MAT2001 – Statistics for Engineers - ELA (R Code Studio), Winter Semester 2020-2021

LAB FAT EXAM

By: Jonathan Rufus Samuel (20BCT0332) Date: 21.6.2021

Q1) Write R code to compute mean, median, mode, variance, standard deviation and coefficient of variation for the following frequency distribution:

Expenditure	20-30	30-40	40-50	50-60	60-70	70-80	80-90
No. of families	3	61	132	153	140	51	2

Procedure & Result:

MAT2001 - LAB FAT EXAM

BY: JONATHAN RUFUS SAMUEL (20BCT0332) DATE: 21.6.2021

SET - A TIME: 10 am - 11:15 am

Q1) Procedure:

Calculate all the required values using standard formulas, then proceed to display the values.

Mean = $\frac{\Sigma fx}{\Sigma f}$; Median = $l + \frac{h}{(f.cf)} * \left(\frac{N}{2-c}\right)$

Mode = $l + \frac{(f0-f1)}{(2*f0-f1-f2)*h}$ Variance = $\frac{\Sigma(x-\bar{x})^2}{\Sigma f}$

Standard Deviation = sqrt(variance) = $\sqrt{\sigma^2} = \sigma$

Coefficient of Variation => CV = SD / Mean

Code:

```
# Mean
h = 10
x = seq(25, 85, h)
f = c(3.61, 132, 153, 140, 51, 2)
N = sum(f)
Mean = sum(x*f)/N
Mean

# Median
cf = cumsum(f)
mc = min(which(cf >= N/2))
mcf = f[mc]
c = cf[mc-1]
l = x[mc - h/2]
```

P.T.O

```
Median = l + (h/mcf) * (N/2 - c)
Median
# Mode
mo = which (f == max (f))
f0 = f[mo]
f1 = f [mo - 1]
f2 = f [mo + 1]
l = x [mo] - h/2
mode = l + ((f0 - f1) / (2*f0 - f1 - f2)) * h
mode
# Variance
M = c (x - Mean)
M
Var = ((sum (x - M)) ^ 2) / N
Var
# Standard Deviation
SD = sqrt (Var)
SD
# Coefficient of Variation
CV = SD / Mean
CV
```

∴ Result: Mean, Median, Mode, Variance, Standard Deviation and Coefficient of Variation of given data is found.

Output:

A: Code is as follows:

```
#mean
h = 10
x = seq(25,85,h)
f = c(3,61,132,153,140,51,2)
N = sum(f)
Mean = sum(x*f)/N
Mean
#median
cf = cumsum(f)
mc = min(which(cf>=N/2))
mcf=f[mc]
c=cf[mc-1]
l=x[mc]-h/2
Median=l+(h/mcf)*((N/2-c))
Median
#mode
mo=which(f==max(f))
f0=f[mo]
f1=f[mo-1]
f2=f[mo+1]
l=x[mo]-h/2
mode = l+((f0-f1)/(2*f0-f1-f2))*h
mode
#Variance
M = c(x-Mean)
M
Var = ((sum(x-M))^2)/N
Var
#Standard Deviation
SD = sqrt(Var)
SD
```

