Experiment No: 01

Experiment Name: Find out the point estimate of the population mean and interval estimation of the population mean. where 30 students quiz test marks is

(2, 4, 3, 23, 25, 27, 28, 13, 15, 16, 20, 14, 35, 33, 32, 21, 35, 40, 42, 22, 33, 13, 17, 20, 25, 29, 27, 40, 38, 31).

Total marks 50. Here population size $N = 30$ and sample size $n = 10$.

Also illustrate the sample size determination, sampling distribution for mean and check the unbiasedness of the population mean.

Objectives:

1. To calculate the point estimation and interval estimation.
2. To calculate sampling distribution for mean.

3. To check the unbiasedness of the population mean.

4. To comment on the data.

Procedure:

Step-1: First of all we find out the population mean and population variance. Population length is N.

$$\text{mean, } \bar{x} = \frac{\Sigma x_i}{N}$$

$$\text{variance, } s^2 = \frac{1}{N-1}\left[\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{N}\right]$$

Step-2: To calculate point estimation and interval estimation.

interval estimation:

$$\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \;,\; \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

step-3: Sampling Distribution for mean.

We choose the sample size $n=10$ from the population size $N=30$

Then we calculate the mean and unbiasedness.

bias = mean(nsample) - mean(population)

= 0

when bias is 0 then we can say the mean is unbiasedness.

step-4: Sampling Distribution for median. We choose the sample size $n=10$ from the population size $N=30$

The we calculate the median and unbiasedness.

bias = median(nsample) - median(population)

when bias is 0 then we can say the median is unbiasedness.

step-5: Efficiency check

We calculate the mean and the median of sampling distribution.

Mean and median to be two unbiased estimator then which variance is ~~more~~ less than other then we say that this is more efficient than other.

R-Source code:

```
IQ<-c(2,4,3,23,25,27,28,13,15,16,20,14,
35,33,32,21,35,40,42,22,33,13,17,20,25,
29,27,40,38,31)
mean(IQ)
var(IQ)
length(IQ)

set.seed(1246)
x<-sample(IQ,10,replace=TRUE)
mean(x)
sd(IQ)
qnorm(0.025,0.1)
```

```
## lower class interval
21.6 -((1.96*11)/sqrt(10))
## upper class interval
21.6 + ((1.96*11)/sqrt(10))

##Sampling Distribution for mean
choose(30,10)
nsample<-req(0,300000)
for(i in 1:300000){
    nsample[i]<-(mean(sample(IQ,10,
                    replace=TRUE)))

}
mean(nsample)
bias = mean(nsample)-mean(IQ)

##Sampling Distribution for median
choose(30,10)
nsample2<-req(0,300000)
for(i in 1:300000){
    nsample[i]<-(median(sample(IQ,10,
                    replace=TRUE)))

}
```

```
median(IQ)
median(nsample2)
bias = median(nsample2) - median(IQ)

### Efficiency check ###
L1<-length(nsample)
V1<-sum((nsample-mean(IQ))^2)/L1
V1
L2<-length(nsample2)
V2<-sum((nsample2-median(IQ))^2)/L2
V2
```

Input and output :

```
mean(IQ) = 24.1
var(IQ) =121.2655
length(IQ) = 30
mean(x) = 21.6
sd(IQ) = 11.012
qnorm = -1.96
14.78  # lower class interval
28.41  # upper class interval
```

mean(nsample) = 24.097

bias = -0.0024

median(IQ) = 25

median(nsample2) = 25

bias = 0

L1 = 300000

V1 = 11.69

L2 = 300000

V2 = 19.97

Comment : From the R code we can see that the mean is a unbiased estimator and the median also unbiased estimator. The variance of ~~mean is~~ nsample is less then the variance of nsample2. So, the mean is more efficient ~~the~~ than median.

Experiment No: 02

Experiment Name: Two dice rolled, S is the sum of both faces, Find the expectation of S, E(S) and variance of S, V(S). Plot the distribution of s and dice D.

Objectives:

1. To find the expectation of S.
2. To find the variance of S.
3. To plot the distribution of S and dice D.
4. To comment on the data.

Procedure:

step-1: Two dice rolled, S is the sum of both faces. To calculate the expectation of S, E(S).

Step-2: To calculate the variance of S, V(S).

$$V(S) = \frac{\left[\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}\right]}{n-1}$$

step-3: To plot the distribution of S and dice D.

R- Source code:

```
S<- 2:12
A <-c(1:6,5:1)
 PS<-C(1:6,5:1)/36
ES<-Sum(S*PS)
VarS<-Sum((S-C(ES))^2*PS)

## plot distribution of S
barplot(PS, ylim=C(0,0.2),
        ylab="Probability",
        xlab="S",
        col="steelblue",
        space=0,
        main="Sum of two dice rolls")
```

