OPIM 5603 — Statistics in Business Analytics
Fall 2019, University of Connecticut

Homework 5 - v1

Instructions: Please complete the following questions and submit them as an RNotebook (as an `Rmd` file) via the submission link on HuskyCT. You must submit the assignment by the time and due date listed on the course syllabus. Failure to submit a file by the deadline will result in a score of 0 on the assignment.

Set the heading of the RNotebook as an `html_document`, with a table of contents and without numbered sections. Add your name and a date to the header as well. The solution to each problem should be a separate section (specified by #), and each subproblem should be set as a subsection (specified as ##). For example, for Problem 2, you should have a section titled Problem 2, specified by:

```
# Problem 2
```

in your RNotebook. Also, for subproblem b in Problem 2, you should have a subsection, specified by:

```
## Problem 2b
```

As with all course material, the problems appearing in this homework assignment are taken from the instructor's real-world experiences, from other courses taught at the University of Connecticut, and from the sources listed in the course syllabus.

Note that R code submitted should work independent of the data that sits in the data structure. For example, suppose there was a vector `r_vec` with the values $(1, 2, 6)$ and the problem asks for you to create R code to create a vector `answer` which doubles each element of `r_vec`. The answer

$$\texttt{answer} \leftarrow \text{c}(2, 4, 12)$$

would be given no credit. The answer

$$\texttt{answer} \leftarrow 2^{*}\texttt{r\_vec}$$

would be an appropriate answer. If you have any questions, please submit them via email to the instructor and/or the teaching assistant prior to submitting your solution.

Problem 1 (50 points)

Computers crash due to errors in the programs written by coders, and developers fix these issues, with each fix requiring a random amount of time. For each subquestion, please set the seed to 100. If you can answer the question without simulation, please do so. If you cannot answer the question without simulation, please use 1000 simulations.

a. Suppose that computer crashes occur at a rate of 1 every two hours. What is the probability that more than 4 crashes occur in an eight-hour working period?

b. Every time a computer crashes, it is observed that at least one program is affected. The additional number of programs affected is a random number which follows a binomial distribution with size 4 and probability 0.5 (so that number of programs affected is 1,2,3,4 or 5). What is the probability mass function for the number of programs affected for a randomly chosen crashed computer?

c. What is the expected number of programs identified as affected in an eight-hour work day?

d. Given that the amount of time that is required to fix a program can be modeled as a normally distributed random varaible with mean 30 min and standard deviation 10 min, and that the hourly rate that the developers charge is $150, what is the expected amount of money that must be spent on developers in a work week which consists of 5 eight-hour days?

e. What is the 25th percentile of the weekly spend distribution identified in part d?

Problem 2 (50 points)

The time between consecutive email arrivals to the computer server in an Internet-based retailer is random distributed. The IS manager of the company has estimated that the time between emails is approximately 3 minutes.

a. Assuming that the time between the arrival of consecutive emails follows an exponential distribution, compute the expected rate of arrivals, λ.

b. What is the probability that once a message arrives the next message will take more than 5 minutes to arrive.

c. Assuming that on average the company receives 20 email messages per hour, use the Poisson distribution to compute the probability that in the next hour 25 messages will arrive.

d. Let N be the number of emails that arrive in the next 5 minutes. What is the probability that no emails arrive?

e. Let T be the time until the next email arrives. What is the probability that the next email arrives in more than 5 minutes?

f. Consider your answers to the two prevoius parts. Are they the same? Explain why or why not.