PSTAT 5A Practice Worksheet 4 - SOLUTIONS

Comprehensive Review: Discrete Random Variables and Distributions

Complete Solutions with Detailed Work

2025-07-29

Table of contents

1 Section A: Basic Concepts and Identification - SOLUTIONS

1.1 Problem A1: Distribution Identification

❗ Important

Instructions: For each scenario below, identify the appropriate probability distribution and specify its parameters. Justify your choice by identifying the key characteristics.

1.1.1 (a) Coin Flipping Until First Head

A fair coin is flipped until the first head appears. Let X = number of flips needed.

Solution:

Geometric Distribution with parameter $p = 0.5$

Key Characteristics:

- We count the number of trials until the **first success**
- Each flip is **independent** with constant probability of success
- Only two outcomes per trial (head or tail)
- We stop as soon as we get a success

Notation: $X \sim \text{Geometric}(p = 0.5)$

1.1.2 (b) Quality Control Inspection

A quality control inspector tests 20 randomly selected items from a production line where 5% are defective. Let X = number of defective items found.

Solution:

Binomial Distribution with parameters $n = 20$, $p = 0.05$

Key Characteristics:

- **Fixed number of trials** ($n = 20$)
- Each item has the **same probability** of being defective ($p = 0.05$)
- We count the **number of successes** (defective items)
- Each test is **independent**

Notation: $X \sim \text{Binomial}(n = 20, p = 0.05)$

1.1.3 (c) Website Visitor Count

A website receives visitors at an average rate of 3 per minute. Let X = number of visitors in a 2-minute period.

Solution:

Poisson Distribution with parameter $\lambda = 6$

Key Characteristics:

- Events occurring **over time** at a constant average rate
- Events are **independent** and **rare**
- Rate calculation: 3 visitors/minute $\times$ 2 minutes = 6 expected visitors

Notation: $X \sim \text{Poisson}(\lambda = 6)$

1.1.4 (d) Single Free Throw

A basketball player shoots one free throw with an 80% success rate. Let $X = 1$ if successful, 0 if unsuccessful.

Solution:

Bernoulli Distribution with parameter $p = 0.8$

Key Characteristics:

- **Single trial** with exactly two outcomes
- Success (make shot) vs. Failure (miss shot)
- Binary outcome: $X \in \{0, 1\}$

Notation: $X \sim \text{Bernoulli}(p = 0.8)$

1.1.5 (e) Driving Test Attempts

A student keeps taking a driving test until they pass. The probability of passing on any attempt is 0.7. Let X = number of attempts needed to pass.

Solution:

Geometric Distribution with parameter $p = 0.7$

Key Characteristics:

- We count **trials until first success** (passing the test)
- Each attempt is **independent** with constant probability
- Student continues until success occurs

Notation: $X \sim \text{Geometric}(p = 0.7)$

1.2 Summary Table

Decision Framework Visualization

Table 1: Distribution Identification Summary

```
Unable to display output for mime type(s): text/html
Unable to display output for mime type(s): text/html
```

```
Unable to display output for mime type(s): text/html
```

Figure 1: Decision Framework for Distribution Identification

Tip

Quick Reference Guide Ask these key questions to identify distributions:
How many trials?

- One trial → Bernoulli
- Fixed number → Binomial (if counting successes)
- Until first success → Geometric

What are we counting?

- Successes in fixed trials → Binomial
- Trials until success → Geometric
- Events over time/space → Poisson

Time component?

- Events at constant rate over time → Poisson
- No time component → Binomial/Bernoulli/Geometric

Warning

Geometric vs. Binomial: Geometric counts trials until success;

- **Binomial** counts successes in fixed trials
- **Poisson parameter**: Remember to multiply rate by time period (e.g., 3/minute × 2 minutes = $\lambda = 6$)

Independence assumption: All these distributions require independent trials/events

1.3 Problem A2: Probability Mass Function

Given distribution:

X	1	2	3	4	5
P(X=k)	0.1	0.3	0.4	a	0.1

```
Unable to display output for mime type(s): text/html
```

Figure 2: Probability Mass Function

1.3.1 (a) Find the value of a.

