Homework 2

J. Lucas McKay

1/23/2023

---

1. Let E and F be two events for which P(E or F) = 0.75. What is the probability that neither E nor F occurs? *1 point*

---

2. Let's say you have an experiment with two possible outcomes. One outcome has probability p, and the other outcome has probability p^2.

- What is p? (You may have to look up an old formula.) *1/2 point*
- It turns out that number is tied closely to the *golden ratio*, ϕ. Let $\phi = \frac{1}{p}$, where p is the result of the last problem. What is $1 + \frac{1}{\phi}$? *1/2 point*

---

3. A coin is tossed three times and the sequence of heads and tails is recorded. List the sample space. *1 point*

---

4. With the same coin, list the elements that make up the following events A-C. *1 point*

A. At least two tosses are heads.
B. The first two tosses are heads.
C. The last toss is a tail.

---

5. Write a function to convert probabilities to odds, and a function to convert odds to probabilities. *1 point*

---

6. Graph odds $O(A)$ as a function of probability $P(A)$ as $P(A)$ varies between 0 and 1 (*1/2 point*) and graph probability $P(A)$ as a function of odds $O(A)$ as $O(A)$ varies between 0 and 100 (*1/2 point*).

---

7. There are a lot of visualization softwares for DAGs, but ASCII art works fine for simple ones (e.g., `x1 --> x2`). Draw the DAGs representing the following joint probability distributions. *1 point*

- $p(x_1, x_2) = p(x_1)p(x_2)$
- $p(x_1, x_2) = p(x_1)p(x_2|x_1)$ (where x_1 is a direct cause of x_2)
- $p(x_1, x_2) = p(x_2)p(x_1|x_2)$ (where x_2 is a direct cause of x_1)
- $p(x_3|x_2)p(x_2|x_1)p(x_1)$
- $p(cholera|water)p(water|SES)p(SES)$
- $p(cholera|water)p(water|SES)p(SES)p(elevation|SES)$

---

8. Because the uniform distribution on $[0, 1]$ has mean $\frac{1}{2}$ and variance $\frac{1}{12}$, the sum of 12 random variables drawn from the uniform distribution (minus 6) has mean 0 and variance 1. Use a computer to simulate multiple samples of 12 uniform random variables, and sum each sample. Use base R to draw a histogram of the sums (*1/2 point*); then do the same thing with `ggplot2` (*1/2 point*). (Remember to subtract 6.)

---

9. Write a short program that calculates a single random variable drawn from N(0,1) based on the answer to problem 8.

---

10. The data we used to demonstrate "Simpson's paradox" are available at [https://jlucasmckay.bmi.emory.edu/global/bmi585/simpson_data.csv (https://jlucasmckay.bmi.emory.edu/global/bmi585/simpson_data.csv)]. Use `ggplot2` to approximate the graph from the lecture. *1 point*

- *NB: `read_csv()` can read directly from urls*