrac{\max_{\theta \in H_0} L(\underline{\theta})}{\max_{\theta \in H} L(\underline{\theta})} = \dfrac{\text{restricted MLE}}{\text{unrestricted MLE}}$

$= \dfrac{L(\theta)\big|_{\theta=1}}{L(\hat{\theta})}$, $\hat{\theta}$ MLE

for $\theta = 1$, $L(\theta) = \prod_{i=1}^{n} (1)\, r^{(1)}\, x_i^{-1-1}\, I(r < x_i < \infty)$

$= \prod_{i=1}^{n} \dfrac{r}{x_i^2}\, I(r < x_i < \infty)$

using ordered data knowledge,
$X_{(1)} < X_{(2)} < \dots < X_{(n)}$

$\Rightarrow \hat{r}_{MLE} = X_{(1)}$

Date: Mo Tu We Th Fr Sa Su

for $\theta \in H$

$\Rightarrow \hat{\gamma}_{MLE} = X_{(1)}$ and $\hat{\theta}_{MLE} = \dfrac{n}{\ln\left(\prod_{i=1}^{n} x_i / X_{(1)}\right)}$

$$\lambda(x) = \frac{L(\theta = 1, \gamma = X_{(1)})}{L\left(\theta = \dfrac{n}{\ln\left(\prod_{i=1}^{n} x_i / (X_{(1)})^n\right)}, \gamma = X_{(1)}\right)}$$

$$= \frac{\prod_{i=1}^{n} (1)(X_{(1)})^1 x_i^{-2}}{\prod_{i=1}^{n} \hat{\theta}_{MLE} X_{(1)}^{\hat{\theta}_{MLE}} x_i^{-(\hat{\theta}_{MLE}+1)}}$$

$$= \frac{(X_{(1)})^n \prod_{i=1}^{n} x_i^{-2}}{(\hat{\theta}_{MLE})^n X_{(1)}^{n\hat{\theta}_{MLE}} \prod_{i=1}^{n} x_i^{-(\hat{\theta}_{MLE}+1)}}$$

Date: Mo Tu We Th Fr Sa Su

Notes

we reject H_0 if $\lambda(x) < c$ $\quad c \in (0,1)$

$$\Rightarrow \frac{(X_{(1)})^n \prod_{i=1}^{n} x_i^{-2}}{(\hat{\theta}_{MLE})^n (X_{(1)})^{n\hat{\theta}_{MLE}} \prod_{i=1}^{n} x_i^{-(\hat{\theta}_{MLE}+1)}} \leq c$$

where $\hat{\theta}_{MLE} = \dfrac{n}{\ln\left(\prod_{i=1}^{n} x_i / (X_{(1)})^n\right)}$

Date:

Mo	Tu	We	Th	Fr	Sa	Su

3)

Before	After	diff = d
30.5	23	7.5
18.5	21	-2.5
24.5	22	2.5
32	28.5	3.5
16	14.5	1.5
15	15.5	-0.5
23.5	24.5	-1
25.5	21	4.5
28	23.5	4.5
18	16.5	1.5
		21.5

$$\frac{\Sigma d}{n} = \frac{21.5}{10} = 2.15$$

$$\bar{d} = 2.15$$

$$S_d = \sqrt{\frac{\Sigma(d-\bar{d})^2}{n-1}} = \sqrt{\frac{81.025}{9}}$$

$$\underline{S_d \sim 3}$$

Hypothesis : $H_0: \mu_d \leq 0$ v/s $H_1: \mu_d > 0$

(right tailed test)

where μ_d = before value - after value

Assumptions :

- SRS (given)
- diff of the population comes from normal
- samples are dependent
- popⁿ variances of the difference is unknown.