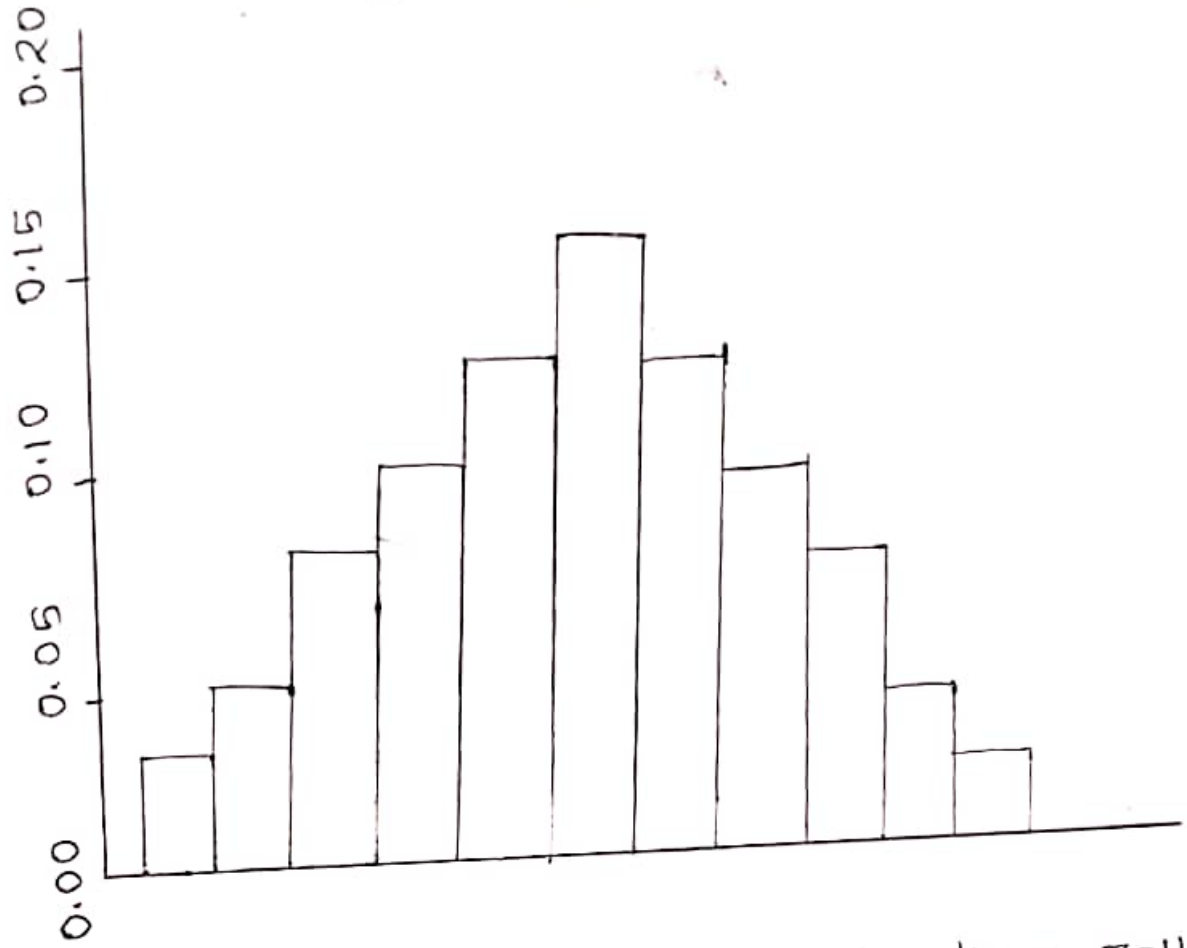
```
### plot distribution of D
Probability <- rep(1/6,6)
names(Probability) <-1:6

barplot(probability,
        ylim=c(0,0.2),
        xlab="D"
        col="steelblue",
        space=0,
        main="outcomes of a single
              dice roll")
```

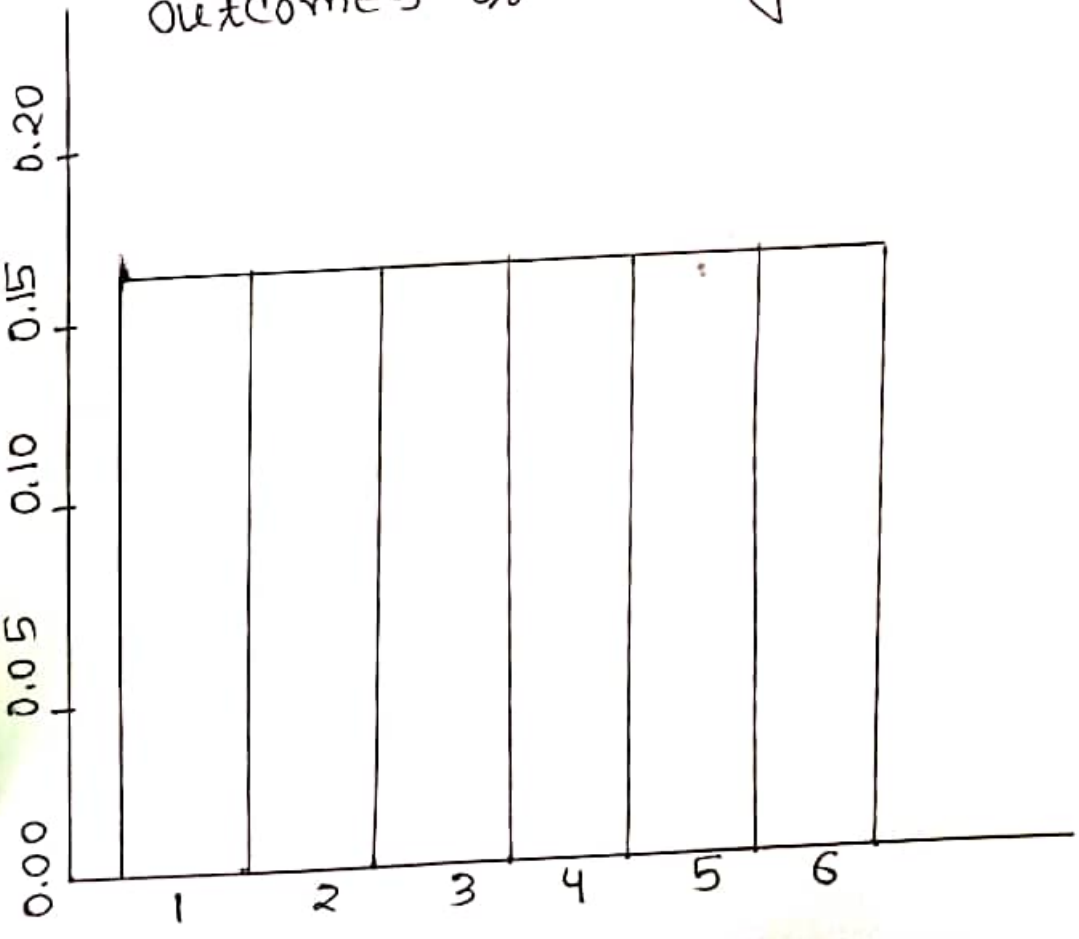
Input and output:

```
Es = 7
vars = 5.833
```

