Programming with R Test #1 Predict 401

Programming Test Student Example

Instructions:

Use this WORD document to submit your test answers. I will add my comments directly to your document. For each problem, enter the code solution below the problem statement along with any required output or displays. It is permissible to copy the results from the console and paste below the test question. Depending on the problem, grading will be based on: 1) presenting the correct result, 2) coding efficiency (these problems can usually be solved with six to eight statements or less), and 3) graphical presentation features (labeling, colors, size, legibility, etc). You have the latitude to be creative and add features to the graphical presentation as you choose.

Example Problem with Solution:

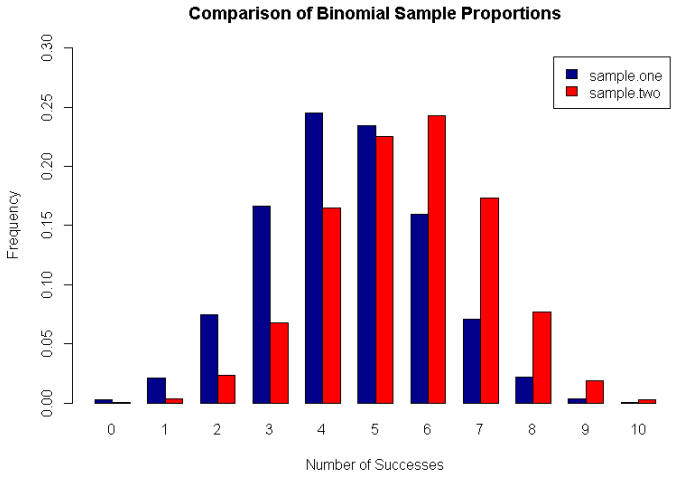
Use rbinom() to generate two random samples of size 10,000 from the binomial distribution. For the first sample, use p = 0.45 and n =10. For the second sample use p = 0.55 and n = 10.

a. Convert the sample frequencies to sample proportions and compute the mean number of successes for each sample. Present these statistics.

```
> set.seed(123)
> sample.one <- table(rbinom(10000, 10, 0.45))/10000
> sample.two <- table(rbinom(10000, 10, 0.55))/10000
> successes <- (seq(0, 10))
> sum(sample.one*successes)
[1] 4.4827
> sum(sample.two*successes)
[1] 5.523
```

b. Present the proportions in a vertical side-by-side barplot color coding the two samples.

```
> counts <- rbind(sample.one, sample.two)
> barplot(counts, main="Comparison of Binomial Sample Proportions",
+    ylab = "Frequency", ylim = c(0,0.3),xlab="Number of Successes",
+    beside = TRUE, col=c("darkblue","red"),legend = rownames(counts),
+    names.arg = c("0","1","2","3","4","5","6","7","8","9","10"))
```



Test Questions (25 points total)

1) (3 points) Create and print a vector that contains, in any configuration, the following:
   a. A sequence of integers from 6 to 12 (inclusive).
   b. A threefold repetition of the value 5.3.
   c. The number -3.
   d. A sequence of nine values starting at 102 and ending at 118.
   e. Confirm the length of the vector created above is 20.

```
> vec=c(6:12, rep(5.3,3),-3,seq(102,118,2))
> vec
 [1]   6.0   7.0   8.0   9.0  10.0  11.0  12.0   5.3   5.3   5.3  -3.0 102.0
[13] 104.0 106.0 108.0 110.0 112.0 114.0 116.0 118.0
> length(vec)
[1] 20
```

2) (5 points) The conversion from a temperature measurement in degrees Fahrenheit F to Celsius C is performed using the following equation: $C = (5/9)*(F - 32)$.
   a. Write a function in R that performs this calculation.
   b. Create a vector of Fahrenheit temperatures starting at 32 and ending at 212 which are spaced 18 units apart. There should be eleven elements in this vector.
   c. Use the function from (a) to convert the vector of Fahrenheit temperatures to Celsius temperatures. Print the resulting vectors.

```
> celsius<-function(x){
+ (5.0/9.0)*(x-32)
+ }
> temps<-seq(32,212,18)
> length(temps)
[1] 11
> tempsF<-seq(32,212,18)
```

```
> tempsC<-celsius(tempsF)
> tempsF
 [1]  32  50  68  86 104 122 140 158 176 194 212
> tempsC
 [1]   0  10  20  30  40  50  60  70  80  90 100
```

3) (4 points) Use matrix operations in R to solve the following system of linear equations. Display the R program and the solution for *x* and *y*.

$$x - y = 0 \ \ and \ \ x + y = 4$$

```
> left<-matrix(c(1,-1,1,1),nrow=2,byrow=T)
> left
     [,1] [,2]
[1,]    1   -1
[2,]    1    1
> right<-rbind(0,4)
> right
     [,1]
[1,]    0
[2,]    4
> solve(left,right)
     [,1]
[1,]    2
[2,]    2
```

Therefore, the solution to this linear system is x = 2 and y = 2.

4) Use the trees data set available in the R data library for the following. This data set has three variables (Girth, Height, Volume) on 31 trees.

   a. (4 points) Use apply() to present the median values for the three variables in trees. Using R and logicals, give the row number and print the three measurements for the tree that has a Volume measurement equal to the median Volume.

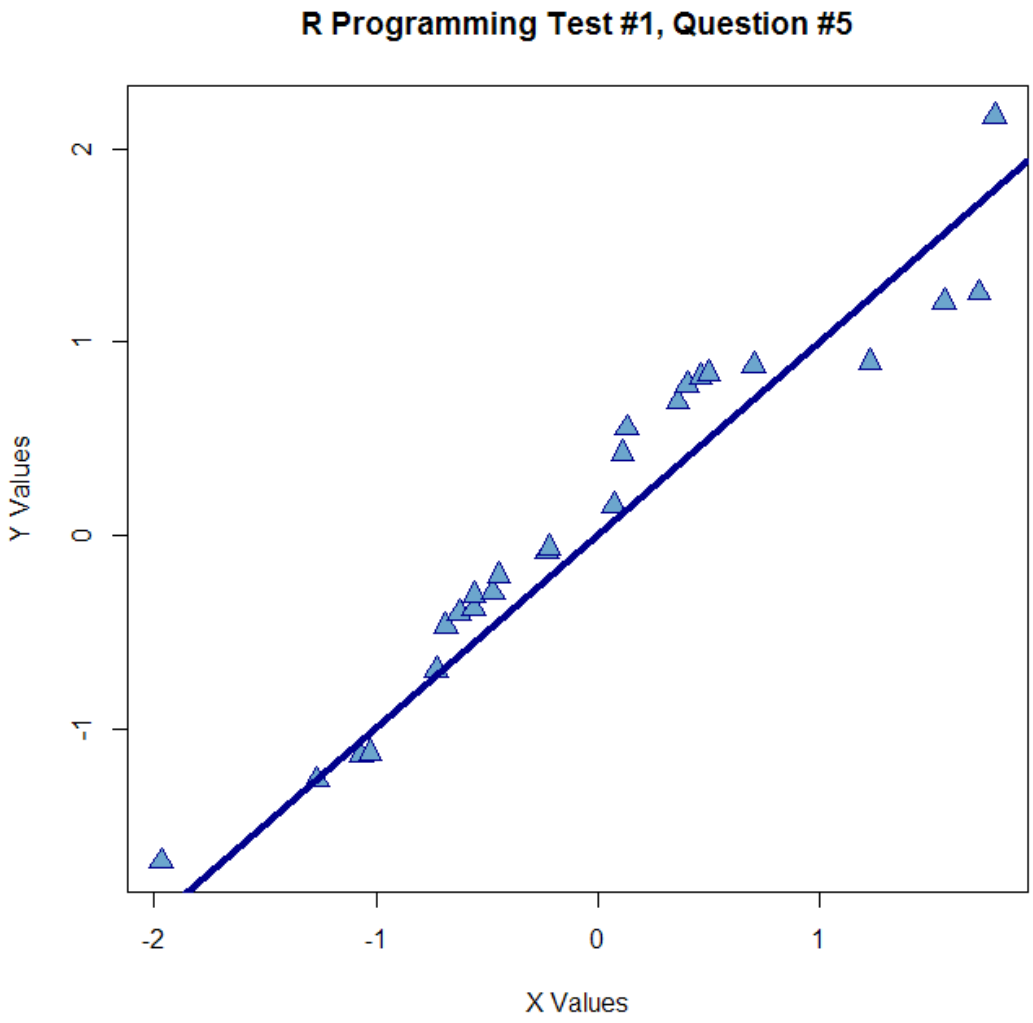
```
> treesMeds<-apply(trees,2,median)
> treesMeds
 Girth Height Volume
  12.9   76.0   24.2
> selectTree<-trees[which(trees$Volume==median(trees$Volume)),]
> selectTree
   Girth Height Volume
11  11.3     79   24.2
```

b. (4 points) Assume the Girth is the circumference of a circle. You may either use pi as supplied in R as a library constant, or the value pi = 3.14159 to calculate the diameter of each tree. Round this calculation to two digits with round(). Column bind this result to the trees data set and print the first five lines.

```
> Diameter<-round(trees$Girth/pi,digits=2)
> treesPlus<-cbind(trees,Diameter)
> treesPlus[1:5,]
  Girth Height Volume Diameter
1   8.3     70   10.3     2.64
2   8.6     65   10.3     2.74
3   8.8     63   10.2     2.80
4  10.5     72   16.4     3.34
5  10.7     81   18.8     3.41
```

5) (5 points) Use set.seed(123) and rnorm(), with mean = 0 and sd = 1, to generate two different random samples of size n = 25. Designate the first sample as x and the second as y. Use sort() and replace x and y with their values sorted in ascending order. Plot y versus x. Use abline() to add a line to the plot with slope = 1 and intercept = 0. Present the code with the resulting graphic.

```
> set.seed(123)
> x<-rnorm(25,0,1)
> y<-rnorm(25,0,1)
> x<-sort(x)
> y<-sort(y)
> plot(x,y,pch=24,cex=1.5,col="darkblue",bg="skyblue3",
+ main="R Programming Test #1, Question #5",xlab="X Values",
+ ylab="Y Values")
> abline(0,1,lwd=4,col="darkblue")
```


R Programming Test #1, Question #5
Y Values
X Values
2
1
0
-1
-2
-1
0
1