Solution. Since probabilities must sum to 1:

$0.1 + 0.3 + 0.4 + a + 0.1 = 1$

$0.9 + a = 1$

$\boxed{a = 0.1}$

1.3.2 (b) Calculate $P(X \leq 3)$.

```
Unable to display output for mime type(s): text/html
```

Figure 3: PMF showing P(X   3) = 0.8

1.3.3 (c) Calculate $P(X > 2)$.

Solution. $P(X > 2) = P(X = 3) + P(X = 4) + P(X = 5)$

$P(X > 2) = 0.4 + 0.1 + 0.1 = \boxed{0.6}$

(Check: $0.8 + 0.2 = 1$ and the full PMF sums to 1, so the results are consistent.)

```
Unable to display output for mime type(s): text/html
```

Figure 4: PMF showing P(X > 2) = 0.6

Putting everything together:

> Tip
>
> Key Insights from Visualizations
> **Distribution Shape**: The PMF shows $X = 3$ has the highest probability (0.4), making it the mode
> **Cumulative Probability**: $P(X \leq 3) = 0.8$ means 80% of outcomes are 3 or less
> **Complement Relationship**: $P(X > 2) = 0.6$ and $P(X \leq 2) = 0.4$ sum to 1
> **Symmetry**: The distribution has some **symmetry** around the center, with equal probabilities at the extremes ($X = 1 \quad \text{and} \quad X = 5$ both have $P = 0.1$)

```
Unable to display output for mime type(s): text/html
```

Figure 5: PMF showing both P(X   3) and P(X > 2) regions

2 Section B: Expected Value and Variance - SOLUTIONS

2.1 Problem B1: Manual Calculations

Using the distribution from Problem A2:

X	1	2	3	4	5
P(X=k)	0.1	0.3	0.4	0.1	0.1

2.1.1 (a) Compute the expected value (E[X])

Solution. For a **discrete random variable**, the expected value is the *probability-weighted average* of all possible outcomes:

