Problem 1:

The time to failure for an electric component used in a flat panel display unit is satisfactory modelled by Weibull distribution $\beta = 1/2$ and $\theta = 5000$. Find the mean time to failure and fraction of components that are expected to survive 20000 hrs.

Objective:- To find the mean time to failure and fraction of components that are expected to survive 20000hrs.

Theory:- The Weibull distribution is most widely used lifetime distribution in reliability engineering and analysis. It is a versatile distribution that can take on the characteristics of other types of distribution.
Weibull Distribution is given as –

$$f(t) = \frac{\beta}{\theta}\left(\frac{t}{\theta}\right)^{\beta-1} e^{-(t/\theta)^{\beta}}$$

where β is shape parameter and θ is scale parameter.

Methodology:- Mean of Weibull distribution = $\theta \times \Gamma(1/\beta + 1)$. To find the mean, we create a mean function in R programming. To find the probability of components that are expected to survive 20000 hrs, we will use
pweibull from the package *stats*.

R-Code:-

```
rm(list=ls())
beta=0.5
theta=5000
1-pweibull(20000,beta,theta)
MTTF=5000*exp(log(1/beta));MTTF
```

R-Console

```
> rm(list=ls())
> beta=0.5
> theta=5000
> 1-pweibull(20000,beta,theta)
[1] 0.1353353
> MTTF=5000*exp(log(1/beta));MTTF
[1] 10000
>
```

Conclusion:- From the above experiment we conclude that the mean lifetime of failure =10000 and probability that lifetime exceeds 2000 is 0.135353.

Problem 2

50 units are subjected to a test and the test is terminated when 35 units fail. Their lifetimes (in weeks) are given below:

22.30	26.80	30.30	31.90	32.10	33.30	33.70	33.90	34.70
36.10	36.40	36.50	36.60	37.10	37.60	38.20	38.50	38.70
38.70	38.90	38.90	39.10	41.10	41.10	41.10	42.40	43.60
43.80	44.00	45.30	45.80	50.40	51.30	51.40	51.50	

Assume Lognormal distribution and estimate the two parameters of the distribution. Also estimate the mean time to failure, median time failure and standard deviation of time to failure.

Objective:- To estimate the parameters of lognormal distribution and mean, median, standard deviation to failure.

Theory: - A random variable is log-normally distributed if the logarithm of the random variable is normally distributed. Lognormal distribution has two parameters μ and σ. The probability density function of lognormal is –

$$f(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}x}\exp\left[-\frac{(\ln x-\mu)^2}{2\sigma^2}\right]; \ \ where \qquad x>0,\sigma>0$$
$$and \ -\infty<\mu<\infty$$

where μ is the mean of the natural logarithms of the times to failure (lifetimes of components) and σ is the standard deviation of the natural logarithms of the times to failure.
If X represents lifetimes of components or times to failure of components, then –

- Mean of $X = \exp\left(\mu + \frac{\sigma^2}{2}\right)$
- Variance of $X = [\exp(\sigma^2)-1]\times\exp(2\mu+\sigma^2)$
- S.D of $X = \sqrt{Var(X)}$
- Median of $X = \exp(\mu)$

Methodology: - If X represents times to failure and $Y = \log(x)$ follows normal distribution then we will take log of X and assign the values in Y and calculate μ, σ; to calculate mean, median, sd time to failure.

R-Code:-

```
rm(list=ls())
units =
c(22.3,26.8,30.3,31.9,32.1,33.3,33.7,33.9,34.7,36.1,36.4,36.5,36.6,37.1,37.6,38.2,38.5,38.7,38.7,38.9,38.9,39.1,41.1,41.1,41.1,42.4,43.6,43.8,44,45.3,45.8,50.4,51.3,51.4,51.5)
y = log(units)
mu_cap = mean(y);mu_cap
sig_cap = sqrt(var(y));sig_cap
mean_time_to_fail<-exp(mu_cap + (sig_cap**2)/2);mean_time_to_fail
standard_tf<-sqrt((exp((sig_cap**2)-1))*exp(2*mu_cap + sig_cap**2));standard_tf
median_tf<-exp(mu_cap);median_tf
```

R-Console:-

```
> rm(list=ls())
> units =
c(22.3,26.8,30.3,31.9,32.1,33.3,33.7,33.9,34.7,36.1,36.4,36.5,36.6,37.1,37.6,38.2,38.5,38.7,38.7,38.9,38.
9,39.1,41.1,41.1,41.1,42.4,43.6,43.8,44,45.3,45.8,50.4,51.3,51.4,51.5)
> y = log(units)
> mu_cap = mean(y);mu_cap
[1] 3.647185
> sig_cap = sqrt(var(y));sig_cap
[1] 0.1788436
> mean_time_to_fail<-exp(mu_cap + (sig_cap**2)/2);mean_time_to_fail
[1] 38.98503
> standard_tf<-sqrt((exp((sig_cap**2)-1))*exp(2*mu_cap + sig_cap**2));standard_tf
[1] 24.02681
> median_tf<-exp(mu_cap);median_tf
[1] 38.36652
>
```

Conclusion:- From the above experiment we conclude that Mean time to failure= 38.98503, Median time to failure =38.36652, Standard time to failure=24.02681 and Parameters of lognormal distribution $\mu = 3.647185; \sigma = 0.1788436$.

Problem: 3

The life time of a medical laser in orthopedic surgery have log-normal distribution $\theta = 6$ and $\omega = 1.2$ hrs. What is the probability that life exceeds 500 hrs.

Objective: - We must find the probability that lifetime of medical laser exceeds 500hrs.

Theory:- A random variable is log-normally distributed if the logarithm of the random variable is normally distributed. Lognormal distribution has two parameters μ and σ. The probability density function of lognormal is –

$$f(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}x}\exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right]; \ where \qquad x > 0, \sigma > 0$$

$$and \ -\infty < \mu < \infty$$

where μ is the mean of the natural logarithms of the times to failure (lifetimes of components) and σ is the standard deviation of the natural logarithms of the times to failure.

If X represents lifetimes of components or times to failure of components, then –

- Mean of $X = \exp\left(\mu + \frac{\sigma^2}{2}\right)$
- Variance of $X = [\exp(\sigma^2) - 1] \times \exp(2\mu + \sigma^2)$
- S.D of $X = \sqrt{Var(X)}$

Methodology: - The log normal distribution has density

$$f(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}x}\exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right]$$

where μ and σ are the mean and standard deviation of the logarithm. The mean is $\mathbb{E}(X) = \exp(\mu + \sigma^2/2)$ and the variance $Var(X) = \exp(2 * \mu + \sigma^2) * (\exp(\sigma^2) - 1)$. we will use **plnorm** function from package **stats**.

R-Code:-

```
rm(list=ls())
x=log(500)
mu=6
sg=1.2
prob=1-pnorm((x-mu)/sg);prob
```

R-Console:-

```
> rm(list=ls())
> x=log(500)
> mu=6
> sg=1.2
> prob=1-pnorm((x-mu)/sg);prob
[1] 0.4290316
>
```

Conclusion:- From the above experiment we conclude that probability that lifetime of medical laser exceeds 500hrs. is 0.4290316.

Problem: 4

Ten bearings made by a certain process have a mean diameter of 0.506 cm with standard deviation 0.004. Assuming that data can be regarded as a random sample from a normal population with mean μ and variance σ2. Construct 95% confidence interval for the actual diameter of the bearing made by the process.

Objective: - To construct 95% confidence interval for mean.

Theory: - A confidence interval (CI) is a type of estimate computed from the statistics of the observed data. This gives a range of values for an unknown parameter (for example, a population mean). The interval has an associated confidence level that gives the probability with which an estimated interval will contain the true value of the parameter.
Formula for the confidence interval is-

$$CI = \bar{x} \pm \mathbb{Z}_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

Where $\bar{x}$ is sample mean, n is sample size, s is sample standard deviation, α is the level of significance (in this case $\alpha = 1 - 95\% = 5\%$) and $\mathbb{Z}$ is the critical value for $\alpha/2$ level of significance.

Methodology: - We will create some functions in r studio and do some algebraic calculations to find required result. We will use **qnorm()** function from the package **stats** to find the $\mathbb{Z}$ score for 95% interval.

R-Code:-

```
rm(list=ls())
mue=0.506
stand_dev=0.004
m_l=mue-qnorm(0.975)*(stand_dev^2/10);m_l
m_u =mue+qnorm(0.975)*(stand_dev^2/10);m_u
```

R-Console:-

```
> rm(list=ls())
> mue=0.506
> stand_dev=0.004
> m_l=mue-qnorm(0.975)*(stand_dev^2/10);m_l
[1] 0.5059969
> m_u =mue+qnorm(0.975)*(stand_dev^2/10);m_u
[1] 0.5060031
>
```

Conclusion:- From the above experiment we can conclude that mean diameter will lies in (0.5059969 , 0.5060031) with 95% confidence.

Problem: 5

Suppose that the lifetime distribution of a certain group of experimental mice can be approximated by a normal population with mean μ and variance σ^2 truncated from below at point *A*.

(i) Find the survival function $S(x|X > A)$ and the hazard rate function $\lambda(x|X > A)$.

(ii) Derive formulae for the mean lifetime and variance of this truncated distribution.

Objective:- To find the survival function $S(x|X > A)$ and the hazard rate function $\lambda(x|X > A)$ and Derive formulae for the mean lifetime and variance of this truncated distribution.

Theory:- Suppose that a random variable *X* denotes the lifetime of a living Organism or an inanimate device, it is also called as at death or at failure or briefly age.

The cumulative distribution function (CDF); $F_X(x) = P_t(X \leq x)$; is called the lifetime distribution or failure distribution.

The complementary function of cumulative distribution function *(CDF)* is known as survival distribution function *(SDF)*. And *SDF = 1 – CDF*.

$$SDF = 1 - P_t(X \leq x) = P_t(X > x)$$

Hazard Function is the instantaneous or relative failure at time point x. And given as

$$\begin{aligned}
&\lambda_X(x) = \frac{f_X(x)}{S_X(x)} \\
&\Rightarrow \lambda_X(x) = -\frac{\mathrm{d}}{\mathrm{d}x}(\log S_X(x)) \\
&\Rightarrow \mathrm{d}\log S_X(x) = \lambda_X(x)\mathrm{d}x \\
&\Rightarrow -\log S_X(x) = \int_0^x \lambda_X(x)\mathrm{d}x
\end{aligned}$$

Methodology/solution:-

Let X is a random variable which represents the lifetime distribution of a certain group of experimental mice and given that $X \sim \mathbb{N}(\mu, \sigma^2)$ and truncated from below.

A truncation distribution is a conditional distribution that results from restricting the domain of some other probability distribution(In this case normal distribution).

In general, the truncated SDF is given as follows :

$$S_X(x|a < X \leq b) = \begin{cases} 1 & \text{for } x \leq a \\ \frac{S_X(x) - S_X(b)}{S_X(a) - S_X(b)} & \text{for } a < x \leq b \\ 0 & \text{for } x > b \end{cases}$$

but in this case, we are using only truncation from below. Therefore, we take b=∞ so $S_X(b) = 0$

therefore,

$$S_X(x|x > A) = \begin{cases} I & \text{for } x \leq a \\ \frac{S_X(x)}{S_X(a)} & \text{for } x > a \end{cases}$$

where, $S_X(x) = 1 - F_X(x)$

For normal distribution,

$$F_X(x) = \frac{1}{2}\left[1 + erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$$

therefore,

$$S_X(x) = \frac{1}{2} - \frac{1}{2}erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)$$

and,

$$S_X(a) = \frac{1}{2} - \frac{1}{2}erf\left(\frac{a-\mu}{\sigma\sqrt{2}}\right)$$

and the hazard rate function,

$$\lambda_X(x|a < X \leq b) = \frac{F_X(x)}{S_X(x) - S(b)}$$

For truncation from below, S(b) = 0

$$\lambda_X(x|x > A) = \frac{F_X(x)}{S_X(x)}$$

where,

$F_X(x) = \dfrac{1}{\sigma\sqrt{2\pi}}e^{-\dfrac{1}{2}\left(\dfrac{x-\mu}{\sigma}\right)^2}$

And, $S_X(x) = \dfrac{1}{2} - \dfrac{1}{2}erf\left(\dfrac{x-\mu}{\sigma\sqrt{2}}\right)$

(ii) **Mean and variances :-**

let $\alpha = \dfrac{a-\mu}{\sigma}$ and $\beta = \dfrac{b-\mu}{\sigma}$ then,

$$E(X|a < X < b) = \mu + \frac{\phi(\alpha) - \phi(\beta)}{\Phi(\beta) - \Phi(\alpha)}$$

And,

$$V(X|a < X < b) = \sigma^2\left[1 + \frac{\alpha\phi(x) - \beta\phi(\beta)}{\Phi(\beta) - \Phi(\alpha)} - \left(\frac{\phi(\alpha) - \phi(\beta)}{\Phi(\beta) - \Phi(\alpha)}\right)^2\right]$$

Where $\Phi(.)$ is the standard CDF of normal distribution

$$\Phi(x) = \frac{1}{2}\left(1 + erf\left(\frac{x}{\sqrt{2}}\right)\right)$$

And $\phi(\sigma)$ is the standard PDF of normal distribution

$\phi(Z) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}Z^2)$,where $Z \sim \mathcal{N}(0,1)$

therefore, for one sided truncation of lower tail

$b = \infty,\ \phi(\beta) = 0,\ \Phi(\beta) = 1$ then,

$$E(X|X>a)=\mu+\sigma(\frac{\phi(\alpha)}{1-\Phi(\alpha)})$$

And,

$$V(X|X>a)=\sigma^2[1+\frac{\sigma\phi(\alpha)}{1-\Phi(\alpha)}-(\frac{\phi(\alpha)}{1-\Phi(\alpha)})^2]$$

Conclusions:-

a.

i.) Survival function $S(x|X>A)$ = $S_X(x)=\frac{1}{2}-\frac{1}{2}erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)$

ii.) Hazard rate function $\lambda(x|X>A)$ = $\frac{\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\frac{1}{2}-\frac{1}{2}erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)}$

b.

$$E(X|X>a)=\mu+\sigma(\frac{\phi(\alpha)}{1-\Phi(\alpha)})$$

And,

$$V(X|X>a)=\sigma^2[1+\frac{\sigma\phi(\alpha)}{1-\Phi(\alpha)}-(\frac{\phi(\alpha)}{1-\Phi(\alpha)})^2]$$

Problem: 6

Following are the activity value in micromoles per minute per gram of a tissue of a certain enzyme measured in gastric tissue of 30 patient with gastric carcinoma:

0.360	**1.189**	**0.614**	**0.273**	**2.464**	**0.571**	**1.827**	**0.537**	**0.374**	**0.262**
0.448	0.971	0.372	0.898	0.411	1.925	0.550	0.622	0.610	0.319
0.406	0.767	0.385	0.674	0.521	0.603	0.533	1.177	0.307	1.499

Explain the concept and method of construction of confidence intervals and also find out the confidence interval of mean.

Theory: - A confidence interval (CI) is a type of estimate computed from the statistics of the observed data. This gives a range of values for an unknown parameter (for example, a population mean). The interval has an associated confidence level that gives the probability with which an estimated interval will contain the true value of the parameter.

