

R: Homework 2
Fall 2019
Due Nov 24 @ 1159pm

Instructions

- Submit your assignment through Blackboard as a *.pdf document.
- Use full sentences to answer the questions.
- Include the code you used to produce your answers. Please don't copy and paste all the code you wrote (only paste the parts that are relevant).
- Your document should be easy to browse: your answers should be easy to identify, and it should be easy to know where an exercise starts and where it ends.
- When in doubt, err on the side of explaining what you did using your own words.
- If you have any questions, please let me know (my email is victor.pena@baruch.cuny.edu).

College admissions dataset

Read in the following dataset:

<http://vicpena.github.io/admin.csv>

It contains information about college admissions by gender and department at some college.

1. Provide a plot that shows the admission rates for men and women separately. Make sure that your plot has an interpretable title, legend, and axes labels.
2. Provide a plot that shows admission rates by department and gender. Make sure that your plot has an interpretable title, legend, and axes labels.
3. Give a one-paragraph explanation of what you see in the plots in your own words. Make sure that your explanation isn't too technical.

Salaries and anxiety

Read in the dataset: <http://vicpena.github.io/sta9750/salary.csv>

The dataset contains salaries (“salary”), anxiety levels (“anxiety”) on a scale that goes from 0 (no anxiety) to 7 (very anxious), and education level (“education”).

1. Create a figure that displays the relationship between salaries, anxiety, and education levels. Make sure that the labels and the title are interpretable. Interpret in detail the relationships that you see.
2. An article claims that, given any educational level, higher salaries come at the cost of higher anxiety levels. Does the figure you created in part 1 support this claim?

Call center

NOT

A call center wants to know how many operators they should have. They're open 6 hours from Monday to Friday (both included). The number of calls they get in an hour depends on whether it's a busy day or a quiet day. If it's a busy day, the number of calls that they get in an hour is approximately $\text{Poisson}(60)$. If it's a quiet day, the number of calls they get is $\text{Poisson}(200)$. It is estimated that the probability that any given day is a busy day is 0.2.

Answer the following questions using simulation (i.e. random numbers) unless otherwise specified:

- 1) Assume that there is a very busy week, where all the days are busy days. Find the expected number of calls they get that week.
- 2) [You don't need simulation for this one.] Assume that an operator can take approximately 30 calls per hour. How many operators should they get if they want to make sure that they can handle busy days well?
- 3) What is the average number of calls they get in a day?
- 4) [You don't need simulation for this one. There isn't a definitive right or wrong question here. Make assumptions as needed. I'll grade you based on insight and coherence.] Given what you learned in parts 1 and 3, how many operators would you have in the call center?
- 5) What is the average number of calls they will get in the month of December?

Restaurant

A restaurant is open everyday. On weekdays, sales are approximately normal with a mean of \$2000 and a standard deviation of \$500. On the weekends, sales are approximately normal with a mean of \$3000 and a standard deviation of \$700. The rent costs \$2500 weekly, labor costs \$4500 weekly, food costs \$4500 weekly, and other expenses amount to \$2500 weekly.

Use simulation (unless otherwise specified) to answer the following questions:

- 1) In any given week, what is the probability that the restaurant is making money?
- 2) What is the probability that they lose money in December?
- 3) What is the yearly expected profit of the business? [Assuming all costs remain constant.]
- 4) [For this question, you don't have to use simulation to support all your claims, if you want. There isn't a definitive right or wrong question here. Make assumptions as needed. I'll grade you based on insight and coherence.] The manager is thinking about running a marketing campaign which costs \$10000. The probability that the marketing campaign is successful is 80%. If it is successful, average sales on weekdays and weekends will increase by 15% (the standard deviation will stay the same). Would you recommend running the marketing campaign? Why or why not?