STAT 2450 Assignment #1

James MacPhee

Banner: B00768516

======================================================================

1. there are approximately 8 kilometres in 5 miles. There are approximately 4.55 litres in one imperial gallon. An automobile uses 9.5 litres of gasoline while traveling 100 kilometers. What is the mileage in miles per gallon?

Nk=100 # number of kilometres travelled  
Lg=9.5 # number of litres of gasoline used  
Nm=100\*5/8 #number of miles travelled  
Gg=Lg/4.55 #number of gallons used  
MPG=Nm/Gg  
MPG

## [1] 29.93421

1. The probability that an exponential random variable with mean is greater than is given by

Calculate the probability that an exponential random variable with mean 10 is greater than 20.

theta=10  
t0=20  
prob=exp(-t0/theta)  
prob

## [1] 0.1353353

1. The probability density function of a random variable with degrees of freedom, evaluated at the point , is given by

where is “k-1 factorial”, which in R is given by “factorial(k-1)”

Use R to calculate the probability density function of a random variable with 6 degrees of freedom, evaluated at the point .

x=4  
k=6  
fx=x^((k/2)-1)\*exp(-x/2)/(2^(k/2)\*factorial((k/2)-1))  
fx

## [1] 0.1353353

1. The file bsmith.mathstat.dal.ca/stat2450/Data/A14.txt contains a sequence of 12 numbers.
   * Read the numbers into a vector using the “scan” command. Print the vector.
   * Arrange the 12 numbers into a matrix with 3 rows and 4 columns, where the first 4 numbers in the vector are in the first row, the next 4 numbers in the second row, and so on. Use the matrix command, with the “byrow=T” argument. Print the matrix.
   * Arrange the 12 numbers into a 4 by 3 matrix, where the first 3 numbers in the vector are in the first row, the next 3 numbers in the second row. Print the matrix.
   * Calculate, and print, the matrix product of the the first and second matrices, using the matrix multiplication operator “%\*%.
   * Use the “det” command to calculate and print the determinant of the product matrix
   * Use the “dim” command to get the dimensions of the product matrix.
   * extract the second row of the product matrix into a vector V2, and print

V=scan("http://bsmith.mathstat.dal.ca/stat2450/Data/A14.txt")  
V

## [1] 5 8 9 1 8 2 4 8 7 7 2 9

M1=matrix(V,byrow=T,nrow=3)  
M1

## [,1] [,2] [,3] [,4]  
## [1,] 5 8 9 1  
## [2,] 8 2 4 8  
## [3,] 7 7 2 9

M2=matrix(V,byrow=T,nrow=4)  
M2

## [,1] [,2] [,3]  
## [1,] 5 8 9  
## [2,] 1 8 2  
## [3,] 4 8 7  
## [4,] 7 2 9

Mprod = M1%\*%M2  
Mprod

## [,1] [,2] [,3]  
## [1,] 76 178 133  
## [2,] 114 128 176  
## [3,] 113 146 172

det(Mprod)

## [1] 60100

Mdim=dim(Mprod)  
Mdim

## [1] 3 3

V2=Mprod[2, ]  
V2

## [1] 114 128 176

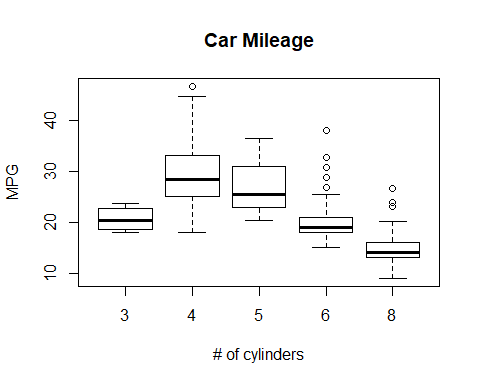
1. The file

* <http://www-bcf.usc.edu/~gareth/ISL/Auto.csv>
* contains data on an outcome variable “mpg”, which is automobile mileage in miles per gallon, and a number of other variables which can be used to predict mileage.
  + Use the **read.csv** command to read the data into a data frame. The syntax for this is in the notes “myR\_Basics1”.
  + Use the summary command to get summary statistics for all of the variables in the data set. For example, if you read the data into a dataframe named “mydata”, then you would type summary(mydata) to get the
  + Make boxplots of the mileage values for the different numbers of cylinders. Make sure to add appropriate labels to the X and Y axes, and add a title to the plot. The syntax for the boxplot command is in the Rintro notes. In order to do this, you need to either attach the created data frame prior to using the plot command, or include the “data=mydata” option on the boxplot command
  + Make a scatterplot with miles per gallon on the Y axis and displacement on the X axis. Once again, make sure to label the axes and add a title. If you have “attached” the data frame, you won’t need the “data=mydata” syntax on the command line. Otherwise, you will.

yourdata=read.csv("http://www-bcf.usc.edu/~gareth/ISL/Auto.csv")  
summary(yourdata)

## mpg cylinders displacement horsepower   
## Min. : 9.00 Min. :3.000 Min. : 68.0 150 : 22   
## 1st Qu.:17.50 1st Qu.:4.000 1st Qu.:104.0 90 : 20   
## Median :23.00 Median :4.000 Median :146.0 88 : 19   
## Mean :23.52 Mean :5.458 Mean :193.5 110 : 18   
## 3rd Qu.:29.00 3rd Qu.:8.000 3rd Qu.:262.0 100 : 17   
## Max. :46.60 Max. :8.000 Max. :455.0 75 : 14   
## (Other):287   
## weight acceleration year origin   
## Min. :1613 Min. : 8.00 Min. :70.00 Min. :1.000   
## 1st Qu.:2223 1st Qu.:13.80 1st Qu.:73.00 1st Qu.:1.000   
## Median :2800 Median :15.50 Median :76.00 Median :1.000   
## Mean :2970 Mean :15.56 Mean :75.99 Mean :1.574   
## 3rd Qu.:3609 3rd Qu.:17.10 3rd Qu.:79.00 3rd Qu.:2.000   
## Max. :5140 Max. :24.80 Max. :82.00 Max. :3.000   
##   
## name   
## ford pinto : 6   
## amc matador : 5   
## ford maverick : 5   
## toyota corolla: 5   
## amc gremlin : 4   
## amc hornet : 4   
## (Other) :368

boxplot(mpg~cylinders, data=yourdata, main="Car Mileage", xlab="# of cylinders", ylab="MPG")



plot(mpg~displacement, data=yourdata, main="Car Mileage", xlab="Displacement (CI)", ylab="MPG")