#Co-efficient of Variation

CV = SD/Mean

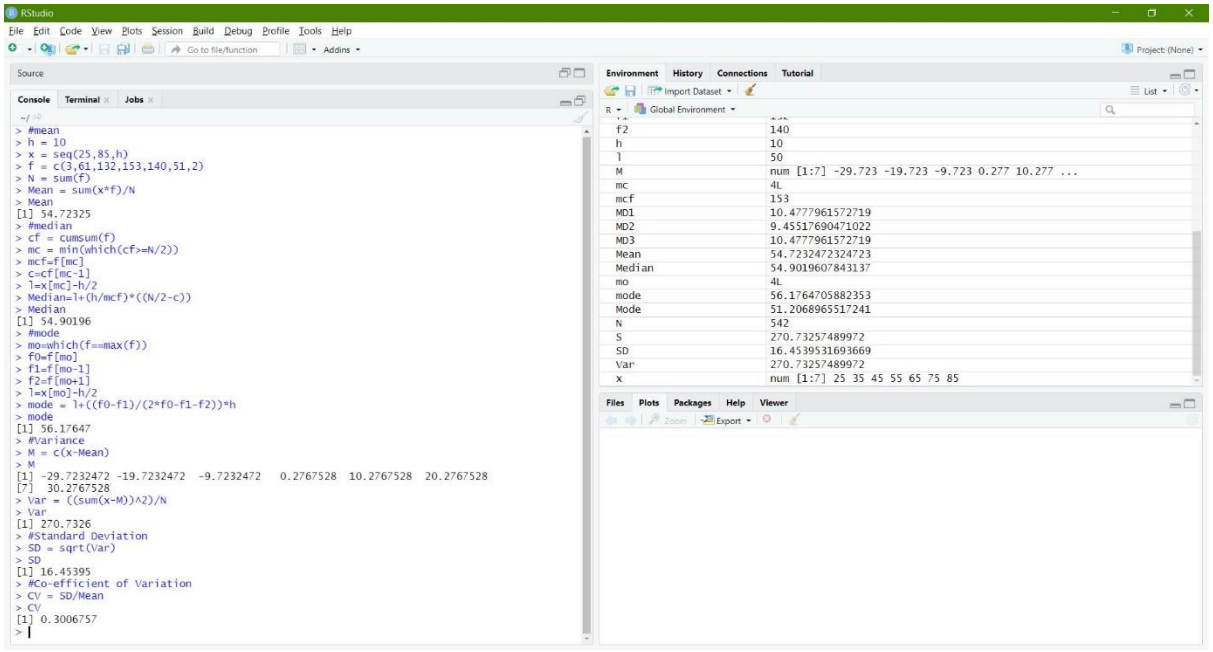
CV

Output (via Command Window):

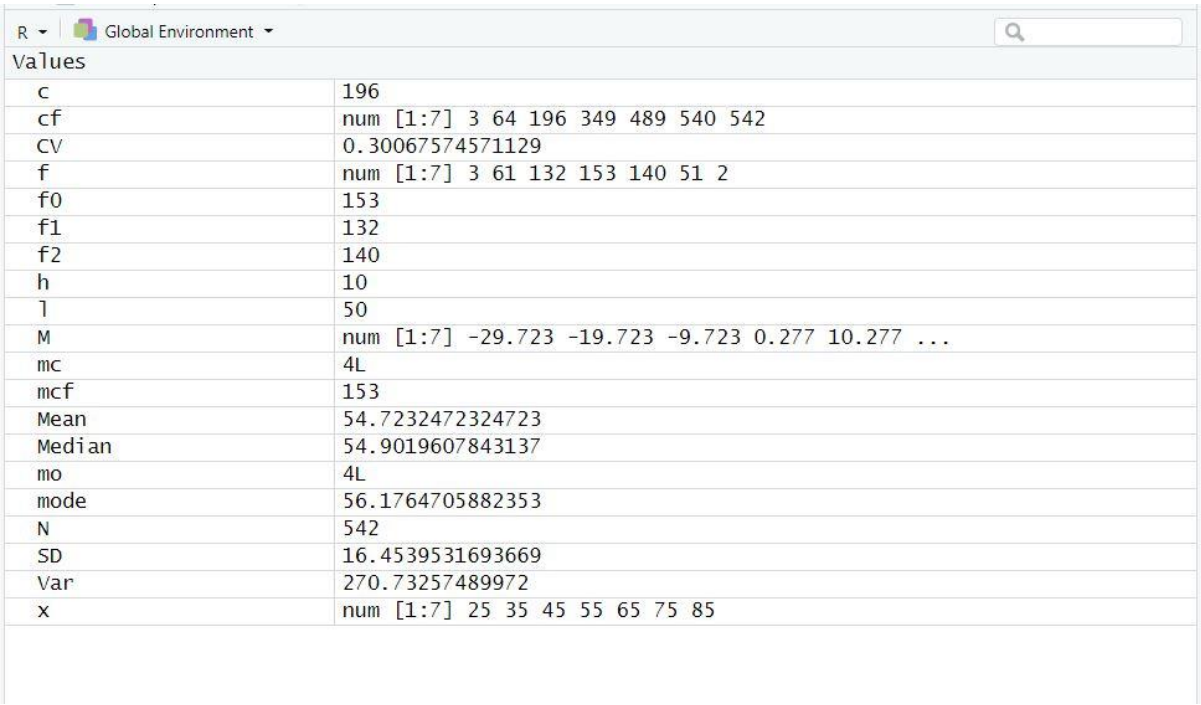
```
> #mean
> h = 10
> x = seq(25,85,h)
> f = c(3,61,132,153,140,51,2)
> N = sum(f)
> Mean = sum(x*f)/N
> Mean
[1] 54.72325
> #median
> cf = cumsum(f)
> mc = min(which(cf>=N/2))
> mcf=f[mc]
> c=cf[mc-1]
> l=x[mc]-h/2
> Median=l+(h/mcf)*((N/2-c))
> Median
[1] 54.90196
> #mode
> mo=which(f==max(f))
> f0=f[mo]
> f1=f[mo-1]
> f2=f[mo+1]
> l=x[mo]-h/2
> mode = l+((f0-f1)/(2*f0-f1-f2))*h
> mode
[1] 56.17647
> #Variance
> M = c(x-Mean)
> M
[1] -29.7232472 -19.7232472  -9.7232472   0.2767528  10.2767528  20.2767528
[7]  30.2767528
> Var = ((sum(x-M))^2)/N
> Var
[1] 270.7326
> #Standard Deviation
> SD = sqrt(Var)
> SD
[1] 16.45395
```

