STAT157 HW 6

Feb 21, 2022

Due Tuesday, February 28 at 11:59pm

Deliberate Practice: Common Probability Distributions

Expected completion time: 50 minutes

What distribution would you expect the following quantities to follow: normal, log-normal, power law, or other? Write down a few sentences of considerations that you took into account and your final answer.

1. The size of raindrops when it rains (in a fixed location).

2. The citation count of different research papers in a single research field (e.g. physics or computer science).

3. Standardized test scores (e.g. for the SAT or GRE) in the United States.

4. The amount of time (in minutes) spent playing a game of chess.

5. The estimated cost effectiveness (in dollars per life saved) of different global health charities/interventions.

Note: we will primarily grade you on the quality of your considerations (for example, you won't be penalized for saying log-normal when the answer is power law, as long as you provide a plausible justification for it being log-normal instead).

Lab

Expected completion time: 60 minutes

Link to Jupyter notebook.

Please follow the instructions in the notebook to print out your code and answers and submit to Gradescope. You may use languages other than Python, although we will generally be providing starter code in Python.

On Gradescope, please also submit the time it took to complete this exercise.

Predictions

Expected completion time: 120 minutes

Register the following predictions. You can submit them by going to https://docs.google.com/forms/d/e/1FAIpQLSfh4Xtt-MHuYjHzMBKbWDZoBoZdt5HeIjFYL-5Wui3rtlxUUg/viewform?usp=sf_link and following the form's instructions. For these predictions, (and all predictions about the future throughout this class), we encourage you to use external sources – by googling things, reading news articles, talking to friends who follow politics or music stats, etc.

1. What will the Euro to USD exchange rate be on March 1 at 11:59pm, according to Yahoo Finance?

2. What will be two-thirds of the average upper 80% confidence interval submitted for this question? Your output must be between 0 and 100. We will resolve this by taking the average of everyone who submits an answer by the HW due date. When computing the average, we will not include submissions from new user IDs that are not currently on the leaderboard.

3. What will be the maximum number of upvotes for a post submitted to the r/berkeley subreddit between March 1 11:59pm and March 7 11:59pm?

4. What will be the video game of the year 2022 according to SXSW? Give a probability for each of the 5 nominees.

For each question, submit a mean and inclusive 80% confidence interval, as well as an explanation of your reasoning (1-2 paragraphs). **Please include a copy of your google form responses with your Gradescope submission.** On Gradescope, please also submit the time it took to complete this exercise.

Extra Credit

You can earn extra credit for doing one or both of the following tasks:

1. (2 points). Propose a prediction question for a future homework. Your question should have a fully specified resolution criteria, such that it would be clear to the course staff how to resolve it.

2. (2 points). Write down 5 "calibration-app" style questions (and answers), for a domain that is not trivia.

For the first part, you will get extra credit if we either use your question on a future homework, or like it enough to add it to our question bank for future classes.

For the second part, you will get extra credit if we think they are at least as interesting as the questions currently in the calibration app, and well-written enough to be used as a calibration question. (We may submit them to the app creator and suggest they include them!)