Formula for the confidence interval is-

$$CI = \bar{x} \pm \mathbb{Z}_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

Where $\bar{x}$ is sample mean, n is sample size, s is sample standard deviation, α is the level of significance (in this case $\alpha = 1 - 95\% = 5\%$) and z is the critical value for $\alpha/2$ level of significance.

Methodology: - First, we will calculate mean and standard deviation of activities by using **mean()** and **var()** function. And then, we will create some functions in r studio and do some algebraic calculations to find required result. We will use **qnorm()** function from the package stats to find the z score for 95% interval.

R-Code:-

```
rm(list=ls())
activities = c(0.360,1.189,0.614,0.273,2.464,0.571,1.827,0.537,0.374,0.262
,0.448, 0.971, 0.372, 0.898, 0.411, 1.925, 0.550, 0.622, 0.610, 0.319, 0.406,
0.767, 0.385, 0.674, 0.521, 0.603, 0.533, 1.177, 0.307, 1.499)
n = length(activities);n
mean_act<-mean(activities);mean_act
sd_act<-sqrt(var(activities));sd_act
# for 95% confidence
alpha <- 0.05
error<-qnorm(1-alpha/2)*(sd_act/sqrt(n))
error
# for Lower limit
Mean_l<- mean_act - error
Mean_l
# for upper limit
Mean_u<- mean_act + error
Mean_u
```

R-Console:-

```
> rm(list=ls())
> activities = c(0.360,1.189,0.614,0.273,2.464,0.571,1.827,0.537,0.374,0.262
,0.448, 0.971, 0.372, 0.898, 0.411, 1.925, 0.550, 0.622, 0.610, 0.319, 0.406,
0.767, 0.385, 0.674, 0.521, 0.603, 0.533, 1.177, 0.307, 1.499)
> n = length(activities);n
[1] 30
> mean_act<-mean(activities);mean_act
[1] 0.7489667
> sd_act<-sqrt(var(activities));sd_act
[1] 0.5422157
> # for 95% confidence
> alpha <- 0.05
> error<-qnorm(1-alpha/2)*(sd_act/sqrt(n))
> error
[1] 0.1940258
> # for Lower limit
> Mean_l<- mean_act - error
> Mean_l
[1] 0.5549408
> # for upper limit
> Mean_u<- mean_act + error
> Mean_u
[1] 0.9429925
```

Conclusion:-

The Confidence interval of mean for given data is [0.5549408,0.9429925]

Problem: 7

Suppose the mean weight of king penguins found in an Antarctic colony last year was 15.4 kg. In a sample of 35 penguins, at the same time this year in the same colony, the mean penguin weight is 14.6 kg. Assuming the population standard deviation is 2.5 kg at 5% significance level, can we reject the null hypothesis that the penguin weight does not differ from last year?

Objective: - To check whether to reject the null hypothesis or not.

Theory: - A hypothesis is statistical about the population. Hypothesis testing can be used to determine whether a statement about the population parameter value should or should not be rejected. Null hypothesis is a tentative assumption about a population parameter, and it is denoted by H_0. The alternative hypothesis is the opposite of what is stated in the null hypothesis denoted by H_1. Hypothesis testing use data from the sample to test the 2 competing statements indicating by H_0 and H_1. Because the hypothesis tests are based on data, we must allow for the possibility of errors the possibility of making a type one error is denoted by alpha (α) and it is the probability of rejecting null hypothesis when null hypothesis is true. The possibility of making type two error denoted by beta, is the possibility of accepting null hypothesis when null hypothesis is false.

There are three approaches to test the hypothesis

- P value approach: - reject H_0 when $p - value \leq \alpha$.
- Critical value approach: - reject H_0 when $z < -z_\alpha \, or \, z > z_\alpha$.
- confidence interval method: - If the hypothesis mean value belong to the confidence interval, do not reject the null hypothesis.

Methodology: -

Null hypothesis of two tailed test :$\mu = \mu_0, \, i.e.,\, \mu = 15.4$

Alternative hypothesis is:$\mu \neq 15.4$

The required test statistic is $Z = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$

We will use critical value approach here.

Reject $H_0 \; if \; Z \leq -Z_{\frac{\alpha}{2}} \; or \; Z \geq Z_{\frac{\alpha}{2}}$;

R-Code:-

```
mean_last_year <- 15.4  # Mean weight last year
mean_this_year <- 14.6  # Mean weight this year
sample_size <- 35       # Sample size
pop_sd <- 2.5           # Population standard deviation
significance_level <- 0.05

# Perform one-sample Z-test
Z_stat <- (mean_this_year - mean_last_year) / (pop_sd / sqrt(sample_size))
Z_stat
p_value <- 2 * pt(abs(Z_stat),sample_size , lower.tail = FALSE)  # Two-tailed p-value

# Check if we can reject the null hypothesis
if (p_value < significance_level) {
  result <- "Reject the null hypothesis"
} else {
  result <- "Fail to reject the null hypothesis"
}
cat("t-statistic:")
cat("Degrees of freedom:")
cat("p-value:", p_value )
cat(result)
```

R-Console:-

```
rm(list=ls())
> mean_last_year <- 15.4  # Mean weight last year
> mean_this_year <- 14.6  # Mean weight this year
> sample_size <- 35       # Sample size
> pop_sd <- 2.5           # Population standard deviation
> significance_level <- 0.05
>
> # Perform one-sample Z-test
> Z_stat <- (mean_this_year - mean_last_year) / (pop_sd / sqrt(sample_size))
> Z_stat
[1] -1.893146
> p_value <- 2 * pt(abs(Z_stat),sample_size , lower.tail = FALSE)  # Two-tailed p-value
>
> # Check if we can reject the null hypothesis
> if (p_value < significance_level) {
+   result <- "Reject the null hypothesis"
+ } else {
+   result <- "Fail to reject the null hypothesis"
+ }
> cat("t-statistic:")
t-statistic:> cat("Degrees of freedom:")
Degrees of freedom:> cat("p-value:", p_value )
p-value: 0.06663033> cat(result)
Fail to reject the null hypothesis>
```

Conclusion:- From above experiment we conclude that we Fail to reject the null hypothesis and hence there is no significance difference bin mean weight of penguin from last year.

Problem: 8

Let $x_1, x_2, \ldots, x_n$ be a random sample from Normal Population $\mathbb{N}(\mu, \sigma^2)$. Obtain C-R inequality for the variance of the unbiased estimator of μ.

Objective: - To obtain C-R inequality for the variance of the unbiased estimator of μ.

Theory:- If t is an unbiased estimator of $\gamma(\theta)$, then under certain regulatory conditions

$$Var(t) \geq \frac{\left\{\frac{d}{d\theta}\gamma(\theta)\right\}^2}{\mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\log L\right)^2\right]} = \frac{\{\gamma'(\theta)\}^2}{I(\theta)}$$

Where $I(\theta)$ is the information on θ, supplied by the sample. And L is the likelihood function

$$L = \prod_{i=1}^{n} f(x_i, \theta_1, \theta_2, \ldots, \theta_k)$$

If t is an unbiased estimator of $\gamma(\theta)$, and $\gamma(\theta)$ = θ this implies $\gamma'(\theta) = 1$.

$$Var(t) \geq \frac{1}{I(\theta)} \tag{1}$$

An estimator T is said to be MVUE of γ(θ) if
$i.$ T is the unbiased estimator for γ(θ).
$ii.$ And T has the smallest variance among the class of all unbiased estimator of $\gamma(\theta)$.

Methodology: First we will find out likelihood function from the pdf of Normal Distribution and then take log of both sides which will give us $logL$ and then partial differentiate with respect to parameter μ and square both sides and then take expectations which will give us-

$$I(\theta) = \mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\log L\right)^2\right]$$

Put the obtained value in equation (1), and we can get the required result.

Solution:

We have, $X \sim \mathbb{N}(\mu, \sigma^2)$

$$f(x,\mu) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \quad ;\sigma > 0, \quad -\infty < x < \infty;$$
$$-\infty < \mu < \infty.$$

So, the likelihood function is given by

$$\begin{aligned} L &= \prod_{i=1}^{n}\frac{1}{\sigma\sqrt{2\pi}}\exp\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \\ &= \left(\frac{1}{2\pi}\right)^{\frac{n}{2}}\left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}}\exp\left(-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}\right) \end{aligned}$$

Now taking log both sides, we get

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

differentiating with respect to μ we get,

$$\begin{aligned}\frac{\partial\log L}{\partial\mu}&=-0-0-\frac{1}{2}.2\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma^2}\right)(-1)\\ \Rightarrow\frac{\partial\log L}{\partial\mu}&=\frac{1}{\sigma^2}\left[\sum_{i=1}^{n}(x_i-\mu)\right]\\ &=\frac{1}{\sigma^2}\left[\sum_{i=1}^{n}(x_i-n\mu)\right]\\ &=\frac{1}{\sigma^2}\left[\sum_{i=1}^{n}(n\bar{x}-n\mu)\right]\\ \Rightarrow\frac{\partial\log L}{\partial\mu}&=\frac{n}{\sigma^2}[\bar{x}-\mu]\end{aligned}$$

So,

$$\begin{aligned}I(\mu)=\mathbb{E}\left(\frac{\partial\log L}{\partial\mu}\right)^2&=\mathbb{E}\left[\frac{n^2}{\sigma^4}(x_i-\mu)^2\right]\\ &=\frac{n^2}{\sigma^4}\mathbb{E}(\bar{x}-\mu)^2\\ &=\frac{n^2}{\sigma^4}.Var(\bar{x})\\ &=\frac{n^2}{\sigma^4}.\frac{\sigma^2}{n}\ \Big|\bar{X}\sim\mathcal{N}(\mu,\sigma^2)\\ I(\mu)&=\frac{n}{\sigma^2}\end{aligned}$$

So by C-R inequality,

$$\begin{gathered}V(T)\geq\frac{1}{I(\mu)}=\frac{\sigma^2}{n}=V(\bar{x})\\ V(T)\geq V(\bar{x})\\ \therefore V(\bar{x})\leq V(T)\quad\forall\,T\end{gathered}$$

$\bar{x}$ is Minimum Variance Unbiased Estimator for μ.

Conclusion: $\bar{x}$ is minimum variance unbiased estimator of μ

Problem: 9

Let $x_1, x_2, \ldots, x_n$ be a random sample from p.m.f $f(x, \theta) = \theta^x (1-\theta)^{1-x}$; $x = 0, 1.$ Obtain C-R lower bound for the variance of unbiased estimator of θ.

Objective: - To obtain C-R lower bound for the variance of the unbiased estimator of θ.

Theory:- If t is an unbiased estimator of $\gamma(\theta)$, then under certain regulatory conditions

$$Var(t) \geq \frac{\left\{\frac{d}{d\theta}\gamma(\theta)\right\}^2}{\mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\log L\right)^2\right]} = \frac{\{\gamma'(\theta)\}^2}{I(\theta)}$$

Where $I(\theta)$ is the information on θ, supplied by the sample. And L is the likelihood function

$$L = \prod_{i=1}^{n} f(x_i, \theta_1, \theta_2, \ldots, \theta_k)$$

If t is an unbiased estimator of $\gamma(\theta)$, and $\gamma(\theta)$ = θ this implies $\gamma'(\theta) = 1.$

$$Var(t) \geq \frac{1}{I(\theta)} \tag{1}$$

An estimator T is said to be MVUE of γ(θ) if
$i.$ T is the unbiased estimator for γ(θ).
$ii.$ And T has the smallest variance among the class of all unbiased estimator of $\gamma(\theta).$

Methodology:- First we will find out likelihood function from the pdf of the given distribution and then take log of both sides which will give us $\log L$ and then partial differentiate with respect to parameter θ and square both sides and then take expectations which will give us-

$$I(\theta) = \mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\log L\right)^2\right]$$

Put the obtained value in equation (1), and we can get the required result.

Solution:- We know that $f(x, \theta) = \theta^x (1-\theta)^{1-x}; x = 0, 1$ likelihood function is given by-

$$\begin{aligned} L &= \prod_{i=1}^{n} \theta^{x_i} (1-\theta)^{1-x_i} \\ &= \theta^{\sum_{i=1}^{n} x_i} . (1-\theta)^{\sum_{i=1}^{n}(1-x_i)} \\ &= \theta^{n\bar{x}} . (1-\theta)^{n-n\bar{x}} \\ &= \theta^{n\bar{x}} . (1-\theta)^{n(1-\bar{x})} \end{aligned}$$

differentiating it with respect to θ, we get-

$$
\begin{aligned}
\frac{\partial \log L}{\partial \theta} &= \frac{\partial}{\partial \theta}\left[n\bar{x}\log\theta + n(1-\bar{x})\log(1-\theta)\right] \\
&= \frac{n\bar{x}}{\theta} + \frac{n(1-\bar{x})}{1-\theta}.(-1) \\
&= \frac{n\bar{x}}{\theta} - \frac{n(1-\bar{x})}{1-\theta} \\
&= \frac{n}{\theta(1-\theta)}\left[(1-\theta)\bar{x} - (1-\bar{x}).\theta\right] \\
&= \frac{n}{\theta(1-\theta)}[\bar{x}-\theta]
\end{aligned}
$$

So,

$$
\begin{aligned}
I(\theta) &= \mathbb{E}\left(\frac{\partial \log L}{\partial \theta}\right)^2 \\
&= \mathbb{E}\left[\frac{n^2}{\theta^2(1-\theta)^2}(\bar{x}-\theta)^2\right] \\
&= \frac{n^2}{\theta^2(1-\theta)^2}.\mathbb{E}(\bar{x}-\theta)^2 \\
&= \frac{n^2}{\theta^2(1-\theta)^2}.V(\bar{x}) \\
&= \frac{n^2}{\theta^2(1-\theta)^2}.\frac{\theta(1-\theta)}{n} \\
&= \frac{n}{\theta(1-\theta)}
\end{aligned}
$$

$\therefore$ C-R lower bound for the variance of the unbiased estimator of θ is -

$$\frac{1}{I(\theta)} = \frac{\theta(1-\theta)}{n} = V(\bar{x})$$

So,

$$V(T) \geq \tfrac{1}{I(\theta)} = V(\bar{x})$$

$$\Rightarrow V(\bar{x}) \leq V(T) \ \ \forall\, T$$

$\bar{x}$ is Minimum variance unbiased estimator for θ.

Conclusion:- The C-R lower bound for the variance of the unbiased estimator of θ is $V(\bar{x}) = \dfrac{\theta(1-\theta)}{n}$.

Practical: 10

Let $x_1, x_2, \ldots, x_n$ be a random sample from Poisson distribution with parameter λ. Find minimum variance unbiased estimate (MVUE) of λ.

Objective:- To find minimum variance unbiased estimate (MVUE) of λ.

Theory:- If t is an unbiased estimator of $\gamma(\lambda)$, then under certain regulatory conditions

$$Var(t) \geq \frac{\left\{\frac{d}{d\lambda}\gamma(\lambda)\right\}^2}{\mathbb{E}\left[\left(\frac{\partial}{\partial\lambda}\log L\right)^2\right]} = \frac{\{\gamma'(\lambda)\}^2}{I(\lambda)} \tag{1}$$

$I(\lambda)$ is the information on λ supplied by the sample. And it can be written as

$$I(\lambda) = \mathbb{E}\left[\left(\frac{\partial}{\partial\lambda}\log L\right)^2\right] = -n \times \mathbb{E}\left(\frac{\partial^2}{\partial\lambda^2}\log f(x,\lambda)\right)$$

An estimator t is said to be MVUE of $\gamma(\lambda)$ if

1. t is unbiased estimator of $\gamma(\lambda)$.
2. And t has the smallest variance among the class of all unbiased estimator of $\gamma(\lambda)$.