Sum of two dice rolls



outcomes of a single dice roll



Comment : Two dic rolled, S is the sum of both faces, the expectation of S, $E(S) = 7$ and variance of S, $V(S) = 5.833$

Experiment No: 03

Experiment Name: A herd of 1500 steer was fed to special high protein gain for a month. A random sample of 29 was weighted and had gained an average of 6.7 pounds. If the sd of weight gain for the entire herd is 2.1. Test the hypothesis at 5% level of significance that the average weight gain per steer for the month was more than 5 pounds. Also comments on the test using the p-value. Create the confidence interval.

Objectives:

1. To construct the average weight gain per steer for the month. ~~was~~
2. To create the confidence interval.
3. To construct p-value.
4. To comment on the data.

Procedure :

Step-1: state the null hypothesis and alternative hypothesis.

The alternate hypothesis was more than 5 pounds.

So, H1 : mu > 5

H0 : mu <= 5

Step-2: Select the level of significance.

The significance level ~~is~~ selected 0.05.

Step-3: select the test statistic.

Hence, the standard deviation are known, so we use z-test.

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$$\bar{x} = \frac{1}{n} \Sigma x_i$$

Step-4: Formulate the decision rule.

If the tabulated value of z is greater than calculated value of z then null

hypothesis is accepted, otherwise null hypothesis is rejected.

R Source code :

```
H0: mu <= 5
H1 : mu > 5

x.bar <- 6.7
mu <- 5
Sd <- 7.1
n <- 29
Z <- (x.bar - mu) / (Sd / sqrt(n))
alpha = 0.05
ztab <- qnorm(0.05, lower.tail = FALSE)
# p-value
pvalue <- pnorm(Z, lower.tail = FALSE)
H0: mu is not equal 5
ztab1 <- qnorm (0.025 )
ztab2 <- qnorm (0.975)
p.value <- 2 * pnorm (Z, lower.tail = FALSE)
CI <- c (x.bar + ztab1 *sd/sqrt(n),
         x.bar + ztab2 *Sd/ sqrt(n) )
```

Input and output :

x.bar
6.7
mu
5
Sd
7.1
n
29
Z
1.28
alpha
0.05
Ztab = 1.64
Pvalue = 0.098
Ztab1 = -1.96
Ztab2 = 1.96
P.value = 0.19
CI = 4.1159 , 9.2840

Comment : From the R code we can see that, Ztab > Zcal and also see that Pvalue > alpha . so H0 is accepted. The average weight gain per steer for the month was less than 5 pounds.

Experiment No : 04

Experiment Name: In order to find out whether children with chronic diarrhea have the same average hemoglobin level (Hb) that is normally seen in healthy children in the same area, a random sample of 10 children with chronic diarrhea are selected and their Hb levels (g/dl) are obtained as follows : 12.3, 11.4, 14.2, 15.3, 14.8, 13.8, 11.1, 15.1, 15.8, 13.2.

Do the data provide sufficient evidence to indicate that the mean Hb level for children with chronic diarrhea is less than the normal value of 14.6 (g/dl)? Test at 0.01 level of significance. Draw a boxplot and normal plot for this data and comments.

Objectives:

1. To calculate the mean Hb level for children with chronic diarrhea is less than the normal value of 14.6 (g/dl).

2. To draw a boxplot and normal plot for this data.

3. To comments for this data.

Procedure:

Step-1: State the null hypothesis and alternate hypothesis. The alternate hypothesis is less than the normal value of 14.6. So the null hypothesis is greater than or equal 14.6.

$H0: \mu >= 14.6$

$H1: \mu < 14.6$

Step-2: Select the level of Significance. The Significance level is selected 0.01.

Step-3: Select the test statistic.

The sample size is less 30 and population standard deviation are unknown. so we use t-test.

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

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

Step-4: Formulate the decision rule.

If the tabulated value of t is greater than calculated value of t then H0 is accepted, otherwise rejected.

