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), nol=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), nrow=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
r
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, 111, 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.