MAT5007 – Applied Statistical Methods

Embedded Lab – R Statistical Software

FALL SEMESTER – 2022~2023 L25+L26 SLOT

E-RECORD

Assignment No.: 9

Submitted By

KAMRAN ANSARI

22MCA0223

MCA– I Year

SITE

DEPARTMENT OF COMPUTER APPLICATIONS

SCHOOL OF INFORMATION TECHNOLOGY AND ENGINEERING

VELLORE INSTITUTE OF TECHNOLOGY

VELLORE – 632 014

TAMIL NADU

INDIA

Date: 25/12/2022

Name: Kamran Ansari

Reg No: 22MCA0223

Experiment 1:

The following table gives the number of fatal road accidents that occurred during the 7 days of a week. Write down the R programming code to test whether the accidents are uniformly distributed over the week at 95 % level of confidence.

Day :	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Number :	8	14	16	12	11	14	9

```
> numberOfAccidents = c(8, 14, 16, 12, 11, 14, 9)

> uniformDistAccidents = c(rep(1/7, 7))

> chisq.test(numberOfAccidents, p = uniformDistAccidents)
```

```
> numberOfAccidents = c(8, 14, 16, 12, 11, 14, 9)
>
> uniformDistAccidents = c(rep(1/7, 7))
>
> chisq.test(numberOfAccidents, p = uniformDistAccidents)

        Chi-squared test for given probabilities

data:  numberOfAccidents
X-squared = 4.1667, df = 6, p-value = 0.6541
```

Interpretation: Here since the p-value > 0.05 we fail to reject the null hypothesis i.e. accidents are uniformly distributed over the week at 5% level of significance.

Name: Kamran Ansari
Reg No: 22MCA0223

Experiment 2:

A total number of 3759 individuals were interviewed according to gender and decision in a public opinion survey on a political proposal with the results as in the following table. Write down the R programming code to test the hypothesis that there is no association between gender and attitude 5 % level of significance.

	Decision		
	Favoured	**Opposed**	**Undecided**
Male	1154	475	243
Female	1103	442	342

```
> data = matrix(c(1154, 475, 243, 1103, 442, 342), ncol = 3, byrow = T)
```

```
> data
```

```
> chisq.test(data)
```

```
> data = matrix(c(1154, 475, 243, 1103, 442, 342), ncol = 3, byrow = T)
>
> data
     [,1] [,2] [,3]
[1,] 1154  475  243
[2,] 1103  442  342
>
> chisq.test(data)

        Pearson's Chi-squared test

data:  data
X-squared = 19.034, df = 2, p-value = 7.358e-05
```

Interpretation: Here since the p-value < 0.05 we reject the null hypothesis (there is no association between gender and attitude) and accept the alternative hypothesis that there is association between gender and attitude at 5% level of significance.

Name: Kamran Ansari
Reg No: 22MCA0223

Experiment 3:

A random sample is selected from each of 3 makes of ropes (Type 1, Type 2 and Type 3) and their breaking strength (in certain units) are measured with the results in the following table. Write down the R programming code to test whether the breaking strengths of the ropes differ significantly at 5 % level of significance.

Type 1 :	70	72	75	80	83		
Type 2 :	60	65	57	84	87	73	
Type 3 :	100	110	108	112	113	120	107

```
> type1BreakingStrength = c(70, 72, 75, 80, 83)
```

```
> type2BreakingStrength = c(60, 65, 57, 84, 87, 73)
```

```
> type3BreakingStrength = c(100, 110, 108, 112, 113, 120, 107)
```

```
> breakingStrengths = c(type1BreakingStrength, type2BreakingStrength,
type3BreakingStrength)
```

```
> breakingStrengths
```

```
> labels = c(rep("Type 1", length(type1BreakingStrength)), rep("Type 2",
length(type2BreakingStrength)), rep("Type 3", length(type3BreakingStrength)))
```

```
> labels
```

```
> crdanova = aov(breakingStrengths ~ labels)
```

```
> summary(crdanova)
```

Name: Kamran Ansari
Reg No: 22MCA0223

```
> type1BreakingStrength = c(70, 72, 75, 80, 83)
>
> type2BreakingStrength = c(60, 65, 57, 84, 87, 73)
>
> type3BreakingStrength = c(100, 110, 108, 112, 113, 120, 107)
>
> breakingStrengths = c(type1BreakingStrength, type2BreakingStrength, type3BreakingStrength)
>
> breakingStrengths
 [1]  70  72  75  80  83  60  65  57  84  87  73 100 110 108 112 113 120 107
>
> labels = c(rep("Type 1", length(type1BreakingStrength)), rep("Type 2", length(type2BreakingStrength)), rep("Type
 3", length(type3BreakingStrength)))
>
> labels
 [1] "Type 1" "Type 1" "Type 1" "Type 1" "Type 1" "Type 2" "Type 2" "Type 2" "Type 2" "Type 2" "Type 2" "Type 3" "T
ype 3"
[14] "Type 3" "Type 3" "Type 3" "Type 3" "Type 3"
>
> crdanova = aov(breakingStrengths ~ labels)
>
> summary(crdanova)
            Df Sum Sq Mean Sq F value   Pr(>F)
labels       2   5838  2919.2   38.89 1.16e-06 ***
Residuals   15   1126    75.1
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Interpretation: Here since the p-value < 0.05 we reject the null hypothesis (breaking strength of the ropes of different types are the same) and accept the alternative hypothesis that the breaking strength of the ropes of different types differ significantly at 5% level of significance.

Name: Kamran Ansari
Reg No: 22MCA0223