We apply the T-test

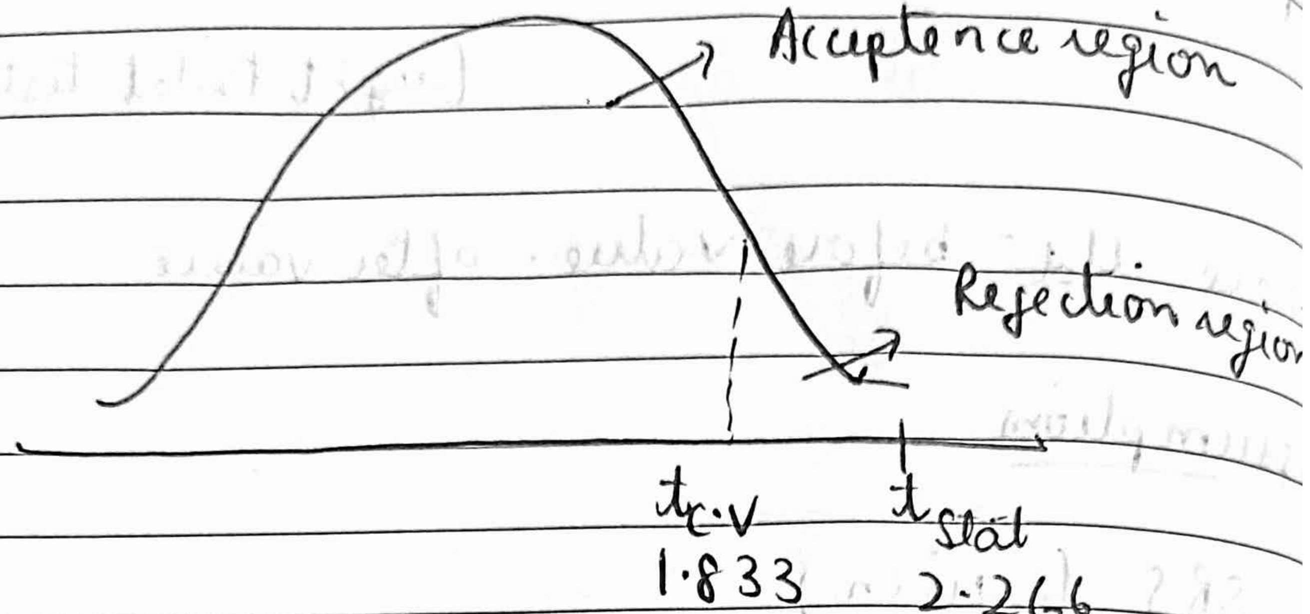
$$T_{stat} = \frac{\bar{d} - \mu_d}{S_d/\sqrt{n}} \sim t_{(n-1)}$$

$$= \frac{2.15 - 0}{3/\sqrt{10}} = 2.266$$

Date:

Mo	Tu	We	Th	Fr	Sa	Su

$T_{c.v} = t_{9,0.05} = 1.833$



Right tailed test

We see that $T_{stat} > T_{c.v}$, it lies in the rejection region

⇒ We can reject H_0.

⇒ Hence, the safety program was effective.

Ans 2)

Here, $X_1,\ X_2,\dots,X_n$ are iid from $N\left(\mu_1, \sigma^2\right)$ and $Y_1,\ Y_2,\dots,Y_n$ are iid from $N\left(\mu_2, 3\sigma^2\right)$.
X and **Y** are independent. Also, $-\ \infty < \mu_1, \mu_2 < \infty;\ 0 < \sigma < \infty$.

Using the properties of normal distribution, we know that the difference of two independent normal random variables is also normal with mean equal to the difference in means and variance equal to the sum of variances. Also, A linear combination of normally distributed random variables is also normally distributed

We can start by finding the distribution of $(\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} Y_i)$. Since $X_1,\ X_2,\dots,X_n$ and $Y_1,\ Y_2,\dots,Y_n$ are independent, we have:

$$Var\left[\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} Y_i\right] = Var\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] + Var\left[\frac{1}{n}\sum_{i=1}^{n} Y_i\right]$$
$$= \frac{\sigma^2}{n} + \frac{3\sigma^2}{n}$$
$$= \frac{4\sigma^2}{n}$$

$$E\left[\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} Y_i\right] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] - E\left[\frac{1}{n}\sum_{i=1}^{n} Y_i\right] = \mu_1 - \mu_2$$

Let, $A = \frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} Y_i$

Therefore, $A \sim N\left(\mu_1 - \mu_2, \frac{4\sigma^2}{n}\right)$, because as A *is* linear combination of normally distributed random variables $X_1,\ X_2,\dots,X_n$ and $Y_1,\ Y_2,\dots,Y_n$, therefore A is also normally distributed.

Next, we can standardize A using the variable $Z = \frac{A-\left(\mu_1-\mu_2\right)}{2\sigma/\sqrt{n}}$
Therefore,
$Z = \left(\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} Y_i - \left(\mu_1 - \mu_2\right)\right)/(2\sigma/\sqrt{n})$. Then, $V_n = Z$.

Here, Z is standardized variable, therefore $Z \sim N(0, 1)$. Therefore, $V_n \sim N(0, 1)$ as well.

