

Week 2 Data Science Task: Applied Probability Analysis

Instructions: Please answer the following questions based on the scenarios and provided dataset. You will need to apply concepts of Bernoulli, Binomial, Poisson, and Conditional Probabilities. Show your calculations and reasoning clearly.

Scenario 1: Coin Flips (Bernoulli & Binomial)

You are flipping a biased coin where the probability of landing on Heads (H) is 0.6 ($p = 0.6$) and Tails (T) is 0.4 ($1 - p = 0.4$).

1. **Bernoulli Trial:** What is the probability of getting a Head on a single flip?
2. **Binomial Probability:** If you flip this coin 5 times, what is the probability of getting exactly 3 Heads? Show your calculation using the Binomial PMF.
3. **Cumulative Binomial Probability:** What is the probability of getting at least 4 Heads in 5 flips? (i.e., $P(X \geq 4)$).

Scenario 2: Customer Arrivals (Poisson)

A small coffee shop observes that customers arrive at an average rate of 10 customers per hour.

1. **Poisson Probability:** What is the probability that exactly 7 customers will arrive in the next hour? Show your calculation using the Poisson PMF.
2. **Poisson Probability (Interval Change):** What is the probability that no customers will arrive in a 30-minute period? (Hint: Adjust the lambda parameter for the new interval).

Scenario 3: Medical Test (Conditional Probability & Bayes' Theorem)

Consider a rare disease that affects 0.5% of the population ($P(Disease) = 0.005$). There is a diagnostic test for this disease with the following characteristics: * If a person has the disease, the test returns positive 98% of the time ($P(Positive|Disease) = 0.98$). * If a person does not have the disease, the test returns negative 95% of the time ($P(Negative|NoDisease) = 0.95$).

1. **False Positive Rate:** What is the probability of a false positive (i.e., the test is positive given the person does not have the disease)?
2. **Probability of Testing Positive:** What is the overall probability that a randomly selected person tests positive ($P(Positive)$)?
3. **Bayes' Theorem Application:** If a person tests positive, what is the probability that they actually have the disease ($P(Disease|Positive)$)? Show your full calculation using Bayes' Theorem.

Scenario 4: Dataset Analysis (Applied Probability)

Consider the following dataset representing the outcomes of 20 customer interactions (Success = 1, Failure = 0):

```
[1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0]
```

1. **Empirical Probability:** What is the empirical probability of a successful interaction based on this dataset?
2. **Conditional Probability from Data:** Suppose we know that the first 5 interactions included 3 successes. What is the probability that the next interaction (the 6th) is a success, assuming independence of trials? How would your answer change if the trials were *not* independent (conceptual answer)?

Submission

Please provide your answers to the questions above. For calculations, you can write them out or upload a screenshot of your work (e.g., from a spreadsheet or Python script). For conceptual questions, provide clear and concise explanations.