R Source code:

```
H0: mu >= 14.6
H1: mu < 14.6
mu <- 14.6
data <- c(12.3, 11.4, 14.2, 15.3, 14.8, 13.8, 11.1,
15.1, 15.8, 13.2)
n <- length(data)
```

```
x.bar <- mean(data)
Sd.est <- sd(data)
t <- (x.bar - mu)/(Sd.est/sqrt(n))
tab <- qt(0.01, n-1)
Pvalue.t <- pt(t, df = n-1)
boxplot(data, ylab = "Hb lebel", col = "gray")
qqnorm(data, main = "Normal Q-Q plot
of Hb level")
qqline(data)
```

Input and output :

```
mu = 14.6
n = 10
x.bar = 13.7
Sd.est = 1.655
t = -1.71
tab = -2.82
Pvalue.t = 0.059
```

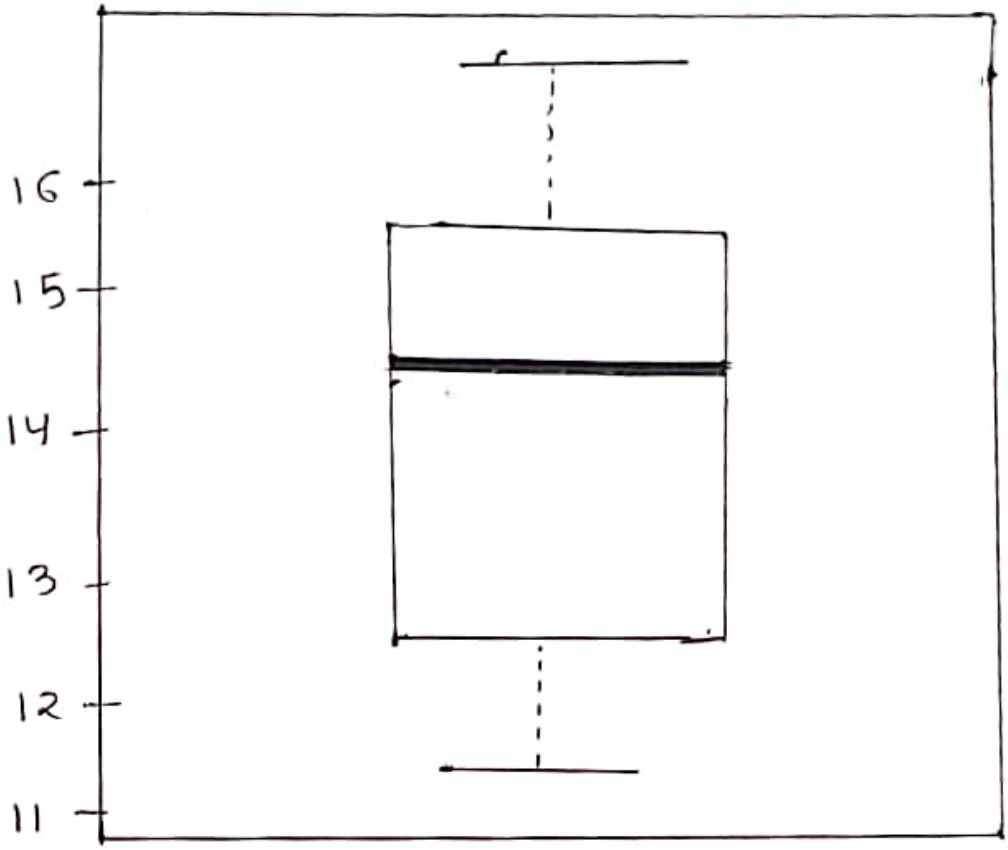


Fig: Boxplot

Normal Q-Q plot of Hb level

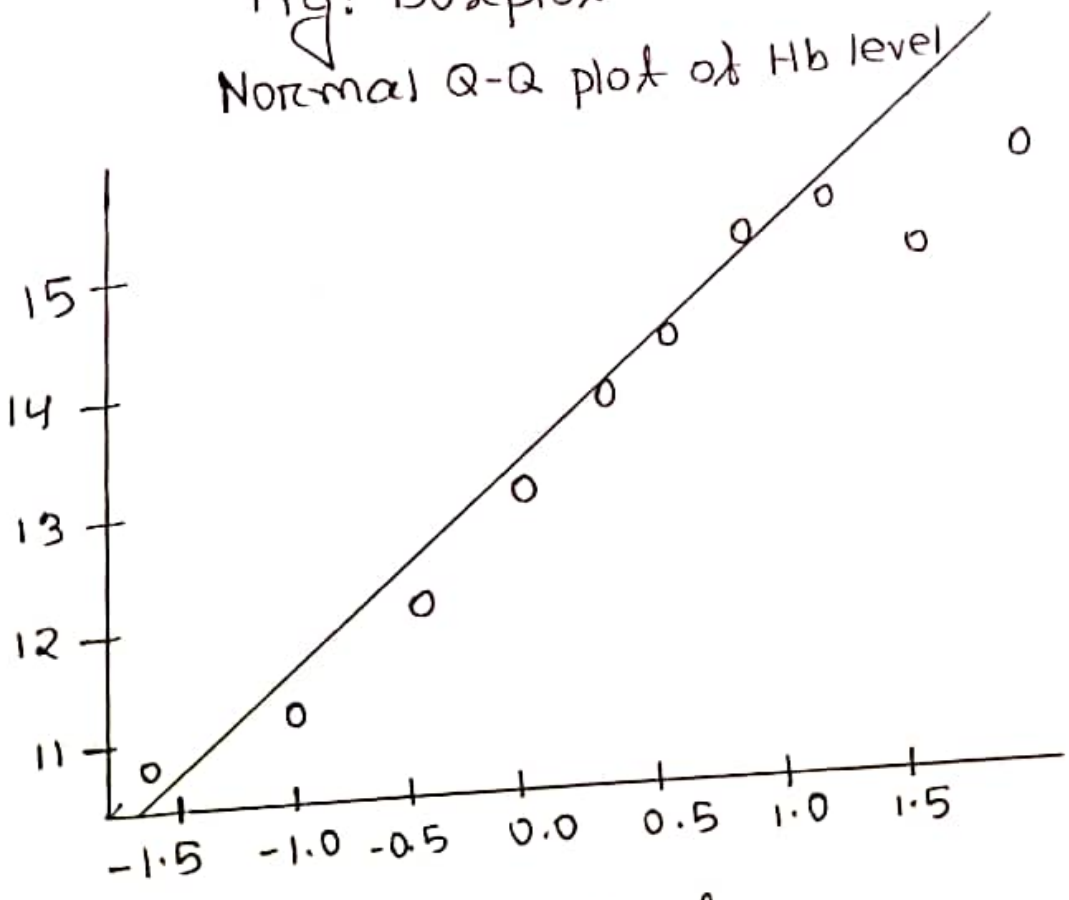


Fig: Normal plot

Comment: From the R code we can see that $t_{tab} > t_{cal}$ and also pvalue > alpha so H0 is accepted. Then the mean Hb level for children with chronic diarrhea is more than the normal value of 14.6.

Experiment No: 05

Experiment Name: In order to find out wheather children with chronic diarrhea have the same average hemoglobin level (Hb) that is normally seen in healthy children in the same area a random sample of 10 children with chronic diarrhea are selected and their Hb level (g/dl) are obtained as follows: 12.3, 11.4, 14.2, 15.3, 14.8, 13.8, 11.1, 15.1, 15.8, 13.2 another random sample of 12 children with chronic diarrhea are 11.1, 17.2, 13.4, 15.2, 14.1, 13.0, 12.5, 11.5, 12.7, 14.5, 15.3, 14.0. Is there any difference in the mean Hb label between the two groups of children

Objectives :

1. To calculate any difference in the mean Hb label between the two groups of children.

2. To comment on the data.

Procedure :

Step-1 : State that the null hypothesis the mean Hb lebel between the two group of children is equal and the alternate hypothesis state that the mean Hb lebel between two group of children is not equal.

$$H0: \mu_1 = \mu_2$$

$$H_A: \mu_1 \neq \mu_2$$

Step-2 : Select the level of significance. The selected level of significance is 0.05.

step-3: Select the test statistics.

The sample size is less than 30 and population standard deviation is unknown also the variance is ~~not~~ equal of the two groups of data. so we use t-test.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