Methodology:- In this problem, we take log of pdf of poisson distribution and then double differentiate it with respect to parameter λ and take expectation of both sides, which gives us $\mathbb{E}\left(\frac{\partial^2}{\partial\theta^2}\log f(x,\lambda)\right)$ and we can use $I(\lambda) = -n \times \mathbb{E}\left(\frac{\partial^2}{\partial\theta^2}\log f(x,\lambda)\right)$ and put the obtained value in equation (1) to get the required result.

Solution:- We have $X \sim P(\lambda)$,

So,

$$f(x,\lambda) = \frac{e^{-\lambda}\lambda^x}{x!} \quad ; x = 0, 1, 2, \ldots, \; \lambda > 0$$

$$\therefore \log f(x,\lambda) = -\lambda + x\log\lambda - \log x!$$

differentiating it with respect to λ, we get -

$$\begin{aligned}
\frac{\partial}{\partial\lambda}\log f(x,\lambda) &= -1 + \frac{x}{\lambda} - 0 \\
\frac{\partial}{\partial\lambda}\log f(x,\lambda) &= -1 + \frac{x}{\lambda} \\
\frac{\partial^2}{\partial\lambda^2}\log f(x,\lambda) &= \frac{-x}{\lambda^2}
\end{aligned}$$

$$\begin{aligned}
I(\lambda) &= -n.\,\mathbb{E}\left(\frac{\partial^2}{\partial\lambda^2}\log f\right) \\
&= -n.\,\mathbb{E}\left(\frac{-x}{\lambda^2}\right) \\
&= \frac{n}{\lambda^2}\mathbb{E}(x) \\
&= \frac{n}{\lambda^2}.\lambda
\end{aligned}$$

$$\therefore \quad I(\lambda) = \frac{n}{\lambda}$$

By C-R inequality,

$$V(T) \geq \frac{1}{I(\lambda)} = \frac{\lambda}{n} = V(\bar{x})$$

So, $V(\bar{x}) \leq V(T) \;\; \forall\, T$,

So, $\bar{x}$ is Minimum variance unbiased estimator for λ.

Conclusion:- $\bar{x}$ is Minimum variance unbiased estimator for λ.

Problem: 11

Let $x_1, x_2, \ldots, x_n$ be a random sample from Normal Population $N(\mu, \sigma^2)$. Show that $x - bar$ is MVBE of μ and obtain its variance.

Theory:- If t is an unbiased estimator of $\gamma(\theta)$, then under certain regulatory conditions

$$Var(t) = \frac{\left\{\frac{d}{d\theta}\gamma(\theta)\right\}^2}{\mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\log L\right)^2\right]} = \frac{\{\gamma'(\theta)\}^2}{I(\theta)}$$

Where $I(\theta)$ is the information on θ, supplied by the sample. And L is the likelihood function

$$L = \prod_{i=1}^{n} f(x_i, \theta_1, \theta_2, \ldots, \theta_k)$$

If t is an unbiased estimator of $\gamma(\theta)$, and $\gamma(\theta)$ = θ this implies $\gamma'(\theta) = 1$.

$$Var(t) = \frac{1}{I(\theta)} \tag{1}$$

An estimator T is said to be MVBE of γ(θ) if
$i.$ T is the estimator for γ(θ).
$ii.$ And T has the smallest variance among the class of all estimator of $\gamma(\theta)$.

Methodology: First we will find out likelihood function from the pdf of Normal Distribution and then take log of both sides which will give us $logL$ and then partial differentiate with respect to parameter μ and square both sides and then take expectations which will give us-

$$I(\theta) = \mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\log L\right)^2\right]$$

Put the obtained value in equation (1), and we can get the required result.

Solution:-

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

$$f(x, \mu) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]; \quad -\infty < x < \infty; \\ -\infty < \mu < \infty; \\ \sigma > 0.$$

So, the likelihood function is give by,

$$\begin{aligned} L &= \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}}\exp\left[\frac{-1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2\right] \\ &= \left(\frac{1}{2\pi}\right)^{\frac{n}{2}}\left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}}.\exp\left(\frac{-1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}\right) \end{aligned}$$

Now taking log both sides, we get -

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

differentiating with respect to μ we get,

$$\begin{aligned}\frac{\partial}{\partial\mu}\log L &= 0 - 0 - \frac{1}{2}.2\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}(-1)\\ &= \frac{1}{\sigma^2}\left[\sum_{i=1}^{n}(x_i-\mu)\right]\\ &= \frac{1}{\sigma^2}\left[\sum_{i=1}^{n}x_i - n\mu\right]\\ &= \frac{1}{\sigma^2}[n\bar{x}-n\mu]\\ \therefore\ \frac{\partial}{\partial\mu}\log L &= \frac{n}{\sigma^2}[\bar{x}-\mu]\end{aligned}$$

So,

$$\begin{aligned}I(\mu) = \mathbb{E}\left(\frac{\partial logL}{\partial\mu}\right)^2 &= \mathbb{E}\left[\frac{n^2}{\sigma^4}(\bar{x}-\mu)^2\right]\\ &= \frac{n^2}{\sigma^4}\mathbb{E}(\bar{x}-\mu)^2\\ &= \frac{n^2}{\sigma^4}.Var(\bar{x})\\ &= \frac{n^2}{\sigma^4}.\frac{\sigma^2}{n}\\ I(\mu) &= \frac{n}{\sigma^2}\end{aligned}$$

So, by C-R inequality-

$$\begin{aligned}V(T) \geq \frac{1}{I(\mu)} = \frac{\sigma^2}{n} = V(\bar{x})\\ V(T) \geq V(\bar{x})\\ \therefore V(\bar{x}) \leq V(T) \quad \forall\, T\end{aligned}$$

$\bar{x}$ is minimum variance bound estimator of μ.

Problem 12

Solve Graphically:

$MaxZ = 80x_1 + 55x_2$

Subject to:-

$$4x_1 + 2x_2 \leq 40$$
$$2x_1 + 4x_2 \leq 32$$
$$x_1 \geq 0, x_2 \geq 0$$

Objective:- To Maximize given objective function which are subject to given constraints by using Graphical method.

Theory:- Linear programming problems which involve only two variables can be solved by graphical method. If the problem has three or more variables, the graphical method is impractical. We just graphically find that for which point the objective function satisfy.

Solution:-

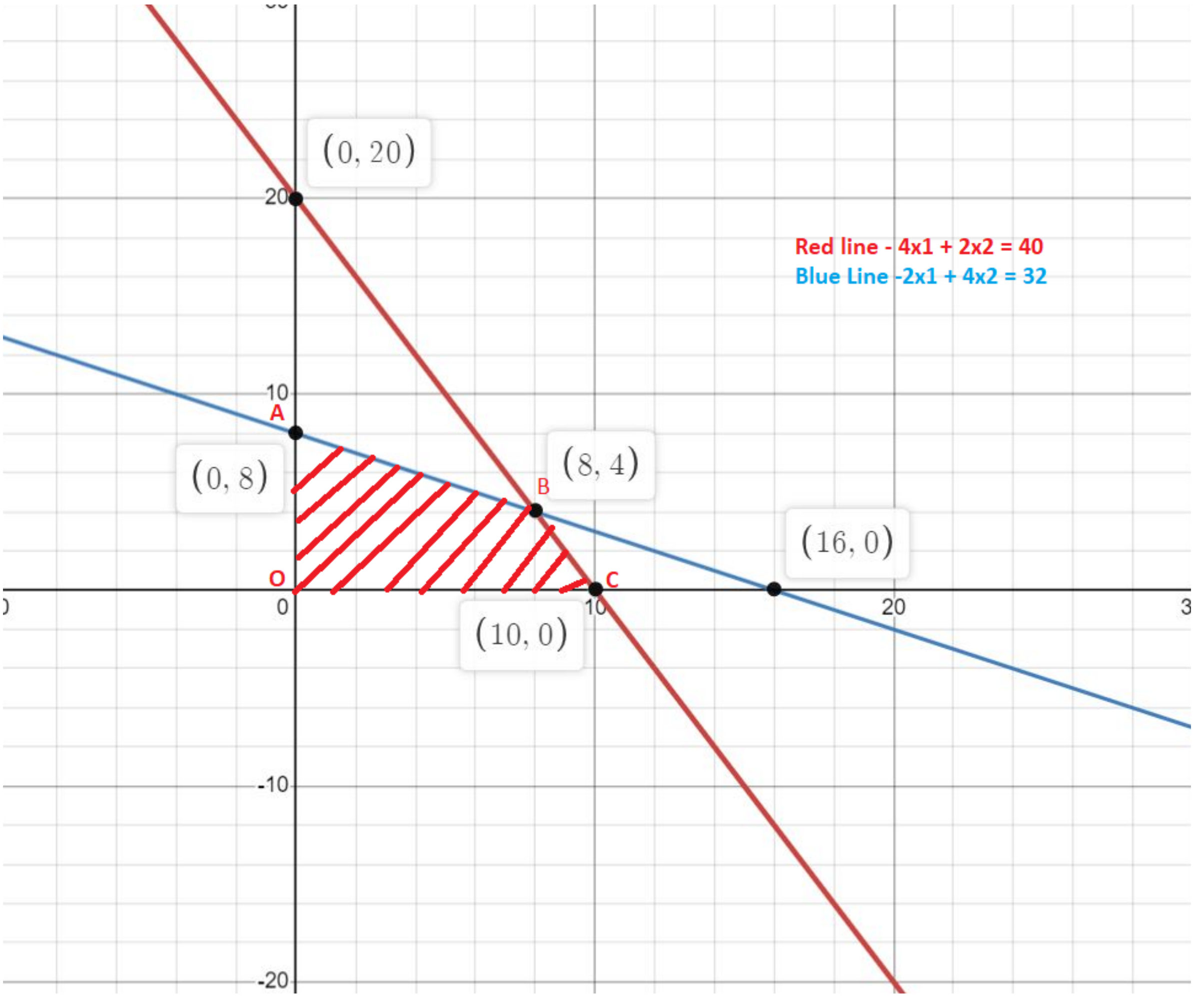
For equation $4x_1 + 2x_2 = 40$

x_1	0	8	10
x_2	20	4	0

For equation $2x_1 + 4x_2 = 32$

x_1	0	8	16
x_2	8	4	0

Graph:-



Conclusion:- Here area of (O,A,B,C) shows unbounded feasible Region which satisfy all the constraints and Objective function is satisfied at Point B = **(8,4)** for which maximum of Z will be **860**.

Problem 13

Solve by using Simplex Method:

$MaxZ = 80x_1 + 55x_2$

Subject to:-

$$4x_1 + 2x_2 \leq 40$$
$$2x_1 + 4x_2 \leq 32$$
$$x_1 \geq 0, x_2 \geq 0$$

Objective:- To Maximize given objective function which are subject to given constraints by using Simplex method.

Theory:- The Simplex method is an approach to solving linear programming models by using slack variables, surplus variables, tableaus, and pivot variables as a means to finding the optimal solution of an optimization problem(here maximization).

Solution:-

$MaxZ = 80x_1 + 55x_2$

Subject to:-

$$4x_1 + 2x_2 \leq 40$$
$$2x_1 + 4x_2 \leq 32$$
$$x_1 \geq 0, x_2 \geq 0$$

adding slack variables in the constraints of the given Linear Programming Problem. The Standard form of the LP problem becomes:-

$$MaxZ = 80x_1 + 55x_2 + 0S_1 + 0S_2$$

		C_j	80	55	0	0		
C_B	B	X_B	X_1	X_2	S_1	S_2	Min Ratio	
0	S_1	40	4	2	1	0	10	$\rightarrow$
0	S_2	32	2	4	0	1	16	
		Z_j	0	0	0	0		
		$Z_j - C_j$	-80	-55	0	0		
			$\uparrow$					

		C_j	80	55	0	0		
C_B	B	X_B	X_1	X_2	S_1	S_2	Min Ratio	
80	X_1	10	1	$1/2$	$1/4$	0	20	
0	S_2	12	0	3	$-1/2$	1	4	$\rightarrow$
		Z_j	80	40	20	0		
		$Z_j - C_j$	0	-15	20	0		
				$\uparrow$				

		C_j	80	55	0	0
C_B	B	X_B	X_1	X_2	S_1	S_2
80	X_1	8	1	0	$1/3$	$-1/6$
55	X_2	4	0	1	$-1/6$	$1/3$
		Z_j	80	55	$105/6$	5
		$Z_j - C_j$	0	0	$105/6$	5

Conclusion:- Here we see that all $Z_j - C_j \geq 0$

$X_1 = 8, X_2 = 4$

$MaxZ = 80 \times 8 + 55 \times 4$

$= 640 + 220 =$ **860** .

Problem 14

Solve by using Two-phase Simplex Method:

$MaxZ = 3x_1 - x_2$

Subject to:-

$$
\begin{gathered}
2x_1 + x_2 \geq 2 \\
x_1 + 3x_2 \geq 2 \\
x_2 \geq 4 \\
x_1 \geq 0, x_2 \geq 0
\end{gathered}
$$

Objective:- To Maximize given objective function which are subject to given constraints by using Two-phase Simplex method.

Theory:- In the Two-phase simplex method, we add artificial variables to the constraints and find a Basic feasible solution to the original Linear Programming by solving the Phase I Linear Programming. In the Phase I Linear Programming, the objective function is to minimize the sum of all artificial variables. At the completion of Phase I, we reintroduce the original LP's objective function and determine the optimal solution to the original Linear Programming.

Solution:-

$MaxZ = 3x_1 - x_2$

Subject to:-

$$
\begin{gathered}
2x_1 + x_2 \geq 2 \\
x_1 + 3x_2 \geq 2 \\
x_2 \geq 4 \\
x_1 \geq 0, x_2 \geq 0
\end{gathered}
$$

adding slack variable, surplus variable and artificial variable in the constraints of the given LP problem. The Standard form of the LP problem becomes:

$$MaxZ = 3x_1 + x_2 + 0S_1 + 0S_2 + 0S_3 + a_1$$

$$2x_1 + x_2 - S_1 + a_1 = 2$$

$$x_1 + 3x_2 + S_2 = 2$$

$$x_2 + S_3 = 4$$

Thus, $MaxZ = -a_1$

Phase 1

		C_j	0	0	0	0	0	-1	
C_B	B	X_B	X_1	X_2	S_1	S_2	S_3	a_1	min Ratio
-1	a_1	2	2	1	-1	0	0	1	$1 \rightarrow$
0	S_2	2	1	3	0	1	0	0	2
0	S_3	4	0	1	0	0	1	0	$-$
		Z_j	-2	-1	-1	0	0	-1	
		$Z_j - C_j$	-2	-1	1	0	0	0	
			$\uparrow$						

$R_1 \rightarrow \frac{r_1}{2}$

$R_2 \rightarrow r_2 - R_1$

		C_j	0	0	0	0	0
C_B	B	X_B	X_1	X_2	S_1	S_2	S_3
0	X_1	1	1	$1/2$	$-1/2$	0	0
0	S_2	1	0	$5/2$	$1/2$	1	0
0	S_3	4	0	1	0	0	1
		Z_j	0	0	0	0	0
		$Z_j - C_j$	0	0	0	0	0

Phase II

		C_j	3	-1	0	0	0	
C_B	B	X_B	X_1	X_2	S_1	S_2	S_3	Min Ratio
3	X_1	1	1	$1/2$	$-1/2$	0	0	$-$
0	S_2	1	0	$5/2$	$1/2$	1	0	2 $\rightarrow$
0	S_3	4	0	1	0	0	1	4
		Z_j	3	$3/2$	$-3/2$	0	0	
		$Z_j - C_j$	0	$5/2$	$-3/2$	0	0	
					$\uparrow$			

$R_2' \rightarrow 2R_2$

$R_1' \rightarrow r_1 + \frac{R_2'}{2}$

		C_j	3	-1	0	0	0
C_B	B	X_B	X_1	X_2	S_1	S_2	S_3
3	X_1	2	1	3	0	1	0
0	S_1	2	0	5	1	2	0
0	S_3	4	0	1	0	0	1
		Z_j	3	9	0	3	0
		$Z_j - C_j$	0	10	0	3	0

Conclusion:-

Here, In Big-M Method the objective function will satisfy for $x_1 = 2$ and $x_2 = 0$.

