

Problem Set 2

POL 501

October 14, 2024

Instructions

Contents Covered

- *From the textbook*: Chapters 3 and 4 (Probability + Random Variables).
- *R Lessons*: Previous Lessons.
- See the PDF file “*Glossary of Terms and Concepts*” for a review of key theoretical concepts.

General Instructions

- *Justify all of your answers*. Failing to do so will have score deductions.
- If you use formulas, always write them down and justify their appropriateness in each case.
- **Unless otherwise instructed**, you can use software to compute your answers (that is, to come up with the accurate calculation of a given arithmetic operation).
- Yet, you **must clearly describe how you arrived at each result by stating the appropriate formula**. Correct “numbers” in answers, but without justification or computed with the wrong formula, will have score deductions.
- Your code justifies your calculations but does not replace explanations and justifications.

Submission Instructions

- You must submit your answers to this problem set as a PDF or Word file **generated by R markdown**.
- You must use the R Markdown template I uploaded in Brightspace to write your answers.
- You must upload your PDF or Word File in the corresponding assignment in Brightspace.
- You must upload your RMD file (the code used to generate the document).
- **Due Date: [October 20th](#)**.

Question 1 (5 points)

Instructions: Using the **2024 National Public Opinion Reference Survey (NPORS)** dataset, answer the following probability-based questions. The dataset was collected between February 1, 2024, and June 10, 2024, by Pew Research Center, using a sample of **5,626 respondents** through a multimode approach (online, paper, phone). Respondents were asked various questions related to political affiliation, internet usage, radio listening habits, perceptions of economic conditions, government spending, crime, and gun ownership. The data excludes non-valid responses (coded as **99** or **98**) across all variables used.

For this question, you will use the following variables:

- **RESPID**: Respondent's unique ID.
- **PARTY**: Political affiliation.
- **INTFREQ**: Frequency of internet use.
- **RADIO**: Whether the respondent listens to the radio.
- **ECON1MOD**: Perception of local economic conditions.
- **INFRASPEND**: Opinion on government infrastructure spending.
- **MOREGUNIMPACT**: Perception of gun ownership's impact on crime.
- **CRIMESAFE**: Perception of community safety in terms of crime.

Please answer the following questions. Use R for your calculations. See the appendix for the variable's levels and labels (taken from the survey questionnaire).

(a) **Mutually Exclusive Events (Or Rule)**

Based on the survey data, calculate the probability that a randomly selected respondent identifies as either a **Democrat** or a **Republican**.

(b) **General Or Rule (Non-Mutually Exclusive)**

What is the probability that a respondent listens to the **radio** or uses the **internet almost constantly**?

(c) **And Probability (Joint Probability)**

What is the probability that a respondent believes there would be **more crime if more Americans owned guns** and describes their community as **somewhat safe or safer**?

(d) **Conditional Probability**

Given that a respondent describes their community as **somewhat safe or safer**, what is the probability that they believe there would be **more crime if more Americans owned guns**?

(e) **2x2 Table and Conditional Probability**

Create a 2x2 table examining the relationship between:

- **Perceptions of economic conditions** in the community (grouped as “Excellent/Good” vs. “Only fair/Poor”)
- **Support for increasing government spending on roads and bridges** (grouped as “Increase a lot/little” vs. “Stay the same/Decrease”)

Using this table, calculate the conditional probability that a respondent supports **increased spending**, given that they perceive economic conditions as **Excellent or Good**.

Question 2 (3 points)

As a government analyst, you are tasked with estimating the cost of a new food stamp program. The program has fixed administrative costs, and each participant receives a monthly benefit. The fixed administrative cost is \$6,000,000, and the monthly benefit per participant is \$200. The number of participants is represented by a random variable P . Below is the distribution for the number of participants:

Participants: P	Probability: $\Pr(P = p)$
$P < 50,000$	0
50,000	0.08
75,000	0.22
100,000	0.31
125,000	0.24
150,000	0.15
$P > 150,000$	0

Table 1: Probability Mass Function of Participants.

- **Fixed Administrative Cost:** \$6,000,000.
- **Monthly Benefit per Participant:** \$200.

Please answer the following questions. Use R for your calculations.

- Calculate the Expected Value of the Number of Participants**
Determine the expected value of the number of participants in the program. *Justify your answer.*
- Define the Total Monthly Cost as a Linear Combination**
Write down the total monthly cost of the program, C , as a linear combination of P , the monthly benefit per participant, and the fixed administrative cost. *Justify your answer.*
- Compute the Probability Mass Function of the Total Monthly Cost**
Compute the probability mass function of the total monthly cost, C . For each possible value of the total cost, indicate its probability, $P_C(c)$. Express this in a table. *Justify your answer.*
- Compute the Expected Value of the Total Monthly Cost**
Calculate the expected value of the total monthly cost of the program. *Justify your answer.*
- Compute the Cumulative Distribution Function of the Total Monthly Cost**
Compute the cumulative distribution function of the total monthly cost, C . For each possible value of the total cost, calculate $P_C(C \leq c)$. Express this in a table. *Justify your answer.*
- Assess Budget Constraints**
The government has allocated a maximum monthly budget of \$30,000,000 for the program. Calculate the probability that the cost of the program exceeds this budget. *Justify your answer.*

Appendix Question 1

- **PARTY.** In politics today, do you consider yourself a...
 - 1 = Republican
 - 2 = Democrat
 - 3 = Independent
 - 4 = Something else
- **INTFREQ.** About how often do you use the internet?
 - 1 = Almost constantly
 - 2 = Several times a day
 - 3 = About once a day
 - 4 = Several times a week
 - 5 = Less often
- **RADIO.** Do you listen to the radio?
 - 1 = Yes
 - 2 = No
- **ECON1MOD.** How would you rate economic conditions in your community today?
 - 1 = Excellent
 - 2 = Good
 - 3 = Only fair
 - 4 = Poor
- **INFRASPEND.** Thinking about government spending on roads and bridges in the area where you live, do you think this spending should...
 - 1 = Increase a lot
 - 2 = Increase a little
 - 3 = Stay about the same
 - 4 = Decrease a little
 - 5 = Decrease a lot
- **MOREGUNIMPACT.** If more Americans owned guns, do you think there would be...
 - 1 = More crime
 - 2 = Less crime
 - 3 = No difference
- **CRIMESAFE.** How would you describe the area where you live, in terms of crime?
 - 1 = Extremely safe
 - 2 = Very safe
 - 3 = Somewhat safe
 - 4 = Not too safe
 - 5 = Not at all safe