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 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)

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.

$$HO: \mu_1 = \mu_2$$
$$H1: \mu_1 \neq \mu_2$$

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.

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

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

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

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

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

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

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, $t.tab < t cal$.

Ho 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

```
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 unknown. So the test statistics is chi-square distribution.

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

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

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 $d_0 = 0$. 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, 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 hypertension are same and the alternate hypothesis state that

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 HO 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.

$$H0: median = 110$$

$$H1: median \neq 110$$

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.