And Thus, The optimum solution for the given problem will be $3 \times 2 - 0 = 6$ (as objective function is $3x_1$-x_2).

Problem 15

Solve by using Big-M Method:

$MaxZ = -2x_1 - x_2$

Subject to:-

$$\begin{gathered} 3x_1 + x_2 = 3 \\ 4x_1 + 3x_2 \geq 6 \\ x_1 + 2x_2 \leq 4 \\ and, x_1 \geq 0, x_2 \geq 0 \end{gathered}$$

Objective:- To Maximize given objective function which are subject to given constraints by using Big-M method.

Theory:- The Big M method is a version of the Simplex Algorithm that first finds a Basic Feasible Solution by adding "artificial" variables to the problem. The objective function of the original Linear Programming Problem must be modified to ensure that the artificial variables are all equal to 0 at the conclusion of the Simplex algorithm, the only difference is that here we use "M" in I phase only.

Solution:-

$MaxZ = -2x_1 - x_2$

Subject to:-

$$\begin{gathered} 3x_1 + x_2 = 3 \\ 4x_1 + 3x_2 \geq 6 \\ x_1 + 2x_2 \leq 4 \\ and, x_1 \geq 0, x_2 \geq 0 \end{gathered}$$

adding slack variable S_3; surplus variable S_2 and artificial variable a_1 and a_2 in the constraints of the given LP problem, the standard form of the LP problem becomes:-

$$MaxZ = -2x_1 - x_2 + 0S_2 + 0S_3 - Ma_1 - Ma_2$$

$$\begin{gathered} 3x_1 + x_2 + a_1 = 3 \\ 4x_1 + 3x_2 - S_2 + a_2 = 6 \\ x_1 + 2x_2 + S_3 = 4 \end{gathered}$$

		C_j	-2	-1	$-M$	0	$-M$	0		
C_B	B	X_B	X_1	X_2	a_1	S_2	a_2	S_3	min Ratio	
$-M$	a_1	3	3	1	1	0	0	0	1	$\rightarrow$
$-M$	a_2	6	4	3	0	-1	1	0	1.5	
0	S_3	4	1	2	0	0	0	1	4	
		Z_j	$-7M$	$-4M$	$-M$	M	$-M$	0		
		$Z_j - C_j$	$-7M-2$	$-4M-1$	0	M	0	0		
			$\uparrow$							

$MaxZ = -M$

$R_1 \rightarrow \frac{r_1}{3}$

$R_2 \rightarrow r_2 - 4R_1$

$R_3 \rightarrow r_3 - R_1$

		C_j	-2	-1	0	$-M$	0		
C_B	B	X_B	X_1	X_2	S_2	a_2	S_3	min Ratio	
-2	X_1	1	1	$1/3$	0	0	1	3	$\rightarrow$
$-M$	a_2	2	0	$5/3$	-1	1	0	$6/5$	
0	S_3	3	0	$5/3$	0	0	1	$9/5$	
		Z_j	-2	$\frac{-5M-2}{3}$	M	$-M$	0		
		$Z_j - C_j$	0	$\frac{-5M+1}{3}$	M	0	0		
				$\uparrow$					

$MaxZ = -2 - 2M$

$R_2' \rightarrow \frac{3R_2}{5}$

$R_1' \rightarrow R_1 - \frac{R_2'}{3}$

$R_3' \rightarrow R_3 - \frac{5R_2'}{3}$

		C_j	-2	-1	0	0
C_B	B	X_B	X_1	X_2	S_2	S_3
-2	X_1	$3/5$	1	0	$1/5$	0
-1	X_2	$6/5$	0	1	$-3/5$	0
0	S_3	1	0	1	1	1
		Z_j	-2	-1	$1/5$	0
		$Z_j - C_j$	0	0	$1/5$	0

all $Z_j - C_j \geq 0$

Conclusion:-

Here, $MaxZ = -2 \times \frac{3}{5} + (-1 \times \frac{6}{5}) + 0 = \frac{-12}{5}$

Therefore, the objective function is Maximum at $X_1 = \frac{3}{5}$ and $X_2 = \frac{6}{5}$ with optimum value of $\frac{-12}{5}$ **(-2.4)**.

Problem 16

Consider a clinical trial in which 10 lung cancer patients are followed till death. The survival data are given below. Plot the Survival function graph $t = 4, 5, 6, 8, 8, 8, 10, 10, 11, 12$

Objective:- We must plot Survival function with respect to time for given failure time.

Theory:- The survival curve can be created assuming various situations. It involves computing of probabilities of occurrence of event at a certain point of time and multiplying these successive probabilities by any earlier computed probabilities to get the final estimate.

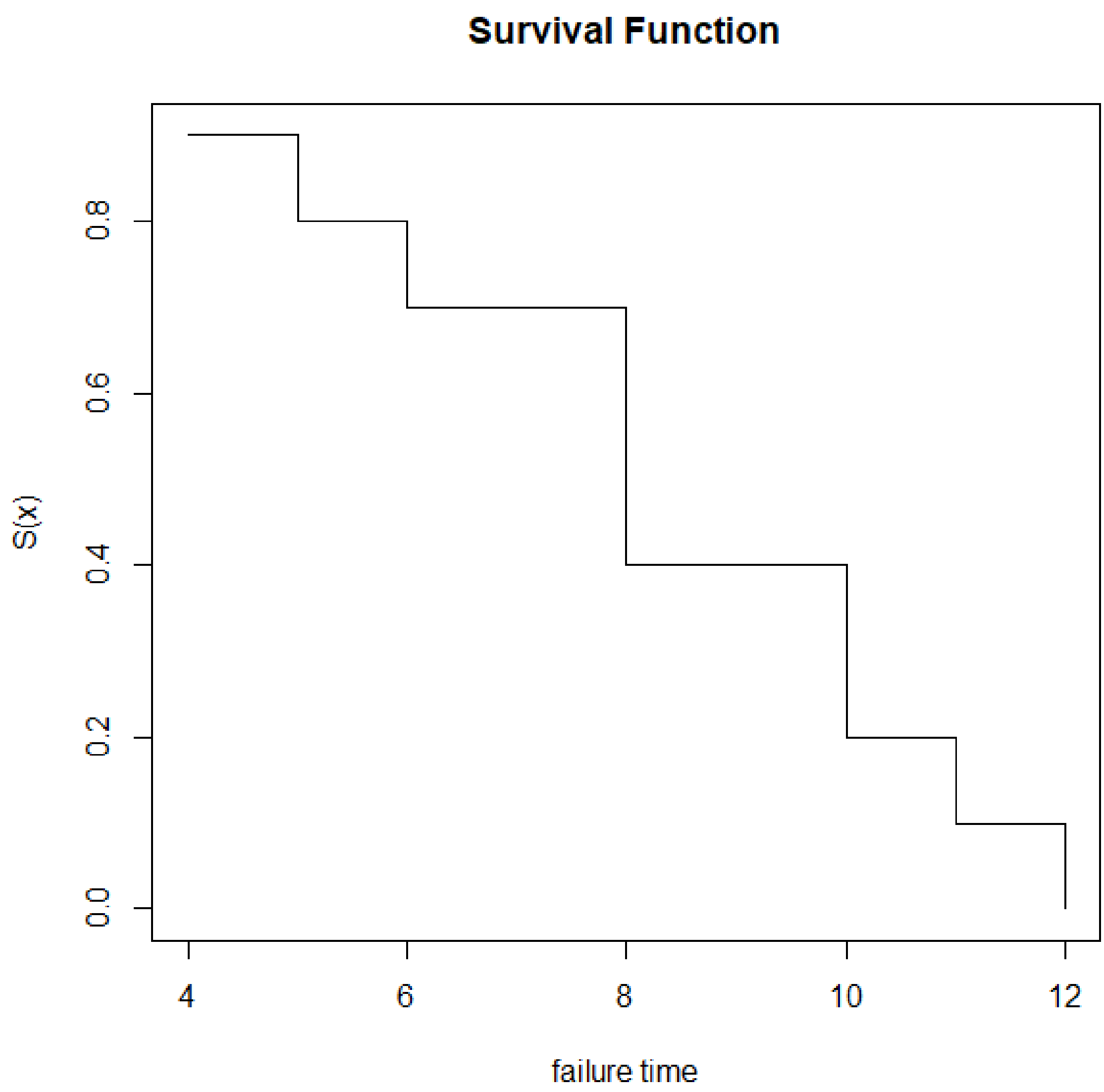
R-Code:-

```
ti = c(4,5,6,8,8,8,10,10,11,12)
b=(rank(-ti,ties.method = 'min')-1)/10
c = data.frame(ti,b);c
plot(ti,b,type='s', xlab = "failure time", ylab = "S(x)", main = "Survival
Function")
```

R-Console:-

```
> ti = c(4,5,6,8,8,8,10,10,11,12)
> b=(rank(-ti,ties.method = 'min')-1)/10
> c = data.frame(ti,b);c
   ti   b
1   4 0.9
2   5 0.8
3   6 0.7
4   8 0.4
5   8 0.4
6   8 0.4
7  10 0.2
8  10 0.2
9  11 0.1
10 12 0.0
> plot(ti,b,type='s', xlab = "failure time", ylab = "S(x)", main = "Survival
Function")
```

Graph:-



Practical: 17

Let $x_1, x_2, \ldots, x_n$ be a random sample from Poisson distribution with parameter λ. Find minimum variance Bound estimate (MVBE) of λ.

Objective:- To find minimum variance Bound estimate (MVBE) of λ.

Theory:- If t is an unbiased estimator of $\gamma(\lambda)$, then under certain regulatory conditions

$$Var(t) \geq \frac{\left\{\frac{d}{d\lambda}\gamma(\lambda)\right\}^2}{\mathbb{E}\left[\left(\frac{\partial}{\partial\lambda}\log L\right)^2\right]} = \frac{\{\gamma'(\lambda)\}^2}{I(\lambda)} \tag{1}$$

$I(\lambda)$ is the information on λ supplied by the sample. And it can be written as

$$I(\lambda) = \mathbb{E}\left[\left(\frac{\partial}{\partial\lambda}\log L\right)^2\right] = -n \times \mathbb{E}\left(\frac{\partial^2}{\partial\lambda^2}\log f(x,\lambda)\right)$$

An estimator t is said to be MVBE of $\gamma(\lambda)$ if

1. t is unbiased estimator of $\gamma(\lambda)$.
2. And t has the smallest variance among the class of all unbiased estimator of $\gamma(\lambda)$.

Methodology:- In this problem, we take log of pdf of poisson distribution and then double differentiate it with respect to parameter λ and take expectation of both sides, which gives us $\mathbb{E}\left(\frac{\partial^2}{\partial\theta^2}\log f(x,\lambda)\right)$ and we can use $I(\lambda) = -n \times \mathbb{E}\left(\frac{\partial^2}{\partial\theta^2}\log f(x,\lambda)\right)$ and put the obtained value in equation (1) to get the required result.

Solution:- We have $X \sim P(\lambda)$,

So,

$$f(x,\lambda) = \frac{e^{-\lambda}\lambda^x}{x!} \quad ; x = 0, 1, 2, \ldots, \ \lambda > 0$$

$$\therefore \log f(x,\lambda) = -\lambda + x\log\lambda - \log x!$$

differentiating it with respect to λ, we get -

$$\begin{aligned}
\frac{\partial}{\partial\lambda}\log f(x,\lambda) &= -1 + \frac{x}{\lambda} - 0 \\
\frac{\partial}{\partial\lambda}\log f(x,\lambda) &= -1 + \frac{x}{\lambda} \\
\frac{\partial^2}{\partial\lambda^2}\log f(x,\lambda) &= \frac{-x}{\lambda^2} \\
I(\lambda) &= -n.\,\mathbb{E}\left(\frac{\partial^2}{\partial\lambda^2}\log f\right) \\
&= -n.\,\mathbb{E}(\frac{-x}{\lambda^2}) \\
&= \frac{n}{\lambda^2}\mathbb{E}(x) \\
&= \frac{n}{\lambda^2}.\lambda \\
\therefore \quad I(\lambda) &= \frac{n}{\lambda}
\end{aligned}$$

$$V(T) = \frac{\lambda}{n} = V(\bar{x})$$

So, $V(\bar{x}) = V(T) \ \forall T$,

So, $\bar{x}$ is Minimum variance bound estimator for λ.

Conclusion:- $\bar{x}$ is Minimum variance bound estimator for λ.

Problem 18:

Let X follows $B(1,\theta), x=0,1.$ Examine the completeness of the distribution.

Objective:- We are to find the complete sufficient statistics of the given distribution.

Theory:- A sufficient statistics t is said to be complete for any measurable function $\phi(t)$

1. $E\{\phi(t)\}=0; \forall\theta$
2. And $\phi(t)=0$; almost everywhere

Methodology/Solution:-

As we know that $T=\sum_{i=1}^{n} x_i$ is a sufficient statistic. So the distribution of T is-

$$\binom{n}{t}.\theta^t(1-\theta)^{n-t}$$

$$\mathbb{E}[\phi(T)]=0$$

$$\Rightarrow \sum_{t=0}^{n}\phi(t).\binom{n}{t}\theta^t(1-\theta)^{n-t}=0$$

$$\Rightarrow \sum_{t=0}^{n}\phi(t).\binom{n}{t}\theta^t\frac{(1-\theta)^n}{(1-\theta)^t}=0$$

$$\Rightarrow (1-\theta)^n\sum_{t=0}^{n}\psi(t)(\frac{\theta}{1-\theta})^2=0\ [where\ \psi(t)=\phi(t).\binom{n}{t}]$$

$$\Rightarrow \sum_{t=0}^{n}\psi(t)(\frac{\theta}{1-\theta})^t=0$$

$$\Rightarrow \sum_{t=0}^{n}\psi(t)\lambda^t=0 \qquad [where\ \lambda=\frac{\theta}{1-\theta}]$$

$$\psi(o)+\psi(1)\lambda+\psi(2)\lambda^2+.........+\psi(n)\lambda^n=0$$

Since, it is polynomial in λ of degree n and equal to zero, so its all coefficients will be zero.

$$\Rightarrow \psi(t)=0$$

$$\Rightarrow \phi(t)\binom{n}{t}=0$$

$$\Rightarrow \phi(t)=0 \qquad [\because \binom{n}{t}\neq 0]$$

almost everywhere 0.