> #Co-efficient of Variation
> CV = SD/Mean
> CV
[1] 0.3006757

Implementation on R Studio Code (via Command Window):



Global Variables:



Q2) A completely randomized design experiment with 10 plots and 3 treatments (A,B and C) gave the following results:

A:	**5**	**7**	**3**	**1**
B:	**4**	**4**	**7**	
C:	**3**	**5**	**1**	

Write down the R programming code to analyse the results for treatment effects with 95% levels of confidence.

Procedure & Result:

Q2) Procedure

Store values of given data within A, B and C respectively.

i.e A = {5, 7, 3, 1}

B = {4, 4, 7}

C = {3, 5, 1}

As it is 95% Level of Confidence for above testing, note that alpha = 0.05 i.e 100-95% = 5% = 0.05

Get Summary of above data as shown in code below.

Let Hypothesis H0 = Treatment effects > 95% Level of Confidence

& H1 ≠ Treatment effects > 95% Level of Confidence

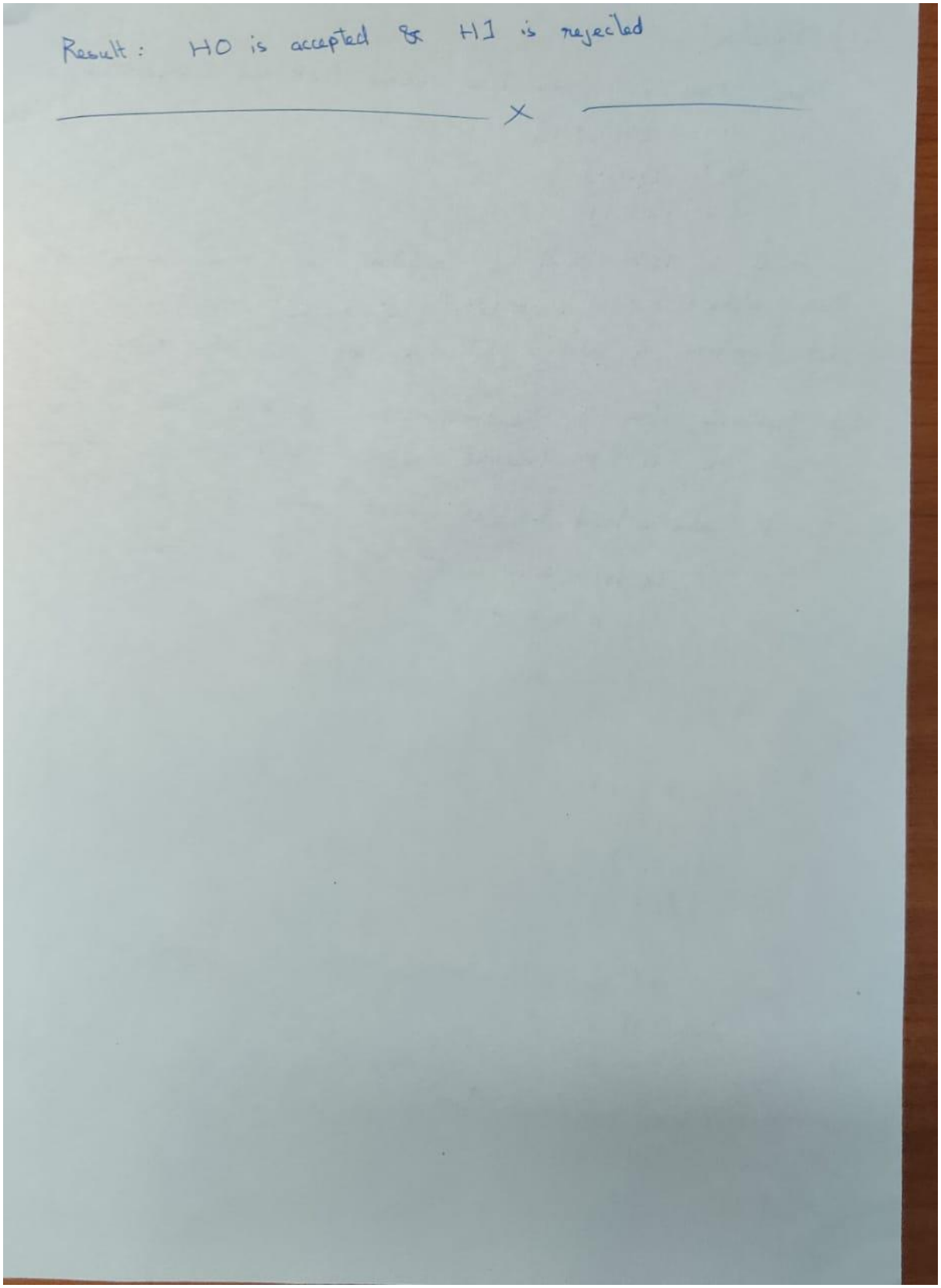
The Pr value is found to be 0.566 (under Pr(>F))