$$\bar{x}_1 = \frac{1}{n_1}\sum x_{i1}$$

$$\bar{x}_2 = \frac{1}{n_2}\sum x_{i2}$$

$$s_1^2 = \frac{1}{n_1-1}\left[\sum x_{i1}^2 - \frac{(\sum x_{i1})^2}{n_1}\right]$$

$$s_2^2 = \frac{1}{n_2-1}\left[\sum x_{i2}^2 - \frac{(\sum x_{i2})^2}{n_2}\right]$$

$$s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$$

step-4: Formulate the decision rule. The tabulated value of t is greater than the calculated value than the null hypothesis is accepted, otherwise rejected.

R-source code:

```
x1<-c(12.3, 11.4, 14.2, 15.3, 14.8, 13.8, 11.1, 15.1,
15.8, 13.2)
x2<-c(11.1, 17.2, 13.4, 15.2, 14.1, 13.0, 12.5, 11.5,
12.7, 14.5, 15.3, 14.0)

s1<-sd(x1)
s2<-sd(x2)
s1/s2

n1<-length(x1)
n2<-length(x2)
x1.bar<-mean(x1)
x2.bar<-mean(x2)
x1.var<-var(x1)
x1.var<-var(x2)
sp<-(((n1-1)*x1.var + (n2-1)*x2.var)/
                          (n1+n2-2))
.
t<-(x1.bar - x2.bar)/sqrt(sp*(1/n1 + 1/n2))
```

```
alpha = 0.05
t.tab <- qt(alpha/2, n1+n2-2)
```

Input and output:

```
S1 = 1.65
S2 = 1.72
S1/S2 = 0.96
n1 = 10
n2 = 12
x1.bar = 2.74  13.7
x2.bar = 2.96  13.708
x1.var = 2.74
x2.var = 2.96
Sp = 2.86
t = -0.011

alpha = 0.05
t.tab = -2.085
```

Comment: From the R code we can see that t.tab > t cal . so H0 is accepted. So we can say There is no difference in the mean Hb label between the two groups of children

Experiment No: 06

Experiment Name: Test the hypothesis that the mean systolic blood pressure of healthy subjects (status-0) and subject with hypertension (status-1) are equal, have $d_0 = 0$. The dataset contains $n_1 = 25$ subjects with status-0 and $n_2 = 30$ with status-1.

Status-0: (120, 115, 94, 118, 111, 102, 102, 131, 104, 107, 115, 139, 115, 113, 114, 105, 115, 134, 109, 109, 93, 118, 109, 106, 125)

Status-1: (150, 142, 119, 127, 141, 149, 144, 142, 149, 161, 143, 140, 148, 149, 141, 146, 159, 152, 135, 94, 161, 130, 125, 141, 148, 153, 145, 137, 147, 169)

Objectives:

1. To calculate ~~any difference~~ in the mean systolic blood pressure of healthy subjects (status-0) and subject with hypertension (status-1) are equal.