$\Rightarrow \phi(t)$ is complete.

Conclusion:- the measurable function $\emptyset(t)$ of t is complete because

1. $E\{\phi(t)\}=0; \forall\theta$
2. And $\phi(t)=0$; almost everywhere

Problem 19:

The data given below represent survival times (in years) after diagnosis of dementia (a mental disorder) in a group of 97 Swiss females with age at diagnosis $70 - 74$ inclusive. This is part of a larger set of data collected at the University Psychiatric Clinic of Geneva over the period January $1, 1960$ to December $31, 1961$. Find the empirical survival function for the data given below with age at diagnosis 74.

Age at diagnosis (x)	No. of cases	Length of survival time in years (t)
70	15	0.50,0.83,1.17,1.25,1.66,1.67,1.83,3.17,4.17,4.92,5.58,6.58,6.92,8.16,1.50
71	18	0.50,0.83,1.25,1.33,1.67,1.83,3.42,3.42,3.92,4.92,5.75,6.92,7.83,7.84,9.00,9.17,13.83,18.08
72	22	0.50,1.00,1.08,1.42,1.58,1.59,2.00,2.33,2.67,3.58,4.08,4.17,5.75,6.25,7.00,7.08,9.92,10.17,11.25,12.67, 21.00,21.83
73	17	0.58,1.08,1.25,1.67,2.00,2.08,2.17,3.75,4.25,4.58,4.92,5.25,6.75,7.83,7.84,9.17,11.50
74	25	0.50,1.00,1.25,1.41,1.42,1.58,1.66,1.67,2.25,2.33,2.92,3.08,3.75,3.92,4.17,4.67,5.0,5.25,5.83,7.25,8.50, 9.33,10.33,11.25,12.50

Objective:- We must find the empirical survival function for the data given below with age at diagnosis 74.

Theory:- Suppose that a random variable X denotes the lifetime of a living Organism or an inanimate device, it is also called as at death or at failure or briefly age.

1. The cumulative distribution function (CDF); $F_x(x) = P_t(X \leq x)$; is called the lifetime distribution or failure distribution.
2. The complementary function of cumulative distribution function (CDF) is known as survival distribution function (SDF). And SDF = 1 – CDF.

$$SDF = 1 - P_t(X \leq x) = P_t(X > x)$$

3. Let $t_1 < t_2 < \ldots < t_N$; represents the N(Distinct) ordered times at death. Let $F_T(t) = P_r(T \leq t)$; be the theoretical (unknown) failure distribution (CDF), and

$$S_T(t) = 1 - P_r(T \leq t) = P_t(T > t)$$

be the corresponding survival distribution function (SOF)

The Empirical Distribution Function is

$$F_N^o(t) = \begin{cases} 0, & \text{for } t < t_1^{'} \\ \dfrac{i}{N}, & \text{for } t_i^{'} \leq t < t_{i+1}^{'} \\ 1, & \text{for } t \geq t_N^{'} \end{cases}$$

Where i is the rank of the i^{th} (ordered) observation, and $F_N^o(t)$ gives the probability $P_r(T \leq t)$.

And the Empirical Survival Function is given as $S_N^o(t) = 1 - F_N^o(t)$.

$$S_N^o(t) = \begin{cases} 1, & \text{for } t < t_1^{'} \\ \dfrac{N-i}{N}, & \text{for } t_i^{'} \le t < t_{i+1}^{'} \\ 0, & \text{for } t \ge t_N^{'} \end{cases}$$

$S_N^o(t)$ gives the probability $P_r(T > t)$.

R-Code:-

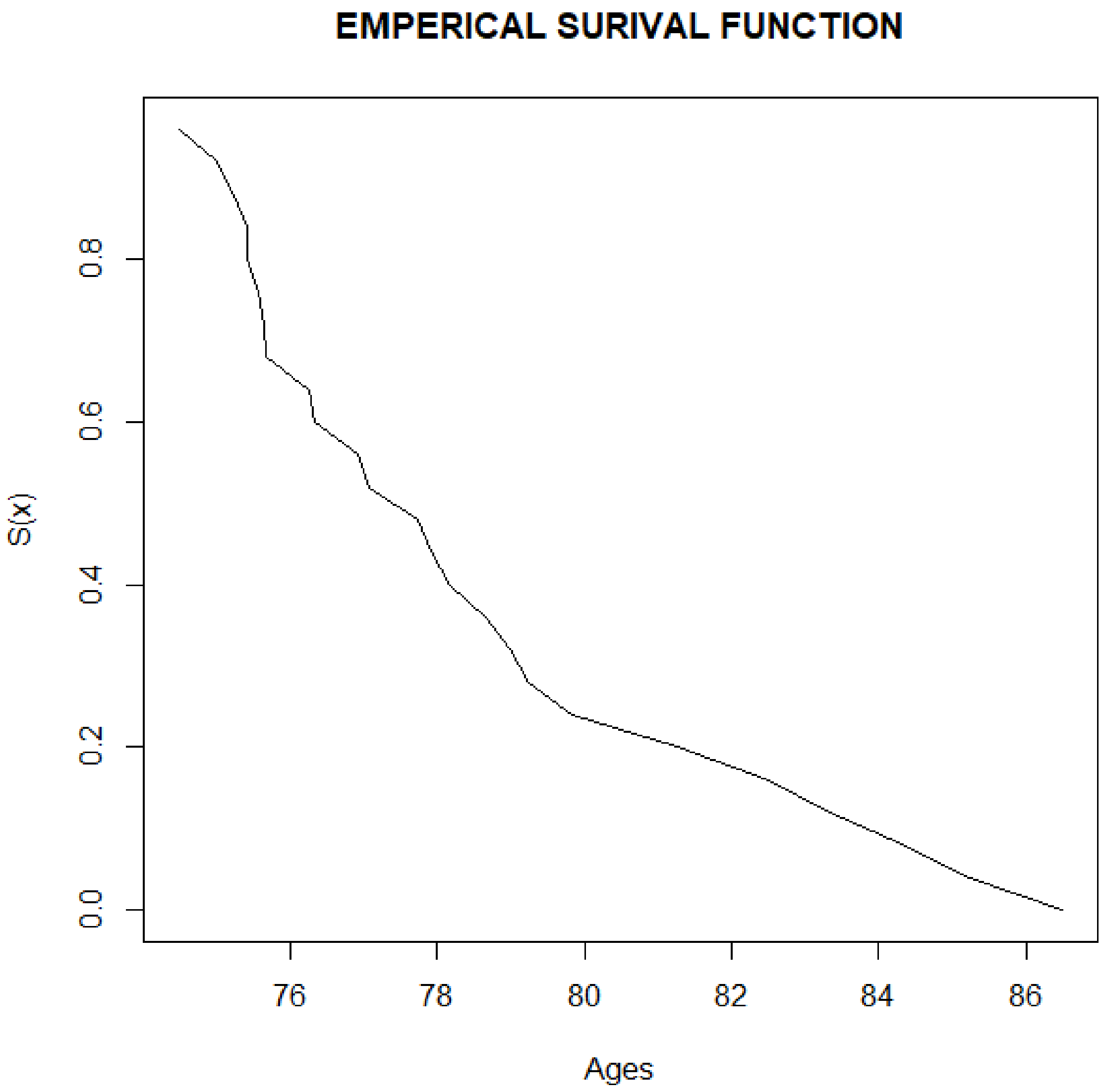
```
dementia_data =
c(0.5,1,1.25,1.41,1.42,1.58,1.66,1.67,2.25,2.33,2.92,3.08,3.75,3.92,4.17,4.67,5,
5.25,5.83,7.25,8.50,9.33,10.33,11.25,12.5)
n_dd = 74 + dementia_data
q = length(dementia_data)
x = rep(1, q)
y = cumsum(x)
z = y / q
est = 1 - z
est
t = data.frame(n_dd,est);t
plot(type='s',n_dd,est,xlab="Ages",ylab="S(x)",main = 'EMPERICAL SURIVAL
FUNCTION')
```

R-Console:-

```
> dementia_data =
c(0.5,1,1.25,1.41,1.42,1.58,1.66,1.67,2.25,2.33,2.92,3.08,3.75,3.92,4.17,4.67,5,
5.25,5.83,7.25,8.50,9.33,10.33,11.25,12.5)
> n_dd = 74 + dementia_data
> q = length(dementia_data)
> x = rep(1, q)
> y = cumsum(x)
> z = y / q
> est = 1 - z
> est
 [1] 0.96 0.92 0.88 0.84 0.80 0.76 0.72 0.68 0.64 0.60 0.56 0.52 0.48 0.44
[15] 0.40 0.36 0.32 0.28 0.24 0.20 0.16 0.12 0.08 0.04 0.00
> t = data.frame(n_dd,est);t
    n_dd  est
1  74.50 0.96
2  75.00 0.92
3  75.25 0.88
4  75.41 0.84
5  75.42 0.80
6  75.58 0.76
7  75.66 0.72
8  75.67 0.68
9  76.25 0.64
10 76.33 0.60
11 76.92 0.56
12 77.08 0.52
13 77.75 0.48
14 77.92 0.44
15 78.17 0.40
16 78.67 0.36
```

```
17 79.00 0.32
18 79.25 0.28
19 79.83 0.24
20 81.25 0.20
21 82.50 0.16
22 83.33 0.12
23 84.33 0.08
24 85.25 0.04
25 86.50 0.00
```

Graph:-



Problem 20:

It appears from the above Table that the length of survival time after diagnosis of dementia is not affected by age of diagnosis in each group. Consider the 97 cases in the above Table as a random sample from a fairly large and homogeneous population. Group the mortality data into fixed intervals $h_i = 1$ year. Estimate q_i and p_i from the grouped data. Calculate standard errors for these estimates.

Objective:- We must estimate q_i and p_i from the grouped data. Calculate standard errors for these estimates.

Theory:- When the sample is large enough, the data can be grouped into M fixed intervals, $[t_i$, $t_{i+1})$, i=0, 1, ... ,M – 1. The length of the interval is h_i = t_{i+1} − t_i. It is convenient to have intervals of the same length, but this is not essential.

N_i is the number of survivors at the beginning of the interval $[t_i , t_{i+1})$, In the present case we have

$$N_i = \sum_{j=1}^{M-1} d_j = \sum_{j=1}^{\infty} d_j$$

since d_j for j greater than $M-1$. In particular,

$$N_0 = \sum_{j=0}^{M-1} d_j = N$$

where, N_0 is the total sample size, that is, the size of the group that starts at time t_0

The observed proportion of survivors at time t_i out of N_0 starters is

$$\hat{p}_i = \frac{N_i}{N_0}$$

We may also estimate the conditional probability of death in the interval $[t_i , t_{i+1})$, given alive at t_i.

$$\hat{q}_i = \frac{d_j}{N_i} \ \& \ \hat{p}_i = 1 - \hat{q}_i$$

R-Code:-

```
rm(list=ls())
Surv_Time70 = c(0.50,0.83,1.17,1.25, 1.66, 1.67, 1.83, 3.17, 4.17, 4.92, 5.58,
6.58, 6.92,8.16, 1.50)
Age_at_death70 = 70+Surv_Time70
Surv_Time71 = c(0.50, 0.83, 1.25, 1.33, 1.67, 1.83, 3.42, 3.42, 3.92, 4.92,
5.75, 6.92,7.83,7.84, 9.00, 9.17, 13.83, 18.08)
Age_at_death71 = 71+Surv_Time71
Surv_Time72 = c(0.50, 1.00, 1.08, 1.42, 1.58, 1.59, 2.00, 2.33, 2.67, 3.58,
4.08, 4.17,5.75,6.25, 7.00, 7.08, 9.92, 10.17, 11.25, 12.67, 21.00, 21.83)
Age_at_death72 = 72+Surv_Time72
Surv_Time73 = c(0.58, 1.08, 1.25, 1.67, 2.00, 2.08, 2.17, 3.75, 4.25, 4.58,
4.92, 5.25,6.75,7.83, 7.84, 9.17, 11.50)
Age_at_death73 = 73+Surv_Time73
Surv_Time74=c(0.50,1.00,1.25,1.41,1.42,1.58,1.66,1.67,2.25,2.33,2.92,3.08,3.75,3
.92,4.17,4.67,5.00,5.25,5.83,7.25,8.50,9.33,10.33,11.25,12.50)
```

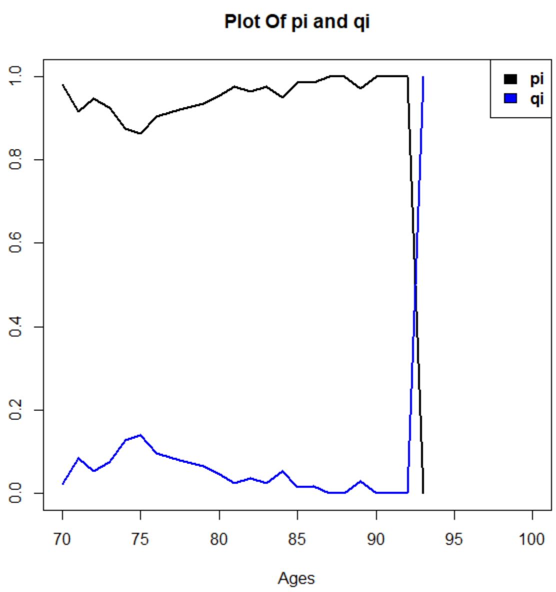
```
Age_at_death74 = 74+Surv_Time74
Age_at_death=c(Age_at_death70,Age_at_death71,Age_at_death72,Age_at_death73,Age_a
t_death74)
N = length(Age_at_death)
Range = round(range(Age_at_death))
Age_Intervals = seq(Range[1],Range[2],by=1)
Age_as_interval = cut(Age_at_death,Age_Intervals,right = F)
Age_Interval = levels(Age_as_interval)
Deaths_Freq = table(Age_as_interval)
di = as.vector(Deaths_Freq)
Ni = sort(cumsum(di),decreasing = T)
qi = di/Ni
pi = 1-qi
data.frame(Age_Interval,di,Ni,qi,pi)
Age_Intervals = Age_Intervals[1:24]
plot(Age_Intervals,pi,type = "l",
 main = paste("Plot Of pi and qi"),xlab="Ages",ylab="",ylim = c(0,1),xlim =
c(70,100),col="black",lwd=2)
lines(Age_Intervals,qi,col="blue",lwd=2)
legend("topright",legend=c(expression(bold("pi")),
 expression(bold("qi"))),fill=c("black","blue"))
```

