

**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:**

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

**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.

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

$$\text{Mode} = l + \frac{(f_0 - f_1)}{(2 * f_0 - f_1 - f_2)} * h$$

$$\text{Variance} = \frac{\sum (x - \bar{x})^2}{\sum f}$$

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

$$\text{Coefficient of Variation} \Rightarrow CV = SD / \text{Mean}$$

Code:

# Mean

n = 10

x = seq(25, 85, n)

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))

mc.f = f[mc]

c = cf[mc - 1]

l = x[mc - n/2]

P.T.O

```

Median =  $l + \left( \frac{h}{mcf} \right) * \left( \frac{N}{2} - c \right)$ 
Median

# Mode
mo = which (j == max (j))
d0 = j[mo]
d1 = j[mo - 1]
d2 = j[mo + 1]
l = x[mo] - h/2
mode =  $l + \left( \frac{(d0 - d1) / (2 * d0 - d1 - d2)}{h} \right) * h$ 
mode

# Variance
M = c (x - Mean)
M
Var =  $(\text{sum} (x - M))^2 / N$ 
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

```

#Coefficient 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):

The screenshot shows the RStudio interface. The console on the left displays the following R code and its output:

```
> #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 = 1+((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
> 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
> |
```

The environment pane on the right shows the following variables and their values:

Variable	Value
f2	140
h	10
l	50
M	num [1:7] -29.723 -19.723 -9.723 0.277 10.277 ...
mc	4L
mcF	153
MD1	10.4777961572719
MD2	9.45517690471022
MD3	10.4777961572719
Mean	54.7232472324723
Median	54.9019607843137
mo	4L
mode	56.1764705882353
N	542
S	270.73257489972
SD	16.4539531693669
Var	270.73257489972
x	num [1:7] 25 35 45 55 65 75 85

### Global Variables:

R - Global Environment	
Values	
c	196
cf	num [1:7] 3 64 196 349 489 540 542
CV	0.30067574571129
f	num [1:7] 3 61 132 153 140 51 2
f0	153
f1	132
f2	140
h	10
l	50
M	num [1:7] -29.723 -19.723 -9.723 0.277 10.277 ...
mc	4L
mcF	153
Mean	54.7232472324723
Median	54.9019607843137
mo	4L
mode	56.1764705882353
N	542
SD	16.4539531693669
Var	270.73257489972
x	num [1:7] 25 35 45 55 65 75 85



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

<b>A:</b>	<b>5</b>	<b>7</b>	<b>3</b>	<b>1</b>
<b>B:</b>	<b>4</b>	<b>4</b>	<b>7</b>	
<b>C:</b>	<b>3</b>	<b>5</b>	<b>1</b>	

**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  $H_0 = \text{Treatment effects} > 95\% \text{ level of confidence}$   
or  $H_1 \neq \text{Treatment effects} > 95\% \text{ level of confidence}$

The P-value is found to be 0.566 (under  $P_n(>F)$ )

$\therefore$  As 0.566 is 95% level of confidence,  
 $H_0$  is accepted &  $H_1$  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
```

Result:  $H_0$  is accepted &  $H_1$  is rejected

\_\_\_\_\_ X \_\_\_\_\_

**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):**

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code for creating vectors A, B, and C, combining them into a data frame, and performing an ANOVA.
- Console:** Displays the output of the R code, including the data frame structure and ANOVA results.
- Environment:** Shows the global environment with variables A, B, C, Data, and Type.

**R Code (Source Editor):**

```

1 #Q2
2 A = c(5,7,3,1) #nA = 4
3 B = c(4,4,7) #nB = 3
4 C = c(3,5,1) #nC = 3
5 Data=c(A,B,C)
6 Data
7 Type=c(rep("Type1",length(A)),rep("Type2",length(B)),rep("Type3",length(C)))
8 Type
9 ANOVA1=ov(Data~Type)
10 summary(ANOVA1)
11 #alpha = 0.05

```

**Console Output:**

```

> #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" "Type2" "Type2" "Type2" "Type3" "Type3" "Type3"
> ANOVA1=ov(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
>

```

**Environment Panel:**

Variable	Class	Values
A	num	[1:4] 5 7 3 1
B	num	[1:3] 4 4 7
C	num	[1:3] 3 5 1
Data	num	[1:10] 5 7 3 1 4 4 7 3 5 1
Type	chr	[1:10] "Type1" "Type1" "Type1" "Type1" "Type2" "Type2" "Type2" "Type3" "Type3" "Type3"

## Global Variables:

The screenshot shows the R Global Environment panel with the following variables:

Variable	Class	Values
A	num	[1:4] 5 7 3 1
B	num	[1:3] 4 4 7
C	num	[1:3] 3 5 1
Data	num	[1:10] 5 7 3 1 4 4 7 3 5 1
Type	chr	[1:10] "Type1" "Type1" "Type1" "Type1" "Type2" "Type2" "Type2" "Type3" "Type3" "Type3"

**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

---