So, the distribution of V_n is a standard normal distribution, N(0,1).

classmate

Date
Page

Q.4)

$x_i \overset{iid}{\sim} N(\underbrace{\mu, \sigma^2}_{\text{unknown}})$

$\bar{x} = 9 \quad ; n = 9 ; \quad s^2 = 9.5$

$$-T_{\alpha/2} \le T_{stat} \le T_{\alpha/2}$$

$$-T_{\alpha/2} \le \frac{\bar{X} - \mu}{S/\sqrt{n}} \le T_{\alpha/2}$$

$$\bar{X} - \frac{T_{\alpha/2} S}{\sqrt{n}} \le \mu \le \bar{X} + \frac{T_{\alpha/2} S}{\sqrt{n}} \quad (1)$$

$\alpha = 1 - 0.95 = 0.05 \quad ; \alpha/2 = 0.025$

Substituting values :-

$$9 - T_{\alpha/2} \sqrt{\frac{9.5}{9}} \le \mu \le 9 + T_{\alpha/2} \sqrt{\frac{9.5}{9}}$$

$$T_{8, 0.025} = 2.306$$

in (1)

$$9 - 2.363 \le \mu \le 9 + 2.306 \sqrt{\frac{9.5}{9}}$$

$$\mu \in [6.637, 11.363]$$

Hence, μ is between $(6.637, 11.363)$

by graph:-

$-t_{\alpha/2} = -2.306$ $\qquad$ $t_{\alpha/2} = t_{0.025,8} = 2.306$

Q5

ans) Let μ_C = mean time a cold lasts when vitamin C tablets are taken

μ_P = mean time a cold lasts when a placebo is taken

We'll be testing

$$H_0: \mu_P \leq \mu_C \quad v/s \quad H_1: \mu_P > \mu_C$$

<u>Assumptions</u>
1. pop^n is normally distributed
2. the sample is SRS
3. pop^n variance is unknown
4. the samples are independent of each other

Also,

$n_1 = 10$ $\qquad$ $n_2 = 12$

$\bar{X} = 6.450$ $\qquad$ $\bar{Y} = 7.125$

$S_x^2 = 0.581$ $\qquad$ $S_y^2 = 0.778$

$\alpha = 0.05$

Page

Based on info. provided and assumptions met, it is clear that it is a 2-sample test about mean.

The samples are independent of each other & pop^n variances are unknown

So,

$$t_{stat} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim t_{d.f}$$

$$= \frac{(6.450 - 7.125) - (0)}{\sqrt{\frac{0.581}{10} + \frac{0.778}{12}}} \quad (\because (\mu_1 - \mu_2) = 0)$$

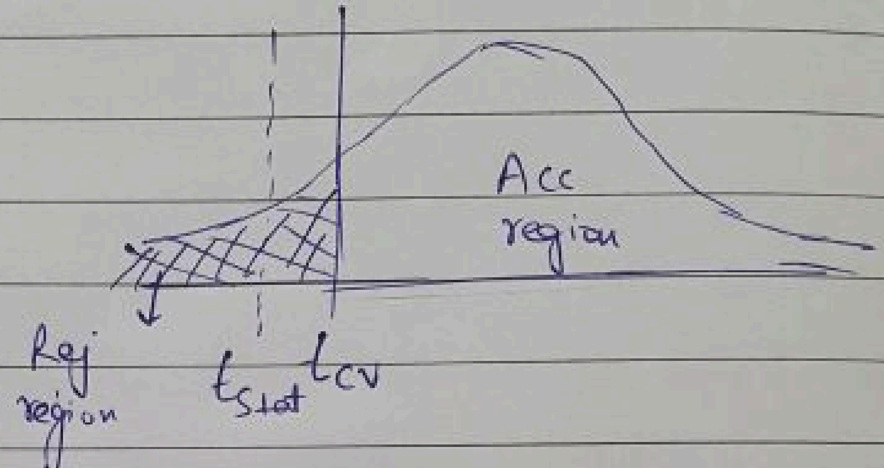
$$= \frac{-0.675}{\sqrt{0.122}} = -1.92$$

Also

$$t_{df} \sim t_{n+m-2,\alpha} = t_{20, 0.05} = -1.725$$

Comparing, we see

$$t_{c.v} > t_{stat}$$



So, t_{stat} lies in rejection region

$\therefore$, null hypothesis is <u>rejected</u>

We infer that there is sufficient evidence to conclude that vitamin C tablets reduce mean time that a cold persists.