R-Console:-

```
> rm(list=ls())
> Surv_Time70 = c(0.50,0.83,1.17,1.25, 1.66, 1.67, 1.83, 3.17, 4.17, 4.92, 5.58,
6.58, 6.92,8.16, 1.50)
> Age_at_death70 = 70+Surv_Time70
> Surv_Time71 = c(0.50, 0.83, 1.25, 1.33, 1.67, 1.83, 3.42, 3.42, 3.92, 4.92,
5.75, 6.92,7.83,7.84, 9.00, 9.17, 13.83, 18.08)
> Age_at_death71 = 71+Surv_Time71
> Surv_Time72 = c(0.50, 1.00, 1.08, 1.42, 1.58, 1.59, 2.00, 2.33, 2.67, 3.58,
4.08, 4.17,5.75,6.25, 7.00, 7.08, 9.92, 10.17, 11.25, 12.67, 21.00, 21.83)
> Age_at_death72 = 72+Surv_Time72
> Surv_Time73 = c(0.58, 1.08, 1.25, 1.67, 2.00, 2.08, 2.17, 3.75, 4.25, 4.58,
4.92, 5.25,6.75,7.83, 7.84, 9.17, 11.50)
> Age_at_death73 = 73+Surv_Time73
> Surv_Time74 =
c(0.50,1.00,1.25,1.41,1.42,1.58,1.66,1.67,2.25,2.33,2.92,3.08,3.75,3.92,4.17,4.6
7,5.00,5.25,5.83,7.25,8.50,9.33,10.33,11.25,12.50)
> Age_at_death74 = 74+Surv_Time74
>
Age_at_death=c(Age_at_death70,Age_at_death71,Age_at_death72,Age_at_death73,Age_a
t_death74)
> N = length(Age_at_death)
> Range = round(range(Age_at_death))
> Age_Intervals = seq(Range[1],Range[2],by=1)
> Age_as_interval = cut(Age_at_death,Age_Intervals,right = F)
> Age_Interval = levels(Age_as_interval)
> Deaths_Freq = table(Age_as_interval)
> di = as.vector(Deaths_Freq)
> Ni = sort(cumsum(di),decreasing = T)
> qi = di/Ni
> pi = 1-qi
> data.frame(Age_Interval,di,Ni,qi,pi)
```

```
   Age_Interval di Ni         qi        pi
1       [70,71)  2 97 0.02061856 0.9793814
2       [71,72)  8 95 0.08421053 0.9157895
3       [72,73)  5 95 0.05263158 0.9473684
4       [73,74)  7 95 0.07368421 0.9263158
5       [74,75) 12 95 0.12631579 0.8736842
6       [75,76) 13 94 0.13829787 0.8617021
7       [76,77)  9 94 0.09574468 0.9042553
8       [77,78)  8 94 0.08510638 0.9148936
9       [78,79)  7 93 0.07526882 0.9247312
10      [79,80)  6 92 0.06521739 0.9347826
11      [80,81)  4 88 0.04545455 0.9545455
12      [81,82)  2 86 0.02325581 0.9767442
13      [82,83)  3 83 0.03614458 0.9638554
14      [83,84)  2 81 0.02469136 0.9753086
15      [84,85)  4 77 0.05194805 0.9480519
16      [85,86)  1 71 0.01408451 0.9859155
17      [86,87)  1 64 0.01562500 0.9843750
18      [87,88)  0 56 0.00000000 1.0000000
19      [88,89)  0 47 0.00000000 1.0000000
20      [89,90)  1 34 0.02941176 0.9705882
21      [90,91)  0 22 0.00000000 1.0000000
22      [91,92)  0 15 0.00000000 1.0000000
23      [92,93)  0 10 0.00000000 1.0000000
24      [93,94)  2  2 1.00000000 0.0000000
> Age_Intervals = Age_Intervals[1:24]
> plot(Age_Intervals,pi,type = "l",
+  main = paste("Plot Of pi and qi"),xlab="Ages",ylab="",ylim = c(0,1),xlim =
+ c(70,100),col="black",lwd=2)
> lines(Age_Intervals,qi,col="blue",lwd=2)
> legend("topright",legend=c(expression(bold("pi")),
+  expression(bold("qi"))),fill=c("black","blue"))
```

Graph:-



Problem 21

Estimate P_i [observed proportion of survivors at time t_i out of N_0 starters] from the group data and calculate the standard errors for the estimate. Estimate the curve of deaths and force of mortality. Represent estimates of survival function, curve of deaths and force of mortality on separate graphs. Draw some conclusions.

Objective:- We must estimate the curve of deaths and force of mortality, and represent estimates of survival function, curve of deaths and force of mortality on separate graphs.

Theory:- Suppose that a random variable X denotes the lifetime of a living Organism or an inanimate device, it is also called as at death or at failure or briefly age.

1. The cumulative distribution function (CDF); $F_x(x) = P_t(X \leq x)$; is called the lifetime distribution or failure distribution.
2. The complementary function of cumulative distribution function (CDF) is known as survival distribution function (SDF). And SDF = 1 – CDF.
$$SDF = 1 - P_t(X \leq x) = P_t(X > x)$$
3. *The Hazard Function(force to mortality):-* it is the instantaneous or relative failure at time point x. And given as
$$\lambda_X(x) = \frac{f_X(x)}{S_X(x)}$$
$$\Rightarrow \lambda_X(x) = -\frac{dlogS_X(x)}{dx}$$
$$\Rightarrow -dlogS_X(x) = \lambda_X(x)dx$$
$$\Rightarrow logS_X(x) = \int_0^x \lambda_X(x)dx$$
4. Let $t_1 < t_2 < ... < t_N$; represents the N(Distinct) ordered times at death. Let $F_T(t) = P_r(T \leq t)$; be the theoretical (unknown) failure distribution (CDF), and

$$S_T(t) = 1 - P_r(T \leq t) = P_t(T > t)$$

be the corresponding survival distribution function (SOF)

The Empirical Distribution Function is

$$F_N^o(t) = \begin{cases} 0, & \text{for } t < t_1^{'} \\ \dfrac{i}{N}, & \text{for } t_i^{'} \leq t < t_{i+1}^{'} \\ 1, & \text{for } t \geq t_N^{'} \end{cases}$$

Where i is the rank of the i^{th} (ordered) observation, and $F_N^o(t)$ gives the probability $P_r(T \leq t)$.

And the Empirical Survival Function is given as $S_N^o(t) = 1 - F_N^o(t)$.

$$S_N^o(t) = \begin{cases} 1, & \text{for } t < t_1^{'} \\ \dfrac{N-i}{N}, & \text{for } t_i^{'} \leq t < t_{i+1}^{'} \\ 0, & \text{for } t \geq t_N^{'} \end{cases}$$

$S_N^o(t)$ gives the probability $P_r(T > t)$.

5. When the sample is large enough, the data can be grouped into M fixed intervals, $[t_i, t_{i+1})$, i=0, 1, ... ,M – 1. The length of the interval is $h_i = t_{i+1} - t_i$. It is convenient to have intervals of the same length, but this is not essential.

   N_i is the number of survivors at the beginning of the interval $[t_i, t_{i+1})$, In the present case we have

$$N_i = \sum_{j=1}^{M-1} d_j = \sum_{j=1}^{\infty} d_j$$

since d_j for j greater than $M - 1$. In particular,

$$N_0 = \sum_{j=0}^{M-1} d_j = N$$

where, N_0 is the total sample size, that is, the size of the group that starts at time t_0

The observed proportion of survivors at time t_i out of N_0 starters is

$$\hat{p}_i = \frac{N_i}{N_0}$$

We may also estimate the conditional probability of death in the interval $[t_i, t_{i+1})$, given alive at t_i.

$$\hat{q}_i = \frac{d_j}{N_i} \text{ \& } \hat{p}_i = 1 - \hat{q}_i$$

R-Code:-

```
rm(list=ls())
Surv_Time70 = c(0.50,0.83,1.17,1.25, 1.66, 1.67, 1.83, 3.17, 4.17, 4.92, 5.58,
6.58, 6.92,8.16, 1.50)
Age_at_death70 = 70+Surv_Time70
Surv_Time71 = c(0.50, 0.83, 1.25, 1.33, 1.67, 1.83, 3.42, 3.42, 3.92, 4.92,
5.75, 6.92,7.83,7.84, 9.00, 9.17, 13.83, 18.08)
Age_at_death71 = 71+Surv_Time71
Surv_Time72 = c(0.50, 1.00, 1.08, 1.42, 1.58, 1.59, 2.00, 2.33, 2.67, 3.58,
4.08, 4.17,5.75,6.25, 7.00, 7.08, 9.92, 10.17, 11.25, 12.67, 21.00, 21.83)
Age_at_death72 = 72+Surv_Time72
Surv_Time73 = c(0.58, 1.08, 1.25, 1.67, 2.00, 2.08, 2.17, 3.75, 4.25, 4.58,
4.92, 5.25,6.75,7.83, 7.84, 9.17, 11.50)
Age_at_death73 = 73+Surv_Time73
Surv_Time74 =
c(0.50,1.00,1.25,1.41,1.42,1.58,1.66,1.67,2.25,2.33,2.92,3.08,3.75,3.92,4.17,4.6
7,5.00,5.25,5.83,7.25,8.50,9.33,10.33,11.25,12.50)
Age_at_death74 = 74+Surv_Time74
Age_at_death =
c(Age_at_death70,Age_at_death71,Age_at_death72,Age_at_death73,Age_at_death74)
N = length(Age_at_death)
Range = round(range(Age_at_death))
hi =1
Age_Intervals = seq(Range[1],Range[2],by=hi)
Age_as_interval = cut(Age_at_death,Age_Intervals,right = F)
Age_Interval = levels(Age_as_interval)
Deaths_Freq = table(Age_as_interval)
di = as.vector(Deaths_Freq)
Ni = sort(cumsum(di),decreasing = T)
qi = di/Ni
pi = 1-qi
Pi = Ni/N
```

```
SE_Pi = Pi*(1-Pi)/N #Standard error of Pi
f_ti_star = di/(hi*N) #curve of deaths
mu_ti_star = (-1/hi)*(log((Ni - di)/Ni)) #force of mortality
#required mortality table
data.frame(Age_Interval,di,Ni,qi,pi,Pi,SE_Pi,f_ti_star,mu_ti_star)
Age_Intervals = Age_Intervals[1:24]
plot(Age_Intervals,Pi,type = "l",
 main = paste("Plot Of Pi"),xlab="Ages",ylab="Pi",ylim = c(0,1),xlim =
c(70,100),col="black",lwd=2)
lines(Age_Intervals,f_ti_star,col = "red")
lines(Age_Intervals,mu_ti_star,col = "blue")
legend("topright",legend = c("Pi","f_ti_star","mu_ti_star"),fill =
c("black","red","blue"))
```

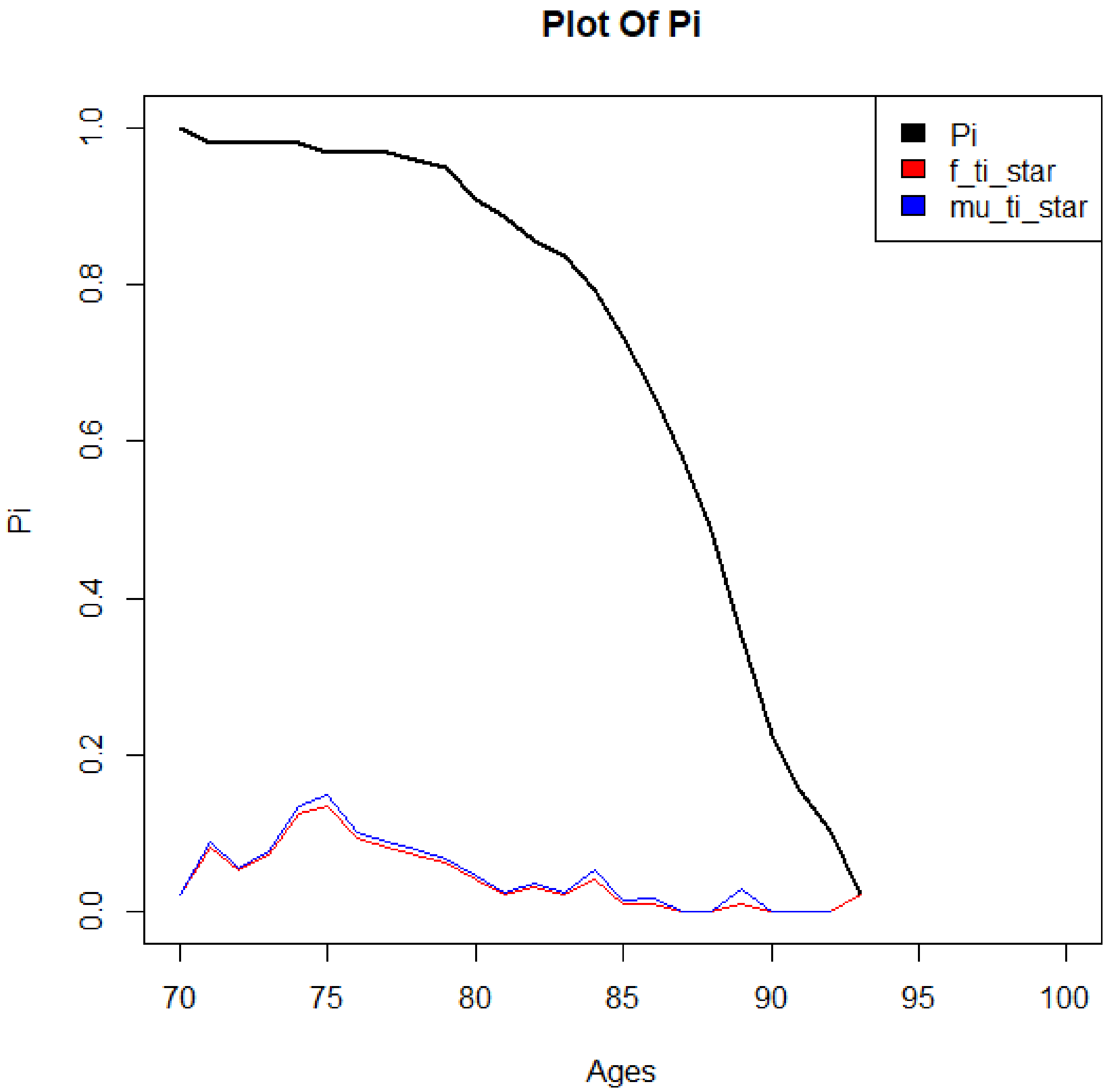
R_Console:-

```
> rm(list=ls())
> Surv_Time70 = c(0.50,0.83,1.17,1.25, 1.66, 1.67, 1.83, 3.17, 4.17, 4.92, 5.58,
6.58, 6.92,8.16, 1.50)
> Age_at_death70 = 70+Surv_Time70
> Surv_Time71 = c(0.50, 0.83, 1.25, 1.33, 1.67, 1.83, 3.42, 3.42, 3.92, 4.92,
5.75, 6.92,7.83,7.84, 9.00, 9.17, 13.83, 18.08)
> Age_at_death71 = 71+Surv_Time71
> Surv_Time72 = c(0.50, 1.00, 1.08, 1.42, 1.58, 1.59, 2.00, 2.33, 2.67, 3.58,
4.08, 4.17,5.75,6.25, 7.00, 7.08, 9.92, 10.17, 11.25, 12.67, 21.00, 21.83)
> Age_at_death72 = 72+Surv_Time72
> Surv_Time73 = c(0.58, 1.08, 1.25, 1.67, 2.00, 2.08, 2.17, 3.75, 4.25, 4.58,
4.92, 5.25,6.75,7.83, 7.84, 9.17, 11.50)
> Age_at_death73 = 73+Surv_Time73
> Surv_Time74 =
c(0.50,1.00,1.25,1.41,1.42,1.58,1.66,1.67,2.25,2.33,2.92,3.08,3.75,3.92,4.17,4.6
7,5.00,5.25,5.83,7.25,8.50,9.33,10.33,11.25,12.50)
> Age_at_death74 = 74+Surv_Time74
> Age_at_death =
c(Age_at_death70,Age_at_death71,Age_at_death72,Age_at_death73,Age_at_death74)
> N = length(Age_at_death)
> Range = round(range(Age_at_death))
> hi =1
> Age_Intervals = seq(Range[1],Range[2],by=hi)
> Age_as_interval = cut(Age_at_death,Age_Intervals,right = F)
> Age_Interval = levels(Age_as_interval)
> Deaths_Freq = table(Age_as_interval)
> di = as.vector(Deaths_Freq)
> Ni = sort(cumsum(di),decreasing = T)
> qi = di/Ni
> pi = 1-qi
> Pi = Ni/N
> SE_Pi = Pi*(1-Pi)/N #Standard error of Pi
> f_ti_star = di/(hi*N) #curve of deaths
> mu_ti_star = (-1/hi)*(log((Ni - di)/Ni)) #force of mortality
> #required mortality table
> data.frame(Age_Interval,di,Ni,qi,pi,Pi,SE_Pi,f_ti_star,mu_ti_star)
```

```
   Age_Interval di Ni         qi        pi          Pi        SE_Pi  f_ti_star
mu_ti_star
1        [70,71)  2 97 0.02061856 0.9793814 1.00000000 0.0000000000 0.02061856
0.02083409
2        [71,72)  8 95 0.08421053 0.9157895 0.97938144 0.0002081797 0.08247423
0.08796877
3        [72,73)  5 95 0.05263158 0.9473684 0.97938144 0.0002081797 0.05154639
0.05406722
4        [73,74)  7 95 0.07368421 0.9263158 0.97938144 0.0002081797 0.07216495
0.07654008
5        [74,75) 12 95 0.12631579 0.8736842 0.97938144 0.0002081797 0.12371134
0.13503628
6        [75,76) 13 94 0.13829787 0.8617021 0.96907216 0.0003089825 0.13402062
0.14884563
7        [76,77)  9 94 0.09574468 0.9042553 0.96907216 0.0003089825 0.09278351
0.10064353
8        [77,78)  8 94 0.08510638 0.9148936 0.96907216 0.0003089825 0.08247423
0.08894749
9        [78,79)  7 93 0.07526882 0.9247312 0.95876289 0.0004075940 0.07216495
0.07825220
10       [79,80)  6 92 0.06521739 0.9347826 0.94845361 0.0005040140 0.06185567
0.06744128
11       [80,81)  4 88 0.04545455 0.9545455 0.90721649 0.0008677807 0.04123711
0.04652002
12       [81,82)  2 86 0.02325581 0.9767442 0.88659794 0.0010365158 0.02061856
0.02353050
13       [82,83)  3 83 0.03614458 0.9638554 0.85567010 0.0012731833 0.03092784
0.03681397
14       [83,84)  2 81 0.02469136 0.9753086 0.83505155 0.0014200048 0.02061856
0.02500130
15       [84,85)  4 77 0.05194805 0.9480519 0.79381443 0.0016873513 0.04123711
0.05334598
16       [85,86)  1 71 0.01408451 0.9859155 0.73195876 0.0020226302 0.01030928
0.01418463
17       [86,87)  1 64 0.01562500 0.9843750 0.65979381 0.0023140818 0.01030928
0.01574836
18       [87,88)  0 56 0.00000000 1.0000000 0.57731959 0.0025156874 0.00000000
0.00000000
19       [88,89)  0 47 0.00000000 1.0000000 0.48453608 0.0025748543 0.00000000
0.00000000
20       [89,90)  1 34 0.02941176 0.9705882 0.35051546 0.0023469523 0.01030928
0.02985296
21       [90,91)  0 22 0.00000000 1.0000000 0.22680412 0.0018078764 0.00000000
0.00000000
22       [91,92)  0 15 0.00000000 1.0000000 0.15463918 0.0013476897 0.00000000
0.00000000
23       [92,93)  0 10 0.00000000 1.0000000 0.10309278 0.0009532439 0.00000000
0.00000000
24       [93,94)  2  2 1.00000000 0.0000000 0.02061856 0.0002081797 0.02061856
       Inf
> Age_Intervals = Age_Intervals[1:24]
> plot(Age_Intervals,Pi,type = "l",
+  main = paste("Plot Of Pi"),xlab="Ages",ylab="Pi",ylim = c(0,1),xlim =
c(70,100),col="black",lwd=2)
> lines(Age_Intervals,f_ti_star,col = "red")
```

```
> lines(Age_Intervals,mu_ti_star,col = "blue")
> legend("topright",legend = c("Pi","f_ti_star","mu_ti_star"),fill =
c("black","red","blue"))
```

