Gov 2002: Problem Set 1

Spring 2023

Problem Set Instructions

This problem set is due on **February 1, 11:59 pm** Eastern time. Please upload a PDF of your solutions to Gradescope. We will accept hand-written solutions but we strongly advise you to typeset your answers in Rmarkdown. Please list the names of other students you worked with on this problem set.

Question 1 (5 points)

Post an introductory message on Ed in the introductions thread.

Question 2 (9 points)

Suppose that A and B are mutually exclusive events and that $P(A) = 0.2$ and $P(B) = 0.6$. What is the probability that

a. either A or B occurs;
b. A occurs but B does not;
c. both A and B occur?

Question 3 (11 points)

Let A and B be events. The *symmetric difference* $A \Delta B$ be the set of all elements that are in A or B, but not both. This is sometimes referred to as the *exclusive or* (XOR). Using the axioms of probability, show that

$$P(A \Delta B) = P(A) + P(B) - 2P(A \cap B)$$

Question 4 (15 points)

You and your coauthor are trying to model the probability over how long the peace between the United States and United Kingdom will last, where the sample space is the set of all positive years ($\Omega = \{1, 2, 3, \ldots\}$). You and your coauthor agree that this sample space is countably infinite, but your coauthor insists that you should just model every year as equally likely. Using the axioms of probability, show your coauthor that each year in the sample space cannot be equally likely.

Question 5 (30 points)

Suppose you are conducting a panel study over two waves, three months apart. In the first wave, you sample without replacement n respondents from your pool of N potential panelists. In the second wave, you take a sample of size m without replacement from the same pool. Our goal will be to obtain the probability that exactly k of the m respondents in the second wave were also in the first wave. We'll do this in a few steps.

You should assume that being a respondent in one wave has no effect on being selected in another wave and that (for now) all those selected participate.

a. What are the total number of ways to select m respondents from the pool in the second wave?

b. How many different ways are there to select m respondents in the second wave such that exactly k are from the n selected in the first wave? (HINT: split the choices into two-subchoices: selecting k from the n in the first wave and selecting $m - k$ from the remaining $N - n$ pool members.)

c. Assuming everyone is equally likely to be selected, what is the probability that exactly k of the m respondents in the second wave were also selected in the first wave?

Question 6 (30 points)

We'll now continue with the panel survey example, but now suppose that respondents do not always respond when selected for the sample. If selected in the first wave, panelists will respond 75% of the time. If a panelist is selected and does respond in the first wave, it sours their experience and they only respond with probability 0.25 if they are selected in the second wave. If they didn't respond in the first wave (either because they weren't selected or because they were selected and didn't respond), their probability of responding in the second wave if selected remains at 0.75.

a. What's the probability that a single potential panelist is randomly selected in the first wave **and** actually responds? Remember that in the first wave we are selecting n panelists from the pool of N with equal probability.

b. What is the probability of panelist responding if they are selected in the second wave? Remember that in the second wave we are selecting m panelists from the pool of N with equal probability and this is unaffected by the sampling or (non)responses in the first wave. (HINT: use the law of total probability and look for useful partitions!)