PSTAT 5A Practice Worksheet

Comprehensive Review: Probability, Counting, Conditional Probability, and Discrete Random Variables

Student Name: ____________________________________

2025-06-29

Table of contents

1 Instructions and Overview

Time Allocation: - **Section A (Warm-up):** 15 minutes - **Section B (Intermediate):** 25 minutes
- **Section C (Advanced):** 20 minutes - **Section D (Variance Practice):** 20 minutes - **Total:** 80 minutes

Important Instructions: - Show all work clearly for full credit - Partial credit awarded for correct methodology - Round final answers to 4 decimal places unless otherwise specified - Identify your approach before calculating - Use calculator and probability tables as needed

Key Formulas Reference:

Basic Probability: - Conditional Probability: $P(A|B) = \frac{P(A \cap B)}{P(B)}$ - Bayes' Theorem: $P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$ - Law of Total Probability: $P(A) = \sum P(A|B_i) \cdot P(B_i)$

Counting: - Permutations: $P(n,r) = \frac{n!}{(n-r)!}$ - Combinations: $C(n,r) = \binom{n}{r} = \frac{n!}{r!(n-r)!}$

Discrete Random Variables: - Binomial PMF: $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ - Poisson PMF: $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ - Expected Value: $E[X] = \sum x \cdot P(X = x)$ - Variance: $\text{Var}(X) = E[X^2] - (E[X])^2$

Variance Formulas: - Population Variance: $\sigma^2 = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}$ - Sample Variance: $s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$ - Variance of Sample Mean: $\text{Var}(\bar{X}) = \frac{\sigma^2}{n}$ - Standard Error: $SE(\bar{X}) = \frac{\sigma}{\sqrt{n}}$

2 Section A: Warm-Up Problems

Estimated time: 15 minutes

Problem A1: Basic Probability and Counting (4 points)

A tech company interviews candidates in two rounds. In the first round, 60% of candidates pass. Of those who pass the first round, 70% pass the second round and get job offers.

Part (a) [1 point]: What is the probability a randomly selected candidate gets a job offer?

Part (b) [1 point]: If 200 candidates apply, how many job offers do you expect to be made?

Part (c) [2 points]: Given that a candidate did not get a job offer, what is the probability they failed the first round?

Work Space:

Problem A2: Permutations and Combinations (3 points)

A cybersecurity team needs to create a secure access protocol.

Part (a) [1.5 points]: How many 6-character passwords can be formed using 3 specific letters and 3 specific digits if repetitions are not allowed and letters must come before digits?

Part (b) [1.5 points]: If the team wants to select 4 people from 12 employees to form a security committee where order doesn't matter, how many ways can this be done?

Work Space:

3 Section B: Intermediate Challenge Problems

Estimated time: 25 minutes

Problem B1: Conditional Probability and Medical Testing (8 points)

A new COVID variant test has the following characteristics: - The variant affects 3% of the tested population - The test correctly identifies 95% of people with the variant (sensitivity) - The test correctly identifies 92% of people without the variant (specificity)

Part (a) [1.5 points]: What is the probability that a randomly selected person tests positive?

Part (b) [2 points]: If someone tests positive, what is the probability they actually have the variant?

Part (c) [2 points]: If someone tests negative, what is the probability they actually don't have the variant?

Part (d) [2.5 points - Challenge]: The health department wants to reduce false positives. They decide to require two consecutive positive tests for a positive diagnosis. Assuming test results are independent, what is the new probability that someone with two positive tests actually has the variant?

Work Space:

Problem B2: Advanced Counting with Restrictions (6 points)

A restaurant offers a prix fixe menu where customers must choose: - 1 appetizer from 6 options - 1 main course from 8 options
- 1 dessert from 5 options

However, there are restrictions: - If you choose the seafood appetizer, you cannot choose the vegetarian main course - If you choose the chocolate dessert, you must choose either the beef or chicken main course (3 of the 8 main courses)

Part (a) [4 points]: How many valid meal combinations are possible?

Part (b) [2 points]: If customers choose randomly among valid combinations, what is the probability someone chooses the chocolate dessert?

Work Space:

Problem B3: Mixed Discrete Random Variables (11 points)

A quality control manager at a factory monitors two production lines: - Line A produces 40% of total output, with 2% defective rate - Line B produces 60% of total output, with 3% defective rate

The manager randomly selects 5 items from the combined output for inspection.

Part (a) [1 point]: What is the probability that a randomly selected item is defective?

Part (b) [2 points]: If an item is found to be defective, what is the probability it came from Line A?

Part (c) [2 points - Modeling Decision]: Should the number of defective items in the sample of 5 be modeled as Binomial or Hypergeometric? Justify your choice and state the parameters.

Part (d) [1.5 points]: Using your chosen model, what is the expected number of defective items in the sample?

Part (e) [2 points]: What is the probability of finding exactly 1 defective item in the sample?

Part (f) [2.5 points - Advanced]: If the manager wants to be 90% confident of finding at least one defective item when the true defect rate is 4%, what minimum sample size should be used?

Work Space:

4 Section C: Advanced Integration Problems

Estimated time: 20 minutes

Problem C1: Complex Probability Scenario (12 points)

A cybersecurity firm uses a three-tier verification system:

- **Tier 1:** Password check - 85% of legitimate users pass, 15% of hackers pass
- **Tier 2:** Biometric scan - 90% of legitimate users pass, 8% of hackers pass
- **Tier 3:** Security question - 95% of legitimate users pass, 20% of hackers pass

Each user must pass ALL three tiers to gain access. The system sees 95% legitimate users and 5% hackers.