∴ As 0.566 is 95% Level of Confidence,

H0 is accepted & H1 is rejected

Code:

```
#Q2
A = c(5,7,3,1)
B = c(4,4,7)
C = c(3,5,1)
Data = c(A,B,C)
Data
Type = c(rep("Type1", length(A)), rep("Type2", length(B)), rep("Type3",
    length(C))
Type
ANOVA1 = aov(Data~Type)
Summary(ANOVA1)
# alpha = 0.05
```



Output:

A: Code is as follows:

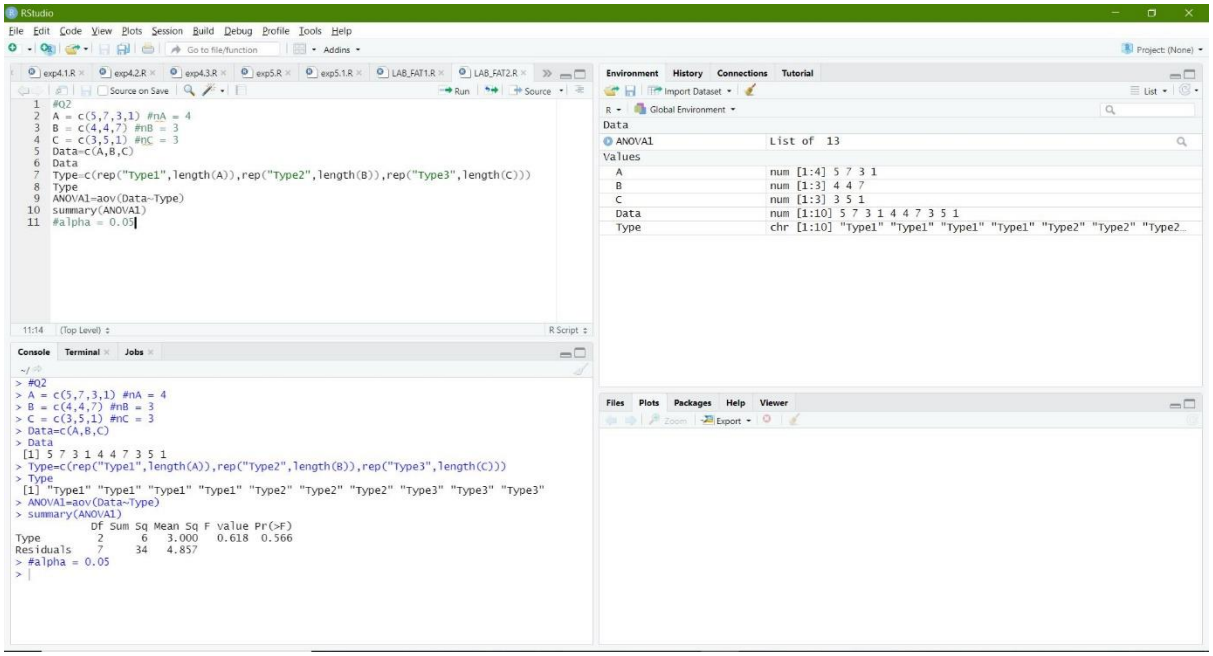
#Q2

```
A = c(5,7,3,1) #nA = 4
B = c(4,4,7) #nB = 3
C = c(3,5,1) #nC = 3
Data=c(A,B,C)
Data
Type=c(rep("Type1",length(A)),rep("Type2",length(B)),rep("Type3",length(C)))
Type
ANOVA1=aov(Data~Type)
summary(ANOVA1)
#alpha = 0.05
```