2. To comment on the data.

Procedure :

Step-1: The null hypothesis state that the mean systolic blood pressure of healthy subjects (status-0) and subject with hypertension (status-1) are equal and the alternate hypothesis state that the mean systolic blood pressure of healthy subjects (status-0) and subject with hypertension (status-1) are not equal.

$$H0: \mu_1 = \mu_2$$

$$H1: \mu_1 \neq \mu_2$$

Step-2: select the level of significance. The selected level of significance is 0.05.

Step-3 : select the test statistics.

The sample size is ~~than~~ less than or equal 30 and population standard deviation is known also the variance is equal of the two groups of data. So we use t-test.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$\bar{x}_1 = \frac{1}{n_1} \Sigma x_{i1}$$

$$\bar{x}_2 = \frac{1}{n_2} \Sigma x_{i2}$$

$$s_1^2 = \frac{1}{n_1 - 1} \left[\Sigma x_{i1}^2 - \frac{(\Sigma x_{i1})^2}{n_1} \right]$$

$$s_2^2 = \frac{1}{n_2 - 1} \left[\Sigma x_{i2}^2 - \frac{(\Sigma x_{i2})^2}{n_2} \right]$$

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Step-4: Formulate the decision rule. The tabulated value of t is greater than the calculated value than the null hypothesis is accepted, otherwise rejected.

R-Source code:

```
x1<-c(120,115,94,118,111,102,102,131,104,107,
115,139,115,113,114,105,115,134,109,109,93,
118,109,106,125)

x2<-c(150,142,119,127,141,149,144,142,149,
161,143,140,148,149,141,146,159,152,135,
134,161,130,125,141,148,153,145,137,147,
169)

s1<-sd(x1)
s2<-sd(x2)
s1/s2

n1<-length(x1)
n2<-length(x2)
x1.bar<-mean(x1)
x2.bar<-mean(x2)
x1.var<-var(x1)
x2.var<-var(x2)
```

```
sp<-(((n1-1)*x1.var+(n2-1)*x2.var)/
                        (n1+n2-2))
t<-(x1.bar - x2.bar)/sqrt(sp*(1/n1+
                              1/n2))
alpha=0.05
t.tab<-qt(alpha/2, n1+n2-2)
```

Input and output :

```
S1 = 11.15
S2 = 10.95
S1/S2 = 1.01
n1 = 25
n2 = 30
x1.bar = 112.92
x2.bar = 144.23
x1.var = 124.41
x2.var = 120.047
Sp = 122.0228
t = -10.46
t.tab = -2.005
```

Comment : From the R code we can see that, $t.tab < t cal$. H0 is rejected. So we can say that, the mean systolic blood pressure of healthy subjects (status-0) and subject with hypertension (status-1) are not equal.

Experiment No : 07

Experiment Name : The 126 people have some doing smoking and some do not smoke. Some of this type of data are tabulated is given below:

Disses / Smoking	Heart disses	Not heart disses	Total
Yes	55	16	71
NO	23	32	55
Total	78	48	N = 126

Is there any association between smoking and heart diseases for the given data?

Objectives:

1. To calculate there is any association between smaking and heart diseases for the given data?

2. To comment on the data.
3. To calculate p value.

Procedure:

Step-1: The null hypothesis state that there is no association between smoking and heart diseases and the alternate hypothesis state that there is any association between smoking and heart diseases.

Step-2: Select the level of significance. The selected level of significance is 0.05.

Step-3: Select the test statistics. These data follow the chi-square distribution.

Step-4: Formulate the decision rule.
If pvalue > alpha then the null hypothesis is accepted otherwise rejected.

R-source code:

```
M<-matrix(c(55,16,23,32), ncol=2,
                          byrow=TRUE)
chisq.test(M)
```

Input and output:

	1	2
1	55	16
2	23	32

X-squared = 15.222, df = 1, p-value = $9.56e^{-05}$

p-value = 0.0000956

Comment: From the R code we can see that, the pvalue is less than alpha. pvalue < alpha. So H0 is rejected. We can say that, there is any association between smoking and heart diseases.

Experiment No: 08

Experiment Name: There are two COVID-19 testing booths, we test some people and their recorded data is below, where the numbers of people of booth-1 are 11 and the numbers of people of booth-2 are 10.

Booth-1: positive, positive, negative, positive, negative, negative, positive, positive, positive, negative, positive.

Booth-2: Negative, negative, negative, positive, ~~negative~~ positive, negative, positive, negative, negative, negative.

Is there any relation between two booth?

Objectives:

1. To calculate the relation between two booth.
2. To calculate p-value.
3. To comment on the data.

Procedure :

step-1: Select the null hypothesis and alternate hypothesis. The null hypothesis state that there is no relation between two booth and alternate hypothesis state that there is relation between two booth.

$$H0: \mu_1 = \mu_2$$

$$H1: \mu_1 \neq \mu_2$$

step-2: Select the level of significance. The selected level of significance is 0.05.

step-3: Select the test statistics. These data follows the chi-square distribution.

step-4: Formulate the decision rule. If chisquare tabulation value is greater than chisquare calculated value then the null hypothesis

is accepted, otherwise the null hypothesis is rejected. Also p-value is greater than alpha then the null hypothesis is accepted.

R-source code:

H0: There is no relation between booth-1 and booth-2.

H1: There is relation between booth-1 and booth-2.

```
booth_1<-c("positive","positive","negative",
"positive","negative","negative","positive",
"positive","positive","negative","positive")

booth_2<-c("negative","negative",
"negative","positive","positive",
"negative","positive","negative",
"negative","negative")
x_table1<- table(booth_1)
x_table2<- table(booth_2)
```

```
m<-matrix(c(4,7,7,3), ncol=2,
byrow=TRUE, dimnames=list(c("Booth-1",
Booth-2"), c("negative", "positive")))
m
c1<-sum(m[1])
c1
c2<-sum(m[2])
c2
r1<-sum[m[1])
r1
r2<-sum(m[2])
r2
n<-sum(m)
n
E11<-(c1*r1)/n
E11
E21<-(c1*r2)/n
E21
E12<-(c2*r1)/n
E12
E22<-(c2*r2)/n
```

```
chi_yates <- (((abs(m[1] - E11) - 0.5)^2)/E11
+(((abs(m[2] - E21) - 0.5)^2)/E21) +
((abs(m[3] - E12) - 0.5)^2/E12 +
((abs(m[4] - E22) - 0.5)^2/E22)
chi_tab <- qchisq(0.05, df=1, lower.tail=
                                    FALSE)
chi_tab
## p-value
P_value <- pchisq(chi_yates, df=1,
                  lower.tail=FALSE)
P_value
```

Input and output :

```
booth_1
  negative  positive
      4         7

booth-2
   negative   positive
       7          3
```

	Negative	positive
Booth-1	4	7
Booth-2	7	3

C1 = 11

C2 = 10

r1 = 11

r2 = 10

n = 21

E11 = 5.76

E21 = 5.23

E12 = 5.23

E22 = 4.76

chi_yates = 1.21

chi_tab = 3.84

p_value = 0.2696

Comment: From R code we can see that, chi_yates is greater than chi_tab also p_value > alpha. So null hypothesis is accepted. So we can say that there is no relation between two booth.

Experiment No: 09

Experiment Name: The number of Systolic blood pressure of healthy subjects. The dataset contains $n = 25$.

120, 115, 94, 118, 111, 102, 102, 131, 104, 107, 115, 139, 115, 113, 114, 105, 115, 134, 109, 109, 93, 118, 109, 106, 125.

Do you think that the sample follows $N(\mu, 400)$.

Objectives:

1. To calculated the variance test.
2. To calculated null hypothesis.
3. To comment on the data.
4. To calculated P-value.

Procedure:

Step-1: Select the null hypothesis and alternate hypothesis.

$$H0: \sigma^2 = \sigma_0^2 = 400$$

$$H1: \sigma^2 \neq \sigma_0^2$$

Step-2: Select the level of significance. The selected level of significance is 0.05.

Step-3: Select the test statistics. It is a one valued variance and μ is ~~un~~known. So the test statistics is chi-square distribution.

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

Step-4: Formulate the decision rule. If p value is greater than alpha then null hypothesis is accepted otherwise, the null hypothesis is rejected.

R-Source code :

```
x<- c(120, 115, 94, 118, 117, 102, 102, 131, 104, 107,
115, 139, 115, 114, 113, 105, 115, 134, 109, 109, 93,
118, 109, 106, 125)

Sigma2 = 400
mu = 130
df = length(x)
chisquare = sum(x-mu)^2/ sigma2
p.value <- 2* min (pchisq (chisquare, df),
1-pchisq ( chisquare, df ))
```

Input and output :

~~mig~~

```
Sigma2 = 400
mu = 130
df = 25
Chisquare = 455.8225
P.value = 0.
```

Comment: From the R code we can see that P value is less than alpha. So the null hypothesis is rejected.

Experiment No : 10

Experiment Name: The systolic blood pressure of healthy subjects (status-0) and subject with hypertension (status-1) are equal, have do=0. The dataset contains n1 = 25 subjects with status-0 and n2 = 30 with status-1.

status-0: (120, 115, 94, 118, 111, 102, 102, 131, 104, 107, 115, 139, 115, 113, 114, 105, 115, 134, 109, 109, 93, 118, 109, 106, 125)

status-1: (150, 142, 119, 127, 141, 149, 144, 142, 149, 161, 143, 140, 148, 149, 141, 146, 159, 152, 135, 134, 161, 130, 125, 141, 148, 153, 145, 137, 147, 169)

Are the variations in systolic blood pressure of healthy subjects and subject with hypertension are same ?

Objectives :

1. To calculated the variations in systolic blood pressure of healthy subjects and subject with hypertension are same.
2. To calculated p value.
3. To comment on the data.

Procedure :

Step-1 : Select the null hypothesis and alternate hypothesis. The null hypothesis state that the variations in systolic blood pressure of healthy subjects and subject with hyper-tension are same and the alternate hypothesis state that

the variations in systolic blood pressure of healthy subjects and subject with hypertension are not same.

$$H0: \sigma_1^2 = \sigma_2^2$$

$$H1: \sigma_1^2 \neq \sigma_2^2$$

Step-2: Select the level of significance. The selected level of significance is 0.05.