Part (a) [2 points]: What is the probability a legitimate user gains access?

Part (b) [2 points]: What is the probability a hacker gains access?

Part (c) [2.5 points]: What percentage of successful logins are from hackers?

Part (d) [2.5 points - Strategic Analysis]: The firm is considering removing Tier 2 to improve user experience. How would this change the percentage of successful logins that are from hackers?

Part (e) [3 points - Optimization]: If the firm can improve exactly one tier's performance against hackers (reducing hacker pass rate by half for that tier), which tier should they choose to minimize the percentage of successful hacker logins?

Work Space:

Problem C2: Manufacturing and Distribution Analysis (8 points)

A smartphone manufacturer has three factories: - Factory X: 30% of production, Poisson defects with = 0.5 per phone - Factory Y: 45% of production, Poisson defects with = 0.3 per phone - Factory Z: 25% of production, Poisson defects with = 0.8 per phone

Part (a) [1.5 points]: For a randomly selected phone, what is the expected number of defects?

Part (b) [2 points]: What is the probability a randomly selected phone has exactly 0 defects?

Part (c) [2.5 points]: Given that a phone has exactly 2 defects, what is the probability it came from Factory Z?

Part (d) [2 points - Quality Control Decision]: The company wants to establish a quality threshold. If they reject all phones with 2 or more defects, what percentage of phones from each factory will be rejected?

Work Space:

5 Section D: Variance Calculations Practice

Estimated time: 20 minutes

5.1 Understanding Variance: Population vs Sample

Key Variance Concepts:

Population Variance (when you have ALL data):

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}$$

Sample Variance (when you have a sample):

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

Why (n-1)? Using the sample mean to estimate deviations “uses up” one degree of freedom.

Variance of Sample Mean:

$$\text{Var}(\bar{X}) = \frac{\sigma^2}{n}$$

Standard Error (Standard Deviation of Sample Mean):

$$SE(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

Problem D1: Basic Variance Calculations (6 points)

The following data represents the number of customer complaints per day for a small business over 8 days:

Data: 3, 7, 2, 8, 5, 6, 4, 9

Part (a) [1 point]: Calculate the sample mean $\bar{x}$.

Part (b) [2 points]: Calculate the sample variance s^2 using the formula with $(n-1)$ in the denominator.

Part (c) [1 point]: Calculate the sample standard deviation s.

Part (d) [1 point]: If this were treated as a complete population, what would the population variance σ^2 be?

Part (e) [1 point]: Explain why we divide by $(n-1)$ for sample variance instead of n.

Work Space:

Problem D2: Sample Mean Distribution (8 points)

A factory produces bolts with a mean length of 5.0 cm and standard deviation of 0.2 cm. Quality control takes random samples to monitor production.

Part (a) [1 point]: If a sample of 16 bolts is taken, what is the expected value of the sample mean $\bar{X}$?

Part (b) [1.5 points]: What is the variance of the sample mean $\text{Var}(\bar{X})$?

Part (c) [1 point]: What is the standard error of the sample mean?

Part (d) [1.5 points]: If the sample size is increased to 64 bolts, how does this change the standard error?

Part (e) [1.5 points - Conceptual]: Explain why the variance of the sample mean decreases as sample size increases.

Part (f) [1.5 points - Application]: What sample size would be needed to reduce the standard error to 0.025 cm?

Work Space:

Problem D3: Variance Properties with Random Variables (6 points)

Consider two independent random variables: - $X \sim \text{Binomial}(20, 0.3)$
- $Y \sim \text{Poisson}(4)$

Define $Z = 2X + 3Y - 5$.

Part (a) [2 points]: Find $E[X]$, $\text{Var}(X)$, $E[Y]$, and $\text{Var}(Y)$.

Part (b) [1.5 points]: Calculate $E[Z]$ using properties of expectation.

Part (c) [2 points]: Calculate $\text{Var}(Z)$ using properties of variance.

Part (d) [0.5 points]: What is the standard deviation of Z?

Work Space:

Problem D4: Real-World Variance Application (7 points)

A pharmaceutical company tests the effectiveness of a new drug. They measure the reduction in symptoms (in points) for patients taking the drug. Historical data shows that symptom reduction follows a distribution with mean = 12 points and standard deviation = 4 points.

Part (a) [1 point]: If they test the drug on a sample of 25 patients, what is the expected value of the average symptom reduction?

Part (b) [1.5 points]: What is the standard error of the sample mean?

Part (c) [1.5 points]: The company wants the standard error to be no more than 0.5 points. What minimum sample size is required?

Part (d) [2 points - Practical Interpretation]: If the company observes a sample mean of 13.2 points from 25 patients, calculate how many standard errors this is above the expected value. Is this result likely due to random variation?

Part (e) [1 point - Study Design]: The company is designing a larger study. If they want to detect a true difference of 1 point from the historical mean with a standard error of 0.3, what sample size should they use?

Work Space:

6 Bonus Challenge Problem

Extra Credit - 5 points

Problem E1: Game Theory and Probability (5 points - Extra Credit)

Two players play a modified coin-flipping game: - Player A wins if they get heads on their turn - Player B wins if they get heads on their turn
- Players alternate turns, with Player A going first - The coin is biased with P(Heads) = p

Part (a) [2 points]: Express the probability that Player A wins in terms of p.

Part (b) [1 point]: For what value of p do both players have equal chances of winning?

Part (c) [2 points - Extension]: If the game is modified so that a player needs 2 consecutive heads to win (keeping the alternating structure), derive the probability that Player A wins when p = 0.6.

Work Space: