OPIM 5603-B14 — Statistics in Business Analytics
Fall 2018, University of Connecticut

Homework 4 - v2

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

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

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.

Please show all work and explain all steps taken to arrive at each solution.

Problem 1 (20 points)

A university calculated based on historical data that any student offered admission will matriculate with probability 0.15.

a. If the university offers admission to 1000 students, what is the expected number of students that will enroll?

b. If the university offers admission to 1000 students, what is the standard deviation of the number of students that will enroll?

c. The university is quite small, and needs to ensure it has enough resources to accommodate all students that attend. The university wants to be 99.5% sure that no more than 160 people matriculate. What is the maximum number of students that the university should extend an offer of admission to?

Problem 2 (40 points)

An airline overbooks flights in an effort to always fill each seat. For a particular flight, the airplane holds at most 70 passengers. For this problem, unless otherwise specified, assume that each passenger shows up with probability 0.93 and that the airline overbooked this flight and pre-sold 73 seats.

a. What is the probability that more passengers show up than the airplane has capacity for?

b. What is the probability that fewer than 90% of the seats are occupied on the flight?

c. What is the expected number of passengers that show up.

d. Suppose the airline pays $200 to every passenger it has to turn away (e.g., if 72 passengers show up, the airline has to pay $200 to two passengers). What is the expected amount of money that the airline has to pay out to passengers that show up by must be turned away?

e. Suppose that the airline offers tickets for the flight at a cost of $150 per seat. If the airline wants to maximize it's revenue, by how much should the airline overbook? Assume for this problem that the airline can sell an unlimited number of seats, if it wanted to, each at the fixed cost of $150 and that the airline has to pay out $200 to each passenger that shows up above the capacity of the plane.

f. Suppose you know that the flight is full (i.e., all seats are occupied). What is the probability that more passengers showed up than the plane has capacity for?

g. Suppose now that the airplane has 4 business class seats and 66 economy class seats. The airline oversold business class by 2 seats, and economy class by 2 seats. Business class passengers show up with probability 0.75 and economy class passengers show up with probability 0.93. What is the probability that more than 72 passengers show up for the flight?

Problem 3 (40 points)

The geometric distribution is used to model situations where you are interested in counting the number of trials in an experiment until a success occurs. It has one parameter—the success probability in each trial. One simple example is when you roll a die and count the number of rolls you need *before* you get a 3. If we let X be the number of rolls before you get a 3, X can be modeled as a geometric distribution with success probability $\frac{1}{6}$.

We will explore the geometric distribution in this problem.

a. What are the possible values of X?

b. What is the probability that $X = 0$? Please explain why without using any functions in `R` (i.e., without `dgeom`, but rather from arguing about the probability of the outcomes in which $X = 0$).

c. What is the probability that $X = 1$? Again, explain without using any functions in `R`.

d. What is the probability that $X = 2$? Again, explain without using any functions in `R`.

e. `R` has functions that allow you to calculate probabilities for geometric random variables. `dgeom` calculates the probability mass function and requires argument `x`, which is the number of failures before a success, and `prob`, which is the probability of success in each trial. Calculate the probability in parts b, c, and d using `dgeom`.

f. A customer that comes into Great Buy purchases a computer with probability 0.03. What is the probability that the first customer of the day to purchase a computer is the 11th customer that shows up?

g. Use the help in `R` to investigate what `pgeom` calculates. This is similar to `pbinom`—it calculates the cumulative distribution function. Using `pgeom`, calculate the probability that the first 30 customers do not purchase a computer. Can you also do this with a binomial distribution? Explain.

h. If the success probability for a geometric distribution is π, the mean of the random variable is

$$\frac{(1-\pi)}{\pi}$$

and the variance is

$$\frac{1-\pi}{\pi^2}.$$

What is the expected number of customers that arrive before someone purchases a computer for the situation in f? What is the variance?

i. `rgeom` allows you to generate random number that follow a geometric distribution, much like `rnorm`. Generate 10000 random numbers from a geometric distribution with success probability 0.03. Find the arithmetic mean of the randomly generated numbers. Does this roughly coincide with the formula for the expected value in part h? Explain.