Assignement- 5

Name: Pranav s reddy

Reg no: 23BCE2075

1)sample test :-

1) An outbreak of salmonella-related illness was attributed to ice produced at a certain factory. Scientists measured the level of Salmonella in 9 randomly sampled batches ice crean.The levels(in MPN/g) were: Is there evidence that the mean level pf Salmonella in ice cream greater than 0.3 MPN/g

Code:-

```
> x=c(0.593,0.142,0.329,0.691,0.231,0.793,0.519,0.392,0.418)
> t.test(x,alternative ="greater",mu=0.3)

        One Sample t-test

data:  x
t = 2.2051, df = 8, p-value = 0.02927
alternative hypothesis: true mean is greater than 0.3
95 percent confidence interval:
 0.3245133       Inf
sample estimates:
mean of x
0.4564444

>
```

2) Suppose that 10 volunteers have taken an intelligence test; here are the results obtained. The average score of the entire population is 75 in the same test. Is there any significant difference (with a significance level of 95%) between the sample and population means, assuming that the variance of the population is not known

Scores: 65, 78, 88, 55, 48, 95, 66, 57, 79, 81

Code:-

```
> a=c(65,78,88,55,48,95,66,57,79,81)
> t.test(a,mu=75)

        One Sample t-test

data:  a
t = -0.78303, df = 9, p-value = 0.4537
alternative hypothesis: true mean is not equal to 75
95 percent confidence interval:
 60.22187 82.17813
sample estimates:
mean of x 
     71.2 

> |
```

Topic:- t-test for two samples

3) Comparing two independent sample means, taken from two populations with unknown variances. The following data shows the heights of individuals of two different countries with unknown population variances. Is there any significant difference between the average heights of two groups.

Code:-

```
x=c(175,168,168,190,156,181,182,175,174,179)
y=c(120,180,125,188,130,190,110,185,112,188)
t.test(x,y)
```

Output:-

```
        Welch Two Sample t-test

data:  x and y
t = 1.8827, df = 10.224, p-value = 0.08848
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -3.95955 47.95955
sample estimates:
mean of x mean of y
    174.8     152.8
```

4) A school athletics has taken a new instructor, and want to test the effectiveness of the new type of training proposed by the new instructor comparing the average times of 10 runners in the 100 meters. The results are given below(time in seconds)

In this case we have two sets of paired samples, since the measurements were made on the same athletes before and after the workout. To see if there was an improvement, deterioration, or if the means of times have remained substantially the same (hypothesis H0), we need to make a Student’s t-test for paired samples, proceeding in this way

Code:-

```
before =c(12.9,13.5,12.8,15.6,17.2,19.2,12.6,15.3,14.4,11.3)
after=c(12.7,13.6,12.0,15.2,16.8,20.0,12.0,15.9,16.0,11.1)
t.test(before,after,paired=TRUE)
```

Output:-

```
        Paired t-test

data:  before and after
t = -0.21331, df = 9, p-value = 0.8358
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.5802549  0.4802549
sample estimates:
mean of the differences
                  -0.05
```

5) Five Measurements of the output of two units have given the following results (in kilograms of material per one hour of operation) .Assume that both samples have been obtained from normal populations, test at 10% significance level if two populations have the same variance

Code:-

```
unit_A=c(14.1,10.1,14.7,13.7,14.0)
unit_B=c(14.0,14.5,13.7,12.7,14.1)
var.test(unit_A,unit_B)
```

Output:-

```
	F test to compare two variances

data:  unit_A and unit_B
F = 7.3304, num df = 4, denom df = 4, p-value = 0.07954
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
  0.7632268 70.4053799
sample estimates:
ratio of variances
         7.330435
```

CHI-SQUARE TEST

1) The below table gives the distribution of students according to the family type and the anxiety level.

Family type	**Anxiety level** [1]
	Low
Joint family	35
Nuclear family	48

Code:-

```
data<-matrix(c(35,42,61,48,51,68),ncol=3,byrow=T)
data

chisq.test(data)
```

Output:-

```
     [,1] [,2] [,3]
[1,]   35   42   61
[2,]   48   51   68

        Pearson's Chi-squared test

data:  data
X-squared = 0.53441, df = 2, p-value = 0.7655
```

2) A biologist is conducting a plant breeding experiment in which plants can have one of four phenotypes. If these phenotypes are caused by a simple Mendelian model, the phenotypes should occur in a 9:3:3:1 ratio. She raises 41 plants with the following phenotypes.

Phenotype	1	2	3	4
Count	20	10	7	4

Should she worry that the simple genetic model doesn't work for her phenotypes?

Code:-

```
piants<- c(20,10,7,4)
chisq.test(piants,p=c(9/16,3/16,3/16,1/16))
```

Output:-

```
        Chi-squared test for given probabilities

data:  piants
X-squared = 1.9702, df = 3, p-value = 0.5786

Warning message:
In chisq.test(piants, p = c(9/16, 3/16, 3/16, 1/16)) :
  Chi-squared approximation may be incorrect
```

3) A survey of 320 families with 5 children each revealed the following distribution:

Number of Boys	**5**	**4**	**3**	**2**	**1**	**0**
No of Girls	0	1	2	3	4	5
No of families	14	56	110	88	40	12

Code:-

```
x=c(5,4,3,2,1,0)
n=5
N=320
p<-0.5
obf<-c(14,56,110,88,40,12)
exf<-dbinom(x,n,p)*320
sum(obf)

sum(exf)

chisq<-sum((obf-exf)^2/exf)
chisq

qchisq(0.95,5)
```

Output:-

```
[1] 320
[1] 320
[1] 7.16
[1] 11.0705
```

4) Fit a Poisson distribution to the following data and test the goodness of fit

X 0 1 2 3 4 5 6

F 275 72 30 7 5 2 1

Code:-

```
x<-0:6
f<-c(275,72,30,7,5,2,1)
lambda<-(sum(f*x)/sum(f))
expf <-dpois(x,lambda)*sum(f)
f1=round(expf)
sum(f)

sum(f1)

obf<-c(275,72,30,15)
exf<-c(242,117,28,6)
chisq<-sum(((obf-exf)^2)/exf)
chisq

qchisq(0.95,2)
```

Output:-

```
[1] 392
[1] 393
[1] 35.45055
[1] 5.991465
```