Graph:-



Problem 22

The data below represent the ordered times at death t_j'(in days) of 43 patients suffering from granulocytic leukemia with t_0' = 0 taken as date of diagnosis.

7, 47, 58, 74, 177, 232, 273, 285, 317, 429, 440, 445, 455, 468, 495, 497, 532, 571, 579, 581, 650, 702, 715, 779, 881, 900, 930, 968, 1077, 1109, 1314, 1334, 1367, 1534, 1712, 1784, 1877, 1886, 2045, 2056, 2260, 2429, 2509

Calculate the empirical survival function and represent it graphically. Group the data in 200 days. Evaluate the SDF for the grouped data and plot it on the same graph. Estimate the PDF from grouped data. Draw a histogram and the estimated PDF on the same graph.

Objective:- We must calculate the empirical survival function and represent it graphically. Group the data in 200 days. Evaluate the SDF for the grouped data and plot it on the same graph. Estimate the PDF from grouped data. Draw a histogram and the estimated PDF on the same graph.

Theory:- Suppose that a random variable X denotes the lifetime of a living Organism or an inanimate device, it is also called as at death or at failure or briefly age.

1. The cumulative distribution function (CDF); $F_x(x) = P_t(X \leq x)$; is called the lifetime distribution or failure distribution.
2. The complementary function of cumulative distribution function (CDF) is known as survival distribution function (SDF). And SDF = 1 – CDF.
$$SDF = 1 - P_t(X \leq x) = P_t(X > x)$$
3. *The Hazard Function(force to mortality):*- it is the instantaneous or relative failure at time point x. And given as
$$\lambda_X(x) = \frac{f_X(x)}{S_X(x)}$$
$$\Rightarrow \lambda_X(x) = -\frac{dlogS_X(x)}{dx}$$
$$\Rightarrow -dlogS_X(x) = \lambda_X(x)dx$$
$$\Rightarrow logS_X(x) = \int_0^x \lambda_X(x)dx$$
4. Let $t_1 < t_2 < \ldots < t_N$; represents the N(Distinct) ordered times at death. Let $F_T(t) = P_r(T \leq t)$; be the theoretical (unknown) failure distribution (CDF), and

$S_T(t) = 1 - P_r(T \leq t) = P_t(T > t)$

be the corresponding survival distribution function (SOF)

The Empirical Distribution Function is

$$F_N^o(t) = \begin{cases} 0, & \text{for } t < t_1' \\ \dfrac{i}{N}, & \text{for } t_i' \leq t < t_{i+1}' \\ 1, & \text{for } t \geq t_N' \end{cases}$$

Where i is the rank of the i^{th} (ordered) observation, and $F_N^o(t)$ gives the probability $P_r(T \leq t)$.

And the Empirical Survival Function is given as $S_N^o(t) = 1 - F_N^o(t)$.

$$S_N^o(t) = \begin{cases} 1, & \text{for } t < t_1' \\ \dfrac{N-i}{N}, & \text{for } t_i' \leq t < t_{i+1}' \\ 0, & \text{for } t \geq t_N' \end{cases}$$

$S_N^o(t)$ gives the probability $P_r(T > t)$.

5. When the sample is large enough, the data can be grouped into M fixed intervals, $[t_i, t_{i+1})$, i=0, 1, ... ,M – 1. The length of the interval is $h_i = t_{i+1} - t_i$. It is convenient to have intervals of the same length, but this is not essential.

   N_i is the number of survivors at the beginning of the interval $[t_i, t_{i+1})$, In the present case we have

$$N_i = \sum_{j=1}^{M-1} d_j = \sum_{j=1}^{\infty} d_j$$

since d_j for j greater than $M-1$. In particular,

$$N_0 = \sum_{j=0}^{M-1} d_j = N$$

where, N_0 is the total sample size, that is, the size of the group that starts at time t_0

The observed proportion of survivors at time t_i out of N_0 starters is

$$\hat{p}_i = \frac{N_i}{N_0}$$

We may also estimate the conditional probability of death in the interval $[t_i, t_{i+1})$, given alive at t_i.

$$\hat{q}_i = \frac{d_j}{N_i} \; \& \; \hat{p}_i = 1 - \hat{q}_i$$

R_Code:-

```
rm(list=ls())
ti = c(7, 47, 58, 74, 177, 232, 273, 285, 317, 429, 440, 445, 455, 468, 495,
497, 532, 571,579, 581 ,650, 702, 715, 779, 881, 900,930, 968, 1077, 1109, 1314,
1334, 1367, 1534,1712, 1784, 1877, 1886, 2045, 2056, 2260, 2429, 2509)
di = table(ti)
N = length(ti)
cum_sum_di = cumsum(di)
cum_sum_survivors = (N - cum_sum_di)
Emp_Surv_Fun = cum_sum_survivors/N
Emp_Surv_Fun =as.vector(Emp_Surv_Fun)
data.frame(ti,Emp_Surv_Fun)
Range = round(range(ti))
hi =200 #interval of Age
Intervals = seq(Range[1],Range[2],by=hi)
Intervals = c(Intervals,2607)
ti_as_interval = cut(ti,Intervals,right = F)
ti_Interval = levels(ti_as_interval)
ti_Freq = table(ti_as_interval)
di = as.vector(ti_Freq)
Ni = sort(cumsum(di),decreasing = T)
qi = di/Ni
pi = 1-qi
Pi = Ni/N
f_ti_star = di/(hi*N) #curve of deaths
```

```
data.frame(ti_Interval,di,Ni,qi,pi,Pi,f_ti_star)
Intervals = Intervals[1:(length(Intervals)-1)]
plot(ti,Emp_Surv_Fun,type = "l",
 main = paste("Plot Of Empirical Survival Function"),xlab="ti",ylab="S(x)",ylim
=
c(0,1),col="black",lwd=2)
lines(Intervals,Pi,type = "l",
 main = paste("Plot Of Pi"),xlab="Ages",ylab="Pi",col="green",lwd=2)
#Drawing The Legend
legend("topright",legend=c(expression(bold("Empirical Survival Function")),
 expression(bold("Survival function of grouped
data"))),fill=c("black","green"))
#Plot of force of mortality
hist(ti,xlab="ti",lwd=2,probability = T,ylim=c(0,0.0015),col="green")
lines(Intervals,f_ti_star,type = "l",
 main = paste("Plot Of f_ti_star"),xlab="ti",ylab="f_ti_star",col="black",lwd=2)
legend("topright",legend=c(expression(bold("histogram of deaths")),
 expression(bold("pdf (curve of deaths)"))),fill=c("green","black"))
```