classmate

Date ______
Page ______

Q6

ans) A two-way contingency table is given and we have to test the hypothesis that smoking and lung cancer are independent. So,

H_0 : Smoking and lung cancer are independent

v/s

H_1 : Smoking and lung cancer are not independent

Assumptions ① ~~Data is randomly chosen~~

② Data is rep. in a 2-way table and there are 2 rows & 2 columns & all of them have finite frequencies

③ $E_i \geq 5 \quad \forall\, i = 1, 2, 3, 4$ (table provided below)

	Smokers	Non-smokers	Total
Lung Cancer	62	14	76
No lung cancer	9,938	19,986	29,924
Total	10,000	20,000	30,000

Also, $\alpha = 0.01$ (given)

classmate

Date

Page

$$E_1 = \frac{(76)(10,000)}{30,000} = 25.33$$

$$E_2 = \frac{(76)(20,000)}{30,000} = 50.67$$

$$E_3 = \frac{(29,924)(10,000)}{30,000} = 9,974.67$$

$$E_4 = \frac{(29,924)(20,000)}{30,000} = 19,949.33$$

Now,

$$\chi^2_{stat} = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i}$$

$$= \frac{(62-25.33)^2}{25.33} + \frac{(14-50.67)^2}{50.67} + \frac{(9938-9974.67)^2}{9974.67} + \frac{(19,986-19,949.33)^2}{19,949.33}$$

$$= 79.83$$

$$\chi^2_{C.V} = \chi^2_{(r-1)(c-1),\,0.01} = \chi^2_{1,0.01} = 6.635$$

We see

$$\chi^2_{Stat} > \chi^2_{C.V} \text{ \& } \chi^2_{Stat} \text{ lies in rejection region}$$

<u>Inference</u>

At $\alpha = 0.01$, $\chi^2_{Stat} > \chi^2_{C.V}$ & χ^2_{Stat} lies in the rejection region. Thus, we <u>reject</u> the null hypothesis and infer that smoking and lung cancer are not independent of each other.

Q7

ans) We are interested in testing

$H_0 : m = 3.7 \quad v/s \quad H_1 : m \neq 3.7$

<u>Assumptions</u> ① sample is random

② data is continuous

③ non-symmetric distr^n/data

Taking into acc. the assumptions and the fact that we are testing about median, we'll be using Wilcoxon Sign Rank Test

The data →

X_i	$X_i - m_0$	$\lvert X_i - m_0 \rvert$	Rank	Sign
5	1.3	1.3	5	+
3.9	0.2	0.2	1	+
5.2	1.5	1.5	6	+
5.5	1.8	1.8	7	+
2.8	-0.9	0.9	3	−
6.1	2.4	2.4	9	+
6.4	2.7	2.7	10	+
2.6	-1.1	1.1	4	−
1.7	-2.0	2.0	8	−
4.3	0.6	0.6	2	+

Page

$W^+ = 5 + 1 + 6 + 7 + 10 + 9 + 2 = 40$
$W^- = 3 + 4 + 8 = 15$

$W_{stat} = \min(W^+, W^-)$
$= 15$

Also,

$W_{c.v.} = 8$

Since

$W_{stat} > W_{c.v}$

We don't reject H_0. At $\alpha = 0.05$, there is no evidence to conclude that the median length of pygmy sunfish differs significantly from 3.7 cm.

Ans 8)

(a)

Source	df	Sum of Squares (SS)	Mean Sum (MS)	F	p-value
Treatment					0.0195
Error	118	29314.047			
Total		31336.777			

To fill in the missing values in the table, we need to use the following formulas:

- Total sum of squares (SST) = sum of squares for treatment (SSTR) + sum of squares for error (SSE)
- Degrees of freedom (df) = number of groups (k) - 1 for treatment, total sample size (n) - k for error, and n - 1 for total
- Mean sum of squares (MS) = sum of squares (SS) / degrees of freedom (df)
- F-ratio (F) = MS for treatment / MS for error

We are given the values for SST and SSE, and we can calculate df for error and total as follows:

df(error) = n - k = 121 - 3 = 118
df(total) = n - 1 = 121 - 1 = 120
df(treatment) = k - 1 = 3 - 1 = 2

Using these values, we can calculate MS for error as follows:

MS(error) = SSE / df(error) = 29314.047 / 118 = 248.424

To find SSTR, we can use the formula:

SSTR = SST - SSE

SSTR = 31336.777 - 29314.047 = 2022.73

And MS for treatment can be calculated as:

MS(treatment) = SSTR / df(treatment) = 2022.73 / 2 = 1011.365

Finally, we can calculate the F-ratio:

F = MS(treatment) / MS(error) = 1011.365 / 248.424 = 4.071

(b)
To test the claim that the three samples come from populations with means that are all equal, we can use an analysis of 1-way ANOVA test with a significance level of $\alpha = 0.05$.

