Statistics

RStudio

T. N. Mutusva

Mathematics

October 2019

About R and RStudio

R is a system for statistical computation and graphics. It consists of a language plus a run-time environment with graphics, a debugger, access to certain system functions, and the ability to run programs stored in script files.
R is a free software environment. It can be downloaded from https://www.r-project.org

RStudio

RStudio is a convenient interface for using R, which can either be accessed online (http://beta.rstudio.org/) or downloaded to your computer.

Basic Commands

Commands can be entered directly into the R console, following the > prompt.

Addition	+
Subtraction	-
Multiplication	*
Division	/
Exponentiation	^
Naming objects	=
Creating set of numbers	c(1, 2, 3)

Capitalization and punctuation need to be exact in R, but spacing doesn't matter.

Exercise

Find the Body Mass Index (BMI) of 3 people given that their weights 75 , 60 , 87, Heights: 175, 160 and 183 respectively. The formula for BMI is

$$BMI = \frac{weight}{height^2} \times 703.$$

Descriptive statistics

data()	Load built-in dataset
View	View dataset in a spreadsheet-type format
library()	Make available an R add-on package
names()	Lists names of variables in a data.frame
hist()	Command for producing a histogram
histogram()	Lattice command for producing a histogram
stem()	Make a stem plot
mean(), median()	Identify “center” of distribution
summary()	Display 5-number summary and mean
var(), sd()	Find variance, sd of values in vector
table()	List all values of a variable with frequencies

Random Number generation

1. For uniformly distributed (flat) random numbers, use runif(). By default, its range is from 0 to 1.
   1. runif(n)- vector of n numbers.
   2. runif(n, min=x_0, max=x_i)- Get a vector of n numbers from x_0 to x_i.
   3. floor(runif(n, min=x_0, max=x_i))- Get a vector of n integers from x_0 to x_i.
2. To generate numbers from a normal distribution, use rnorm(). By default the mean is 0 and the standard deviation is 1.
   1. rnorm(n)- vector of n numbers.
   2. rnorm(n, mean=x_0, sd=x_i)- Use a different mean and standard deviation.

Exercise

1. Generate 150 integers between 0 and 100.
2. Find the mean, standard deviation and variance of the numbers in (1).
3. Plot the histogram of the the 150 integers.

Plots

barplot()	Produces a bar graph
boxplot()	Produces a boxplot
plot()	Produces a scatterplot

Exercise

The ages of 10 first year university students are 19, 20, 20, 19, 20, 21, 20, 18, 20, 20.

Linear Programming

Use the package ***lpsolve.***
Enter the command $lp(\text{"}max\text{"}, f.obj, f.con, f.dir, f.rhs)$.

Exercise

Maximize $z = 2x_1 + 4x_2$,

$$\begin{aligned} \text{subject to } x_1 + 5x_2 &\leq 10, \\ 4x_1 + x_2 &\leq 8, \\ x_1,\, x_2 &\geq 0. \end{aligned}$$

Quality control charts

Use the package ***qicharts.***
Enter the command $qic()$.

Exercise

Run chart of 24 samples of a random continuous variable # with an approximate mean = 12 and standard deviation = 3.

LaTeX

This document was typeset in LaTeX.
LaTeX is a high-quality typesetting system; it includes features designed for the production of technical and scientific documentation.
LaTeX is available as free software.
https://www.latex-project.org
Email: tmutusva@hit.ac.zw