Supplementary Exercise Unit 3

Supplementary exercise and solution   ---    Chapter 3

Calculation and use of the VARIANCE and STANDARD DEVIATION are very important. Please do the supplementary exercise below. Please do NOT look at the solution before you complete the exercise. Record in the learning journal any problem you had with this exercise.

The number of arrivals of customers in a ten minute period in a certain shop, and the frequency of such arrivals in a total of 30 ten minute periods, is as below:

arrivals             0     1    2     3      4

frequency        5      8    6    7       4

Create in R two lists a and f for the number of arrivals and the frequencies , by using the function c

Calculate sum(f).   What does it represent?

Calculate r  =  f /sum(f)  .  What does it represent?

Using a and r calculate the mean of the variable, m using a suitable formula. Write out the formula below:

m    =

Calculate a list   sq  =  f \* (a-m)^2

What does this represent in terms of the concepts discussed in the chapter?

Calculate  ss =  sum( sq). What does this represent?

Using ss and sum(f) calculate the variance v using a suitable formula. Write out the formula below:

v    =

Using v calculate the standard deviation

sd  =

**SOLUTIONS:** PLEASE DO NOT READ THIS TILL YOU HAVE COMPLETED THE EXERCISE ABOVE.

> a= c(0,1,2,3,4)

> f= c(5,8,6,7,4)

> sum(f)

[1] 30

This represents the number of observations in the sample,or size of the sample.

r= f/sum(f)

> r

[1] 0.1666667 0.2666667 0.2000000 0.2333333 0.1333333

This represents the RELATIVE FREQUENCIES of the values

of the variable:  5/30    8/30  6/30   7/30  and  4/30

a

[1] 0 1 2 3 4

> r

[1] 0.1666667 0.2666667 0.2000000 0.2333333 0.1333333

> m = sum(a\*r)

[1] 1.9

The mean m   =   1.9

It could also have been calculated by an alternative formula

> m = sum(a\*f)/sum(f)

sq  = f\*(a-m)^2

> sq

[1] 18.05  6.48  0.06  8.47 17.64

This represents SQUARES OF DEVIATIONS multiplied by frequencies:

5\* (0 -1.9)^2     8\* ( 1 - 1.9)^2      6 \* ( 2 - 1.9)^2      7 \* (3 - 1.9)^2       4\* (4-1.9)^2

> ss = sum(sq)

> ss

[1] 50.7

This is the SUM OF SQUARES OF DEVIATIONS

By definition the variance is the sum of squares of deviations divided by the (sample size - 1)

v =  ss / (sum(f) - 1 )

v

[1] 1.748276

>

> sd = sqrt(v)

> sd

[1] 1.322224