H_0: means of the three populations are equal i.e $\mu_1 = \mu_2 = \mu_3$, where μ_1, μ_2 and μ_3 are the means of the populations for the low, medium and high blood lead levels respectively, versus
H_1: at least one mean is different from the others.

Assumptions:-
i) All the populations have a normal distribution.
ii) Populations variances are equal.
iii) Samples are SRS.
iv) Independent samples (all the 3 samples are independent of each other)

To determine whether to reject or fail to reject the null hypothesis, we need to compare the F-ratio from the ANOVA test to the critical value from the F-distribution with k - 1 and n - k degrees of freedom at the α level of significance. In this case, k = 3 and n = 121, so df(treatment) = 2 and df(error) = 118.

From the F-distribution table, the critical value for F with df(treatment) = 2 and df(error) = 118 at α = 0.05 is 3.08.

Since the calculated F-ratio of 4.071 is greater than the critical value of 3.08, we reject the null hypothesis and conclude that there is evidence to suggest that at least one of the population means is different from the others.

Alternative method using p-value:-

Corresponding to F-ratio = 4.071, df-numerator (df-treatment) = 2, df-denominator (df-error) = 18 and significance level 0.05, we get the p-value to be 0.0195.

Since this p-value is less than the significance level of 0.05, we reject the null hypothesis and conclude that there is evidence to suggest that at least one of the population means is different from the others.

Q.9) Model equation: $y_i = \beta_0 + \beta_1 x_i + \epsilon_i \; ; \epsilon_i \sim N(0, \sigma^2)$

To find:- an unbiased estimator of β_0.

Ordinary (OLSS) .. Finding LSE :-

$$y_i = \underbrace{\beta_0 + \beta_1 x_i}_{\text{Systematic}} + \epsilon_i \rightarrow \text{random}$$

$$\epsilon_i = y_i - (\beta_0 + \beta_1 x_i)$$

$$\min \sum_{i=1}^{n} \epsilon_i^2 = \min \sum_{i=1}^{n} (y_i - \beta_0 + \beta_1 x_i)^2$$

as $\epsilon_i \in N(0, \sigma^2)$

$$E(y_i / x_i) = \beta_0 + \beta_1 x_i$$

$$\sum \epsilon_i^2 = \sum (y_i - \beta_0 - \beta_1 x_i)^2$$

Differentiating & equating to zero :-

$$\frac{dl}{d\beta_0} = 2\sum(y_i - \beta_0 - \beta_1 x_i)(-1) = 0$$

$$\Rightarrow \sum y_i - n\beta_0 - \beta_1 \sum x_i = 0 \quad (a)$$

$$\frac{dl}{d\beta_1} = 2\sum(y_i - \beta_0 - \beta_1 x_i)(-x_i) = 0$$

$$\Rightarrow \sum(y_i - \beta_0 - \beta_1 x_i)(x_i) = 0$$

$$\Rightarrow \sum y_i x_i - \beta_0 \sum x_i - \beta_1 \sum x_i^2 = 0 \quad (b)$$

Multiplying (a) by $\sum x_i$ and (b) by n.

$$\Rightarrow (\sum y_i - n\beta_0 - \beta_1 \sum x_i = 0) \sum x_i$$

$$(\sum y_i x_i - \beta_0 \sum x_i - \beta_1 \sum x_i^2 = 0)\, n$$

$$\Rightarrow \sum x_i \sum y_i - n\beta_0 \sum x_i - \beta_1 (\sum x_i)^2 = 0$$

$$\underset{(-)}{n\sum y_i x_i} \quad \underset{(+)}{- \beta_0 n \sum x_i} \quad \underset{(+)}{- n\beta_1 \sum x_i^2} = 0$$

$$\sum x_i \sum y_i - n\sum x_i y_i = \beta_1 \left[(\sum x_i)^2 - n\sum x_i^2\right]$$

$$\hat{\beta}_1 = \frac{\sum x_i \sum y_i - n\sum x_i y_i}{(\sum x_i)^2 - n\sum x_i^2}$$