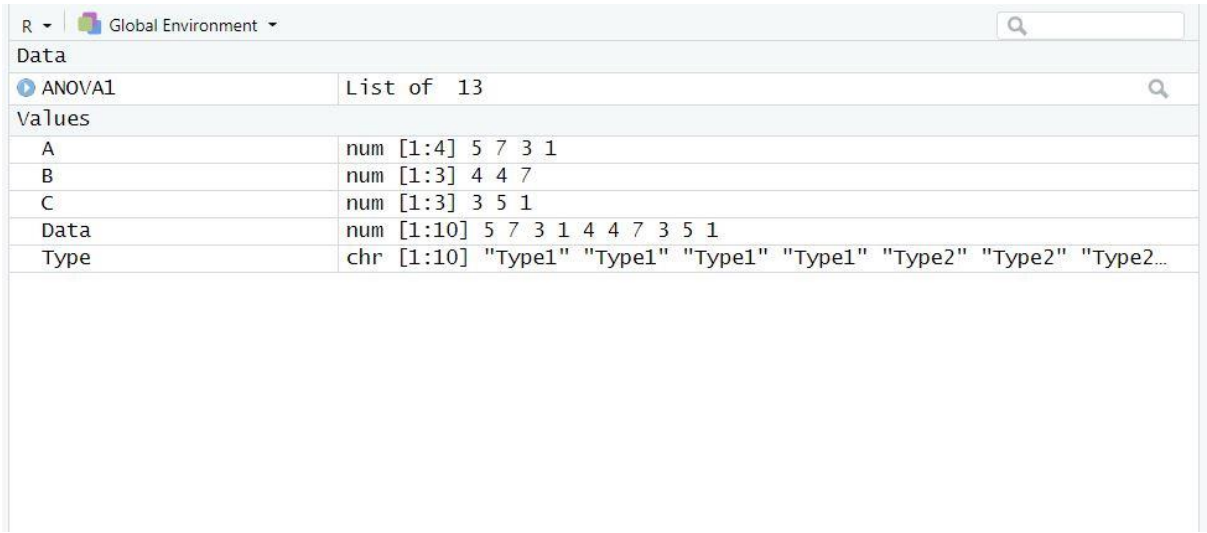
Output (via Command Window):

```
> #Q2
> A = c(5,7,3,1) #nA = 4
> B = c(4,4,7) #nB = 3
> C = c(3,5,1) #nC = 3
> Data=c(A,B,C)
> Data
 [1] 5 7 3 1 4 4 7 3 5 1
> Type=c(rep("Type1",length(A)),rep("Type2",length(B)),rep("Type3",length(C)))
> Type
 [1] "Type1" "Type1" "Type1" "Type1" "Type2" "Type2" "Type2" "Type3" "Type3" "Type3"
> ANOVA1=aov(Data~Type)
> summary(ANOVA1)
            Df Sum Sq Mean Sq F value Pr(>F)
Type         2      6   3.000   0.618  0.566
Residuals    7     34   4.857
> #alpha = 0.05
```

Implementation on R Studio Code (via Command Window):



Global Variables:



Q3) Viva-Voce/Quiz: ANSWERS:

1) t = M*(matrix(c(),nrow=3,ncol=2,byrow=TRUE)

2) Mean = sum(seq(25,82,1))/(82-25)

3) Option (b) - 2

4) Option (b) – not be extracted

5) Option (c) – storage.mode
