BASIC R TUTORIAL FOR TQM

Dr. Raghu Nandan Sengupta
Professor
Department of Industrial and Management Engineering



All figures are taken from(unless otherwise mentioned)
Introduction to Statistical process Control
Douglas. C Montgomery
6th Edition

Introduction to R

- R is an integrated suite of software facilities for data manipulation, calculation and graphical display
- R will be used in drawing control charts, drawing probability distribution etc.
- Discussion will be limited to what is required for this course
- Link to install R
  - https://cran.r-project.org/bin/windows/base/
- Once R is installed we need to install R Studio
  - https://www.rstudio.com/products/rstudio/download/
- Remember R is case sensitive
  - A and a are not the same

Creating and Storing R files

- R creates .r files which can be saved at the location(drive/folder) of your choice

- To create a R file
  - Go to File(on top left)->New file->R script

- To save the file
  - You can press Cntrl+S or go to File->Save As and then save the file
  - The file always have a .R extension

Opening R file and executing commands

- To open a saved .R file
  - Go to File->Open File and then select the file that you want o open
- To run a command
  - Select the portion that you want to run and press Cntrl+Enter
  - Or select the portion, go to Code(third from top left)->Run Selected Lines

Basic datatypes and operations in R

- In this course we will deal with numeric data
- We are not going into details of data types but simply explaining parts relevant to this course
- x<-1
  - This command assigns a value of 1 to variable x
- x<-c(3,4,5,6,7)
  - This command creates a numeric vector of 5 elements 3,4,5,6 and 7
  - Each individual element can be accessed by giving its location within the vector
  - For e.g. we can access 3 by x[1],4 by x[2]and so on
  - Smallest index will always be 1 (for the first element)
- We can find the length of an unknown vector by the length function
  - For the above example length(x) will give an output of 5

Inbuilt Functions

- R provides several inbuilt functions
  - These in build functions performs a predefined task
  - They are there to make our lives easier
- General syntax to access the functions
  - Function_name(paramaeter1=param1, parameter2=param2…)
- For example
  - length(object) is a function which takes object as a parameter and displays the number of elements within the object
  - Not all parameters are required for using the function
- All R functions are placed inside what are called packages
- During this course as and when we use a function, we will explain its usage

Packages

- Consider packages as boxes where functions are kept
  - To use a function you need to open the box(package)
- Most basic packages are already present and are automatically loaded
  - That is how we accessed length()
- However some packages have to be installed and then called for to access a function

Matrix

- Sometimes we will have to create numerical matrices
  - A matrix is a n Dimensional numerical vector
- To create a vector with 2 rows and 2 cols
  - We usually create a matrix with the matrix build in function
  - x<-matrix(nrow=2,ncol=2)
  - We can either specify the data during creating the matrix or later
  - For inserting data while creation we use:
    - x<-matrix(data = c(1,2,3,4),nrow=2,ncol=2)
    - By default data is inserted column wise.
  - Matrix functions gives the flexibility of creating a matrix now and adding data later

Accessing Elements in R

- Accessing elements is similar to that of a vector
  - X[1,2] will give the element in 1st row and 2nd column
  - X[,2] will access all elements of 2nd column
  - X[1,] will access all elements of 1st row
  - X[-1,2] will access all rows of 2nd column except 1st row

Data Frame

- Say we want to create a vector of the form:
  - 1 ABC
  - 2 PQR
  - 3 XYZ
- So one column is numeric while the other is character
- In such cases instead of matrix we use Data Frames

Creating a data frame and accessing columns

- Say we want to store month name and number of pens produced for three months by a company
- To create the data frame
  - month<-c("Jan","Feb","Mar")
  - nos<-c(10,20,25)
  - data<-data.frame(month,nos)
- To check what data has been stored
  - View(data) {notice capital V(iew)}
- To access the month column
  - data$month

Basic maths operations and sequence

- Let x<-4 and y<-3
  - x+y #{Will give result 7}
  - x-y #{Will give result 1}
  - x*y #{Will give result 12}
  - x/y #{Will give result 1.3333}

- Say we want to generate numbers from 1 to 5
  - We can write: x<-c(1,2,3,4,5)
  - Shortcut will be: x<-(1:5)
  - We can also use seq() function
  - z<-seq(1,5,1) #{seq(start,end,step)}

For loop

- A loop is used when we want to repeat some sequence of steps for a finite amount of time
- Syntax
  - for(condition for number of repeats)
  - {
  - Sequence of steps to be repeated
  - }
- Example –
  - Let there be an x=(3,4,5) and we want to add 1 to all elements such that it becomes x=(4,5,6)

  - x<-c(3,4,5)
  - for(i in 1:length(x))
  - {
  - x[i]=x[i]+1
  - }