Step-3: There are two valued variance. So the test statistics is F distribution.

$$s_1^2 = \frac{1}{n_1 - 1}\left[\Sigma x_{i1}^2 - \frac{(\Sigma x_{i1})^2}{n_1}\right]$$

$$s_2^2 = \frac{1}{n_2 - 1}\left[\Sigma x_{i2}^2 - \frac{(\Sigma x_{i2})^2}{n_2}\right]$$

$$F = \frac{s_1^2}{s_2^2} \quad s_1 > s_2$$

step-4: Formulate the decision ~~role~~ rule. When F.tab > Fcal then the null hypothesis accepted, otherwise null hypothesis rejected. Also pvalue is greater than alpha so H0 is accepted.

R-Source Code:

```
x1<-c(120,115,94,118,111,102,102,131,104,107,
115,114,113,105,115,134,109,109,93,118,109,
106,125)
x2<-c(150,142,119,127,141,149,144,142,149,
161,143,148,149,141,146,159,152,135,134,
161,130,125,141,148,153,145,137,147,169)

x1.var<-var(x1)
x2.var<-var(x2)
df1<-length(x1)-1
df2<-length(x2)-1
alpha=0.05
```

```
F.ratio <- x1.var / x2.var
F.tab <- qf(alpha, df1, df2, lower.tail = FALSE)
p.value <- 2 * min(pf(F.ratio, df1, df2), 1 - pf(F.ratio, df1, df2))
```

Input and output:

x1.var = 124.41

x2.var = 120.0471

df1 = 24

df2 = 29

F.ratio = 1.0363

~~F.tab =~~

alpha = 0.05

F.tab = 1.9005

P.value = 0.917

Comment : From the R code we can see that the tabulation value is greater than calculated value of F distribution.

F.tab > F.ratio. Also we can see that p.value is greater than alpha. P.value > alpho. So H0 is accepted. So we can say the variations in systolic blood pressure of healthy subjects and subject with hypertension are same.

Experiment No : ~~10~~ 11

Experiment Name : The sample observations are

X : 122, 145, 120, 45, 98, 67, 109, 100, 107, 106, 93, 125, 130, 90, 34, 108, 80, 48, 65, 56.

The test hypothesis at 5% level of significance that the test of median. Do you think that the median is 110 ?

Objectives :

1. To calculate the test of hypothesis of median.
2. To calculate p value.
3. To comment on the data.

Procedure :

Step-1 : Select the null hypothesis and alternate hypothesis.

H0 : median = 110

H1 : median ≠ 110

step-2: Select the level of significance.
The selected level of significance is 0.05.

step-3: Select the test statistics.
To calculate the median so it is non parametric test. The test statistics is sign test.

step-4: Formulate the decision rule.
If P value is greater than alpha then null hypothesis is accepted, otherwise null hypothesis is rejected.

R-source code:

```
X<-c(122, 145, 120, 45, 98, 67, 109, 100, 107,
106, 93, 125, 130, 90, 34, 108, 80, 48, 65, 56)
H0: median =110
md =110
y<-sum(X>md)
n<-sum(X!=md)
```

p.value <- 1 - pbinom (y-1, n, 0.5)
p.value = 0.99

Input and output :

y = 5
n = 20
p.value = 0.99

Comment : From R code we can see that p value is greater than alpha. so null hypothesis is accepted. so we can say that the median is 110.

Experiment No : 12

Experiment Name : Test the hypothesis that the median systolic blood pressure of healthy subjects (status-0) and subject with hypertension (status-1) are equal have d0 = 0. The dataset contains n1 = 25 subjects with status-0 and n2 = 30 with status-1.

status-0 : (120, 115, 94, 118, 111, 102, 102, 131, 104, 107, 115, 139, 115, 113, 114, 105, 115, 134, 109, 109, 93, 118, 109, 106, 125)

status-1 : (150, 142, 119, 127, 141, 149, 144, 142, 149, 161, 143, 140, 148, 149, 141, 146, 159, 152, 135, 134, 161, 130, 125, 141, 148, 153, 145, 137, 147, 169)

Is there any difference in the median between status-0 and status-1?

Objectives:

1. To calculate the difference in the median between status-0 and status-1.
2. To calculate p-value.
3. To comment on the data.

Procedure:

Step-1: Select the null hypothesis and alternate hypothesis. The null hypothesis state that there is no difference in the median between status-0 and status-1. The alternate hypothesis state that there is difference in the median between status-0 and status-1.

$H0: md1 = md2$

$H1: md1 \neq md2$

step-2: select the level of significance. The selected level of significance is 0.05.

step-3: select the test statistics. There are two valued non parametric so the test statistics is wilcoxon rank sum test.

Step-4: Formulate the decision rule. If P value is greater than alpha then the null hypothesis is accepted otherwise null hypothesis is rejected.

R-Source Code:

```
X1<-c(120, 115, 94, 118, 111, 102, 102, 131, 104, 107,
115, 139, 115, 114, 113, 105, 115, 134, 109, 109, 93,
118, 109, 106, 125)
X2<-c(150, 142, 119, 127, 141, 149, 144, 142, 149, 161,
143, 140, 148, 149, 141, 146, 159, 152, 135, 134, 161, 130,
125, 141, 148, 153, 145, 137, 147, 169)
```

```
wilcox.test(x1, x2, exact = FALSE,
correct = TRUE, alternative = "two.sided")
```

Input and output:

W = 18.

p-value = 1.649×10^{-9}

Comment: From the R code we can see that, p-value is less than alpha. p-value < alpha, so the null hypothesis is rejected. We can say that, There is difference in the median between status-0 and status-1.