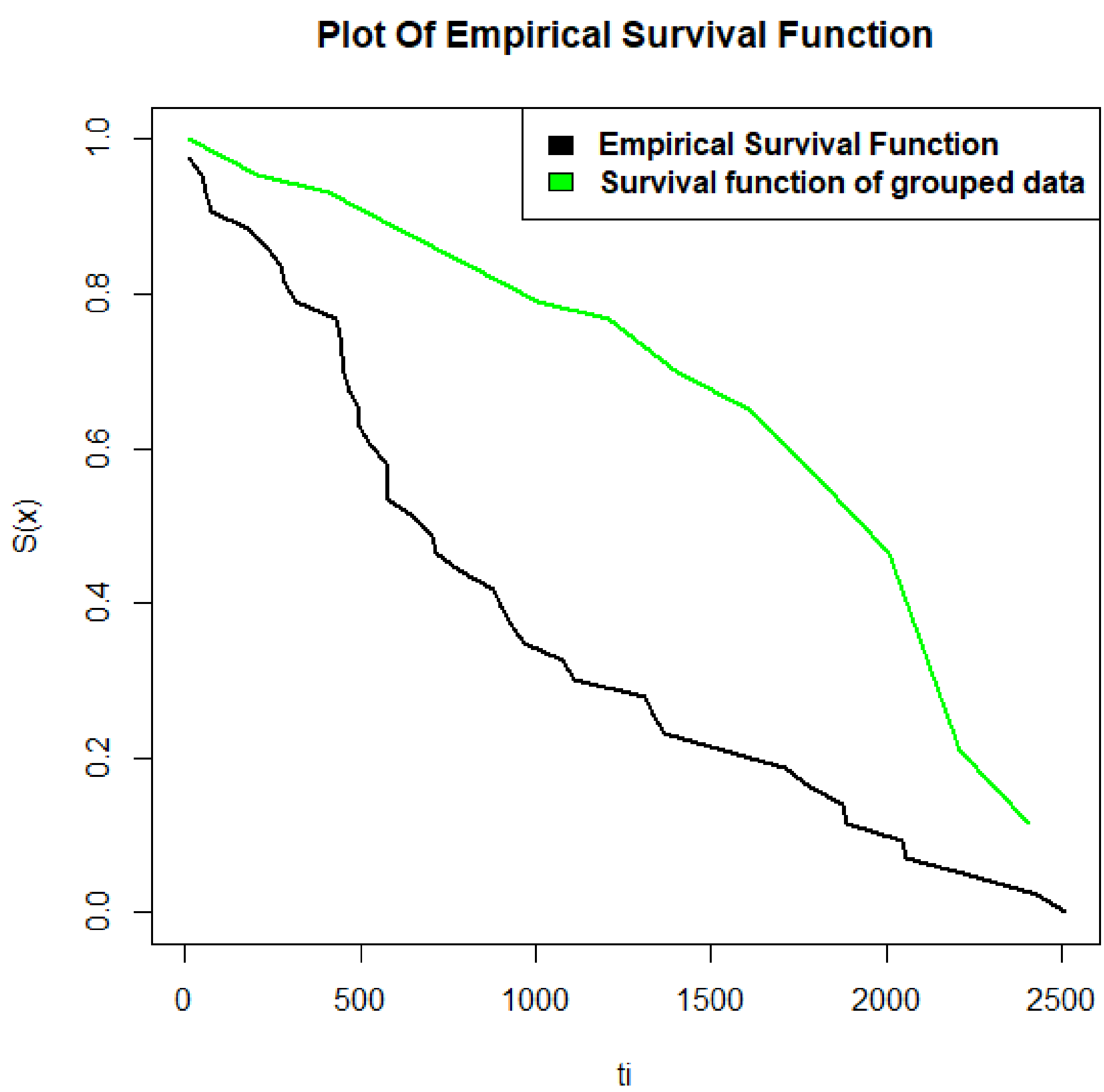
R-Console:-

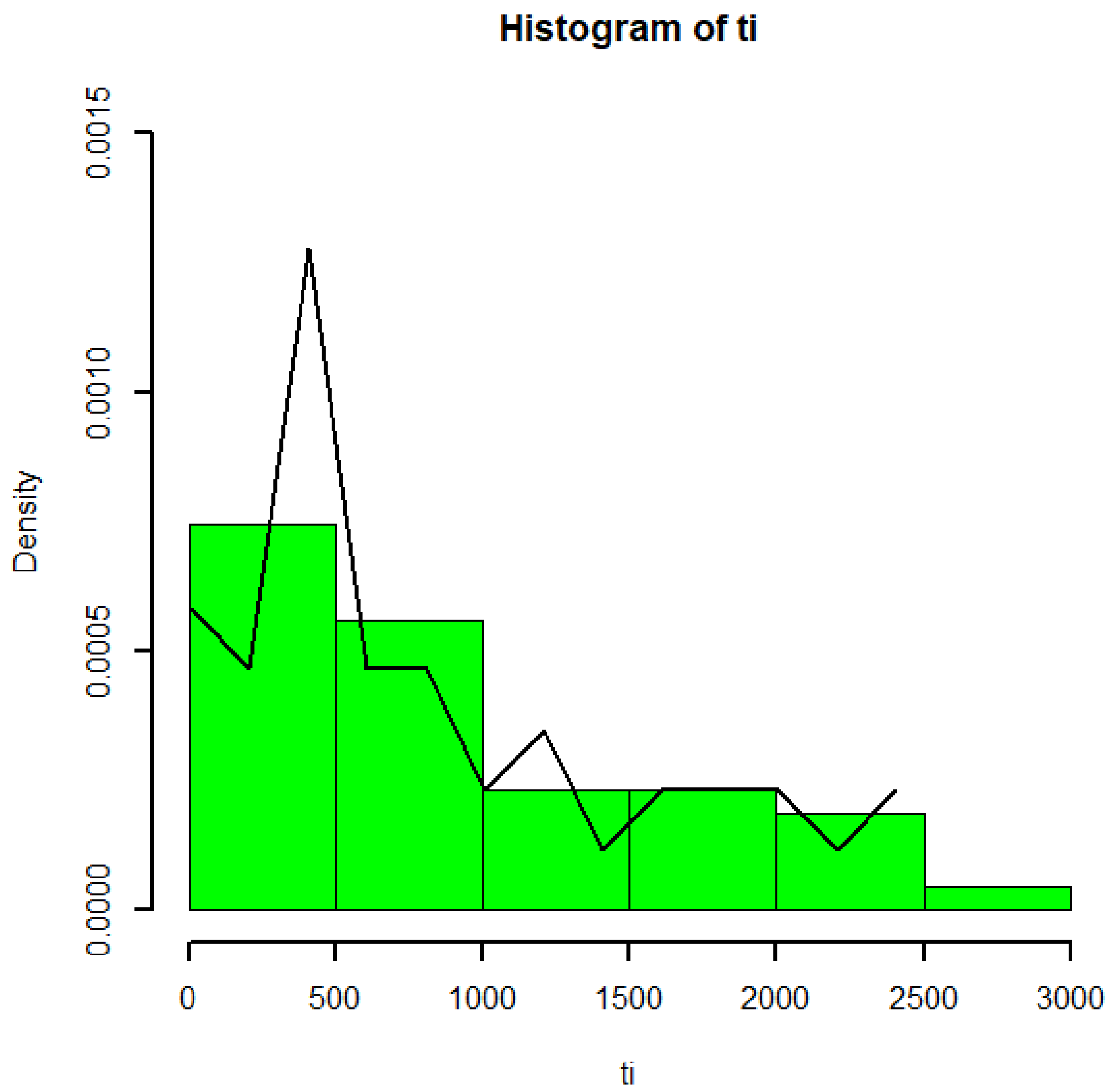
```
> rm(list=ls())
> ti = c(7, 47, 58, 74, 177, 232, 273, 285, 317, 429, 440, 445, 455, 468, 495,
497, 532, 571,579, 581 ,650, 702, 715, 779, 881, 900,930, 968, 1077, 1109, 1314,
1334, 1367, 1534,1712, 1784, 1877, 1886, 2045, 2056, 2260, 2429, 2509)
> di = table(ti)
> N = length(ti)
> cum_sum_di = cumsum(di)
> cum_sum_survivors = (N - cum_sum_di)
> Emp_Surv_Fun = cum_sum_survivors/N
> Emp_Surv_Fun =as.vector(Emp_Surv_Fun)
> data.frame(ti,Emp_Surv_Fun)
     ti  Emp_Surv_Fun
1     7    0.97674419
2    47    0.95348837
3    58    0.93023256
4    74    0.90697674
5   177    0.88372093
6   232    0.86046512
7   273    0.83720930
8   285    0.81395349
9   317    0.79069767
10  429    0.76744186
11  440    0.74418605
12  445    0.72093023
13  455    0.69767442
14  468    0.67441860
15  495    0.65116279
16  497    0.62790698
17  532    0.60465116
18  571    0.58139535
19  579    0.55813953
20  581    0.53488372
```

```
21  650   0.51162791
22  702   0.48837209
23  715   0.46511628
24  779   0.44186047
25  881   0.41860465
26  900   0.39534884
27  930   0.37209302
28  968   0.34883721
29 1077   0.32558140
30 1109   0.30232558
31 1314   0.27906977
32 1334   0.25581395
33 1367   0.23255814
34 1534   0.20930233
35 1712   0.18604651
36 1784   0.16279070
37 1877   0.13953488
38 1886   0.11627907
39 2045   0.09302326
40 2056   0.06976744
41 2260   0.04651163
42 2429   0.02325581
43 2509   0.00000000
> Range = round(range(ti))
> hi =200 #interval of Age
> Intervals = seq(Range[1],Range[2],by=hi)
> Intervals = c(Intervals,2607)
> ti_as_interval = cut(ti,Intervals,right = F)
> ti_Interval = levels(ti_as_interval)
> ti_Freq = table(ti_as_interval)
> di = as.vector(ti_Freq)
> Ni = sort(cumsum(di),decreasing = T)
> qi = di/Ni
> pi = 1-qi
> Pi = Ni/N
> f_ti_star = di/(hi*N) #curve of deaths
> data.frame(ti_Interval,di,Ni,qi,pi,Pi,f_ti_star)
           ti_Interval di Ni         qi        pi        Pi    f_ti_star
1              [7,207)  5 43 0.11627907 0.8837209 1.0000000 0.0005813953
2            [207,407)  4 41 0.09756098 0.9024390 0.9534884 0.0004651163
3            [407,607) 11 40 0.27500000 0.7250000 0.9302326 0.0012790698
4            [607,807)  4 38 0.10526316 0.8947368 0.8837209 0.0004651163
5       [807,1.01e+03)  4 36 0.11111111 0.8888889 0.8372093 0.0004651163
6  [1.01e+03,1.21e+03)  2 34 0.05882353 0.9411765 0.7906977 0.0002325581
7  [1.21e+03,1.41e+03)  3 33 0.09090909 0.9090909 0.7674419 0.0003488372
8  [1.41e+03,1.61e+03)  1 30 0.03333333 0.9666667 0.6976744 0.0001162791
9  [1.61e+03,1.81e+03)  2 28 0.07142857 0.9285714 0.6511628 0.0002325581
10 [1.81e+03,2.01e+03)  2 24 0.08333333 0.9166667 0.5581395 0.0002325581
11 [2.01e+03,2.21e+03)  2 20 0.10000000 0.9000000 0.4651163 0.0002325581
12 [2.21e+03,2.41e+03)  1  9 0.11111111 0.8888889 0.2093023 0.0001162791
13 [2.41e+03,2.61e+03)  2  5 0.40000000 0.6000000 0.1162791 0.0002325581
> Intervals = Intervals[1:(length(Intervals)-1)]
> plot(ti,Emp_Surv_Fun,type = "l",
+  main = paste("Plot Of Empirical Survival
Function"),xlab="ti",ylab="S(x)",ylim =
```

```
+ c(0,1),col="black",lwd=2)
> lines(Intervals,Pi,type = "l",
+  main = paste("Plot Of Pi"),xlab="Ages",ylab="Pi",col="green",lwd=2)
> #Drawing The Legend
> legend("topright",legend=c(expression(bold("Empirical Survival Function")),
+  expression(bold("Survival function of grouped
+ data"))),fill=c("black","green"))
> #Plot of force of mortality
> hist(ti,xlab="ti",lwd=2,probability = T,ylim=c(0,0.0015),col="green")
> lines(Intervals,f_ti_star,type = "l",
+  main = paste("Plot Of
f_ti_star"),xlab="ti",ylab="f_ti_star",col="black",lwd=2)
> legend("topright",legend=c(expression(bold("histogram of deaths")),
+  expression(bold("pdf (curve of deaths)"))),fill=c("green","black"))
```

Graph:-




Histogram of ti
Density
0.0000
0.0005
0.0010
0.0015
0
500
1000
1500
2000
2500
3000
ti

Problem 23

Find the IBFS by using:

i.) North-west corner Rule ii.) Row minimum iii.) Column minimum iv.) Matrix Minimum v.) Vogel's approximation methods:

$W \rightarrow$ $F \downarrow$	W_1	W_2	W_3	W_4	**Factory Capacity**
F_1	19	30	50	10	7
F_2	70	30	40	60	9
F_3	40	8	70	20	18
Warehouse Requirement	5	8	7	14	34

Objective:-

To find the Initial Basic Feasible solution via 5 different methods namely North-West corner Rule, Row Minimum, Column Minimum, Matrix Minimum(Least-Cost) and Vogel's approximation.

Theory:-

1. **North-West Corner Method**:- The Northwest Corner Method is a basic algorithm used in linear programming to find an initial feasible solution for a transportation problem. It is a simplified approach that starts allocating units from the top-left (northwest) corner of the transportation matrix and moves horizontally and vertically.
2. **Row minimum:-** In the row minima method, the first row that is the lowest cost cell is exhausted and we allocate the maximum either at the first source or demand at the destinations or to satisfy both. This process must be continued for all the other reduced transportation costs until and unless the supply and demand are satisfied.
3. **Column minimum:-** In the column minima method, we begin with the first column and allocate gradually moving towards the lowest cost cell of the column. This system is continued until the first destination center is satisfied or the capacity of the second is exhausted, or both happens.
4. **Matrix Minimum:-** The Matrix Minimum or Least Cost Method is used to obtain the initial feasible solution for the transportation problem. Here, the allocation begins with the cell which has the minimum cost. The lower cost cells are chosen over the higher-cost cell with the objective to have the least cost of transportation.
5. **Vogel's approximation:-** Vogel's Approximation Method (VAM) is one of the methods used to calculate the initial basic feasible solution to a transportation problem. However, VAM is an iterative procedure such that in each step, we should find the penalties for each available row and column by taking the least cost and second least cost.

Solution:-

1. North-West Corner Method:

From \ To	D_1	D_2	D_3	D_4	Supply
S_1	19 5	30 2	50	10	7
S_2	70	30 6	40 3	60	9
S_3	40	8	70 4	20 14	18
Demand	5	8	7	14	34

Total cost=$5 \times 19 + 2 \times 30 + 6 \times 30 + 3 \times 40 + 4 \times 70 + 14 \times 20 = 1015$

2. Row Minima Method:

From \ To	D_1	D_2	D_3	D_4	Supply
S_1	19	30	50	10 7	7
S_2	70	30 8	40 1	60	9
S_3	40 5	8	70 6	20 7	18
Demand	5	8	7	14	34

Total Cost=$10 \times 7 + 30 \times 8 + 40 \times 1 + 40 \times 5 + 70 \times 6 + 20 \times 7 = 1110$.

3. Column Minima

From \ To	D_1	D_2	D_3	D_4	Supply
S_1	19 5	30	50	10 2	7
S_2	70	30	40 7	60 2	9
S_3	40	8 8	70	20 10	18
Demand	5	8	7	14	34

Total cost=$19 \times 5 + 10 \times 2 + 40 \times 7 + 60 \times 2 + 8 \times 8 + 20 \times 10 = 779$

4. Least Cost Method

From \ To	D_1	D_2	D_3	D_4	Supply
S_1	19	30	50	10 7	7
S_2	70 2	30	40 7	60	9
S_3	40 3	8 8	70	20 7	18
Demand	5	8	7	14	34

Total cost=$10 \times 7 + 70 \times 2 + 40 \times 7 + 40 \times 3 + 8 \times 8 + 20 \times 7 = 814$

5. Vogel's Approximation method:

	D_1	D_2	D_3	D_4	Supply	Penalty
S_1	19	30	50	10	7	9
S_2	70	30	40	60	9	10
S_3	40	8 8	70	20	18	12
Demand	5	8	7	14	34	
Penalty	21	22	10	10		

	D_1	D_3	D_4	Supply	Penalty
S_1	5 19	50	10	7	9
S_2	70	40	60	9	20
S_3	40	70	20	10	20
Demand	5	7	14		
Penalty	21	10	10		

	D_3	D_4	Supply	Penalty
S_1	50	10	2	40
S_2	40	60	9	20
S_3	70	10 20	10	50
Demand	7	14		
Penalty	10	10		

	D_3	D_4	Supply	Penalty
S_1	50	2 10	2	40
S_2	40	60	9	20
Demand	7	4		
Penalty	10	50		

	D_3	D_4	Supply	Penalty
S_2	40	2 60	9	20
Demand	7	2		
Penalty	40	60		

	D_3	Supply	Penalty
S_2	7 40	7	40
Demand	7		
Penalty	40		

	D_1	D_2	D_3	D_4	Supply
S_1	5 19	30	50	2 10	7
S_2	70	30	7 40	2 60	9
S_3	40	8 8	70	10 20	18
Demand	5	8	7	14	34

Total Cost$= 19 \times 5 + 10 \times 2 + 40 \times 7 + 60 \times 2 + 8 \times 8 + 20 \times 10 = 779$

Conclusion:-

|From the above experiment we find that IBFS for;

1. North-West Corner method is 1015
2. Row minima method is 1110
3. Column minima method is 779
4. Least Cost method is 814
5. Vogel's Approximation method is 779

Problem 24

A department head has four subordinates and four tasks have to be performed. Subordinates differ in efficiency and tasks differ in their intrinsic difficulty. Time each man would take to perform each task is given in the effectiveness matrix. How should the tasks be allocated to each person so as to minimize the total man-hours?

	I	II	III	IV
A	8	26	17	11
B	13	32	4	26
C	38	19	22	15
D	19	26	24	14

Objective:- To allocate 4 Tasks to 4 subordinate in such a way that the effectiveness of the above matrix increases as by minimizing the total man-hours(In other word, Assignment Problem).

Theory:- An assignment problem is a particular case of transportation problem where the objective is to assign a number of resources to an equal number of activities so as to minimize total cost or maximize total profit of allocation. Here, each subordinate will be assigned to a particular task in such a way that there will minimum total man hours.

Solution:-

	I	II	III	IV
A	8	26	17	11
B	13	32	4	26
C	38	19	22	15
D	19	26	24	14

After Row Reduction;

	I	II	III	IV
A	0	18	9	3
B	9	28	0	22
C	23	4	7	0
D	5	12	10	0

After Column Reduction;

	I	II	III	IV
A	0	14	9	3
B	9	24	0	22
C	23	0	7	0
D	5	8	10	0

After Allocating points

	I	II	III	IV
A	0	14	9	3
B	9	24	0	22
C	23	0	7	0
D	5	8	10	0

Conclusion:- From the above Experiment we find the optimum allocation so that it minimizes time hour

$$A \rightarrow I$$
$$B \rightarrow III$$
$$C \rightarrow II$$
$$D \rightarrow IV$$