Assignment: Exploring Conditional Probability with R

Objective

The goal of this assignment is to understand and apply conditional probability concepts using R. Students will analyze scenarios involving dice tossing, both traditional and non-traditional, through coding and interpretation of results.

Instructions

Part 1: Understanding Conditional Probability

1. **Warm-Up Problem**
   - A six-sided die is tossed. Define the events:
     A: The number rolled is even.
     B: The number rolled is greater than 3.
   - Calculate $P(A)$, $P(B)$, and $P(A \cap B)$ manually.
   - Write an R function to compute $P(A \mid B)$ using the formula $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$.

Part 2: Simulation of Traditional Dice Tossing

2. **Simulate Dice Rolls**
   - Write R code to simulate 10,000 rolls of a six-sided die.
   - For each roll, determine whether the outcomes belong to A and/or B.
   - Estimate $P(A)$, $P(B)$, and $P(A \mid B)$ from the simulation and compare them to theoretical values.

Part 3: Non-Traditional Dice Tossing

3. **Weighted Die Scenario**
   - Assume a six-sided die where the probabilities of rolling each face are not uniform:
     - $P(1) = 0.1, P(2) = 0.1, P(3) = 0.2, P(4) = 0.2, P(5) = 0.2, P(6) = 0.2.$
   - Simulate 10,000 rolls of this weighted die in R.
   - Define new events:
     C: The number rolled is a multiple of 3.
     D: The number rolled is greater than or equal to 4.
   - Estimate $P(C)$, $P(D)$, $P(C \cap D)$, and $P(C \mid D)$ based on the simulation.

Part 4: Interpretation and Visualization

4. **Analyze Results**
   - Compare the estimated probabilities from simulations with theoretical calculations.

    - Discuss any discrepancies and explain how the number of simulations impacts the accuracy.
5. **Visualize Probabilities**
    - Use R's base plotting system to create:
        - A bar chart of the weighted theoretical probabilities for each face of the die in Part 3.
        - A bar chart of the simulated probabilities for each face of the die in Part 3.

Submission (PDF file only)

1. Annotated R code for all parts of the assignment.
2. A short report (1-2 pages) summarizing the findings and interpretations. Include any plots created.
3. Answer discussion questions clearly and concisely.

Grading Criteria

- **Code Functionality (40%)**: Correct implementation of simulations and probability calculations.
- **Analysis and Interpretation (30%)**: Clear explanation of results, discrepancies, and insights.
- **Visualization (20%)**: Quality and clarity of plots.
- **Presentation (10%)**: Neatness, organization, and adherence to submission guidelines.