$$\sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \bar{y} n \bar{x} - n \bar{x} \bar{y} + n \bar{x} \bar{y}$$

$$= \sum x_i y_i - n \bar{x} \bar{y}$$

$$\sum (x_i - \bar{x})^2 = \sum x_i^2 + n \bar{x}^2 - 2n\bar{x}^2$$

$$= \sum x_i^2 - n \bar{x}^2$$

$$\hat{\beta}_1 = \frac{n \sum x_i y_i - n^2 \bar{x} \bar{y}}{n \sum x_i^2 - n^2 \bar{x}^2}$$

$$\boxed{\hat{\beta}_1 = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{\sum x_i^2 - n \bar{x}^2} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}}$$

$$= \frac{S_{xy}}{S_{xx}} = \frac{\operatorname{cov}(x, y)}{\operatorname{var}(x)}$$

r_{xy} :- sample correlation coefficient.

$$= r_{xy} \frac{s_y}{s_x}$$

$$\boxed{\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}}$$

$$\hat{\beta}_1 = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$

Numerator of RHS is

$$\sum x_i y_i - n\bar{x}\bar{y}$$

$$= \sum x_i y_i - \bar{x}\sum y_i$$

$$= \sum (x_i - \bar{x}) y_i$$

Now, expectation of $\hat{\beta}_1$'s numerator is

$$E\{\sum(x_i - \bar{x}) y_i\} = \sum (x_i - \bar{x}) E(y_i)$$

$$= \sum (x_i - \bar{x})(\beta_0 + \beta_1 x_i)$$

$$= (\beta_0 \sum x_i - n\bar{x}\beta_0) + \beta_1 \sum x_i^2 - n\bar{x}^2\beta_1$$

$$= \beta_1 (\sum x_i^2 - n\bar{x}^2)$$

Finally, $$E(\hat{\beta}_1) = \frac{E\{\sum(x_i - \bar{x}) y_i\}}{\sum x_i^2 - n\bar{x}^2}$$

$$= \frac{\beta_1\left(\sum x_i^2 - n\bar{x}^2\right)}{\sum x_i^2 - n\bar{x}^2}$$

$$= \beta_1$$

now, $E(\hat{\beta}_0) = E(\bar{y} - \hat{\beta}_1 \bar{x})$

$$= \frac{1}{n} \sum E(y_i) - E(\hat{\beta}_1)\bar{x}$$

$$= \frac{1}{n} \sum [\beta_0 + \beta_1 x_i] - \beta_1 \bar{x}$$

$$= \frac{1}{n} [n\beta_0 + n\beta_1 \bar{x}] - \beta_1 \bar{x}$$

$$= \beta_0$$

(10)

Let $T = X$

$\Rightarrow U = \frac{X}{\theta} \quad \Rightarrow \quad U\theta = X \Rightarrow \left|\frac{dx}{du}\right| = |\theta| = \theta = J.$

$f_X(x) = 2\left(\frac{\theta - x}{\theta^2}\right) \quad ; \quad 0 < x < \theta$

at $x = 0$; $u = 0$
at $x = \theta \Rightarrow u = 1$

$\Rightarrow f_U(u) = 2\left(\frac{\theta - u\theta}{\theta^2}\right) \cdot \theta$

$= 2(1-u) \quad ; \quad 0 < u < 1$

Thus, $U = \frac{X}{\theta}$ is a pivot quantity

Now, for CI

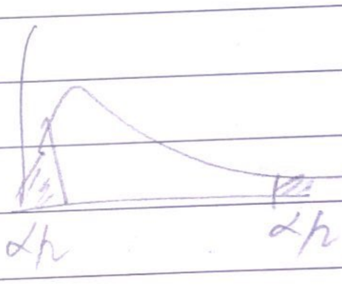
$P(a \leq Q(T,\theta) \leq b) = 1-\alpha$

$P(a \leq U \leq b) = 1-\alpha$

$P(a \leq \frac{X}{\theta} \leq b) = 1-\alpha$

$P\left(\frac{a}{X} \leq \frac{1}{\theta} \leq \frac{b}{X}\right) = 1-\alpha$

$\Rightarrow P\left(\frac{X}{b} \leq \theta \leq \frac{X}{a}\right) = 1-\alpha$



For equal areas of rejection on both sides we consider

$P\left(\frac{X}{\alpha/2} \leq \theta \leq \frac{X}{1-\alpha/2}\right) = 1-\alpha$