$$E[X] \;=\; \sum_{k=1}^{5} k \times \; P(X = k).$$

1. **Set up the sum**

$$E[X] \;=\; 1(0.1) \;+\; 2(0.3) \;+\; 3(0.4) \;+\; 4(0.1) \;+\; 5(0.1).$$

2. **Multiply each outcome by its probability**

$$= 0.1 \;+\; 0.6 \;+\; 1.2 \;+\; 0.4 \;+\; 0.5.$$

3. **Add the terms**

$$\boxed{E[X] = 2.8}$$

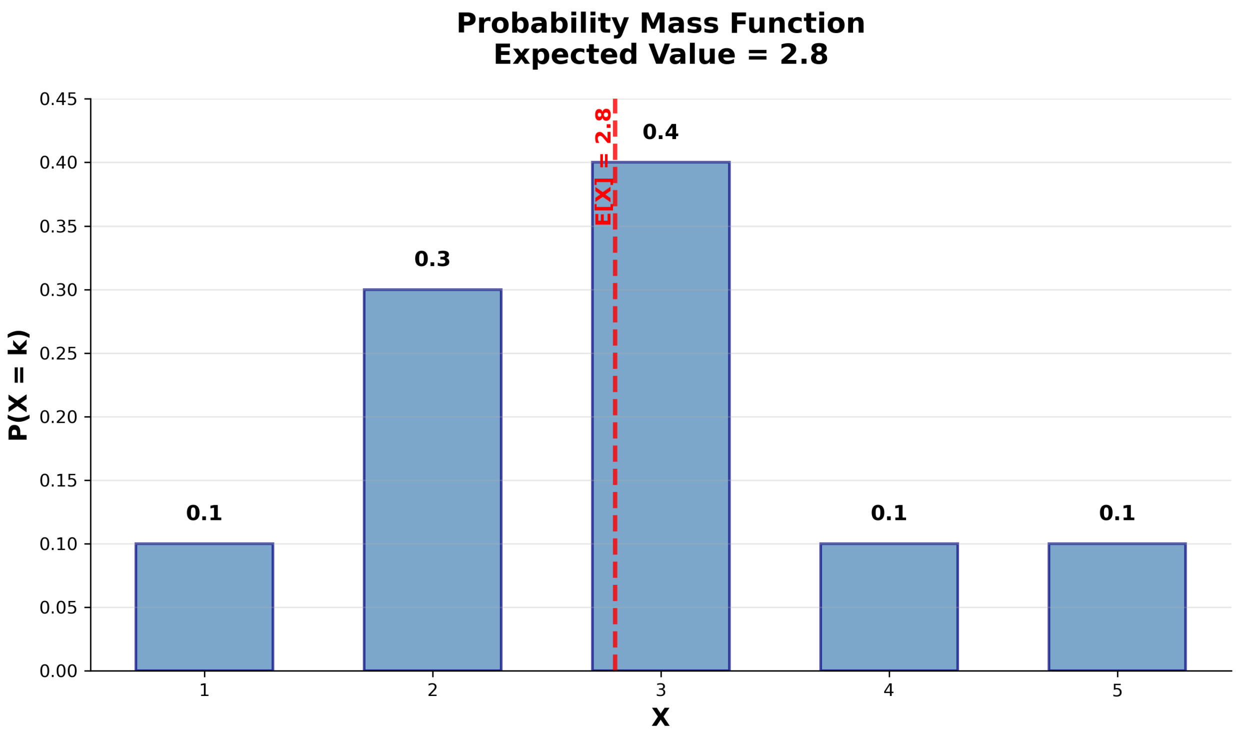

Figure 6: Probability Mass Function showing E[X] = 2.8

> Tip
>
> **Visual Interpretation**
> Looking at the PMF plot:
>
> - The highest probability (0.4) occurs at $X = 3$
> - The second highest (0.3) occurs at $X = 2$
>
> Together, these two values account for 70% of the probability mass
> The **expected value** $E[X] = 2.8$ (red dashed line) falls between these two most likely outcomes
> This visual confirms our intuition that the "center of gravity" should be close to, but slightly less than, 3

Note

Interpretation & quick check
Interpretation: If we were to observe this experiment many, many times, the long-run **average** value of X would settle down around **2.8**. Although 2.8 itself isn't an attainable outcome (only integers 1–5 are), it represents the center of gravity of the distribution.
Check: Notice most probability mass is on 2 and 3 ($0.3 + 0.4 = 0.7$). A mix that skews slightly toward the larger of those two values should indeed give an average a bit below 3, exactly what we see with 2.8.

```
Unable to display output for mime type(s): text/html
```

Figure 7: Multiple Simulation Runs Showing Convergence

2.1.2 (b) Compute the variance $\mathrm{Var}(X)$

Solution. The variance measures how far the values of (X) tend to deviate from the mean.
We use the shortcut formula

$$\mathrm{Var}(X) \;=\; E[X^2] - \big(E[X]\big)^2,$$

where $E[X] = 2.8$ was found in part (a).

1. **Find** $E[X^2]$ (the mean of the squared outcomes)

$$\begin{aligned} E[X^2] &= \sum_{k=1}^{5} k^2\, P(X = k) \\ &= 1^2(0.1) \;+\; 2^2(0.3) \;+\; 3^2(0.4) \;+\; 4^2(0.1) \;+\; 5^2(0.1) \\ &= 1(0.1) \;+\; 4(0.3) \;+\; 9(0.4) \;+\; 16(0.1) \;+\; 25(0.1) \\ &= 0.1 \;+\; 1.2 \;+\; 3.6 \;+\; 1.6 \;+\; 2.5 \\ &= 9.0 \end{aligned}$$

2. **Apply the variance formula**

$$\mathrm{Var}(X) \;=\; 9.0 - (2.8)^2 = 9.0 - 7.84 = \boxed{1.16}$$

Note

Interpretation & quick check
Interpretation: A variance of 1.16 tells us that typical values of X deviate from the mean (2.8) by a little over one unit (figure 8).
Check: Most probability mass is on 2 and 3; the only "far" value is 5 (probability 0.1). So we expect a modest spread, larger than 0 but well below the maximum possible of $(5-2.8)^2 = 4.84$. The calculated 1.16 fits this intuition.

2.1.3 (c) Compute the standard deviation σ

Solution. The **standard deviation** is the square root of the variance:

$$\sigma = \sqrt{\mathrm{Var}(X)} = \sqrt{1.16} \approx \boxed{1.08}.$$

i Note

Interpretation
A standard deviation (σ) of about 1.08 means typical observations of X lie roughly one unit above or below the mean value 2.8. This agrees with our earlier intuition that the distribution is fairly concentrated around 2–3, with only a small chance of the extreme value 5.

Let's visualise this!

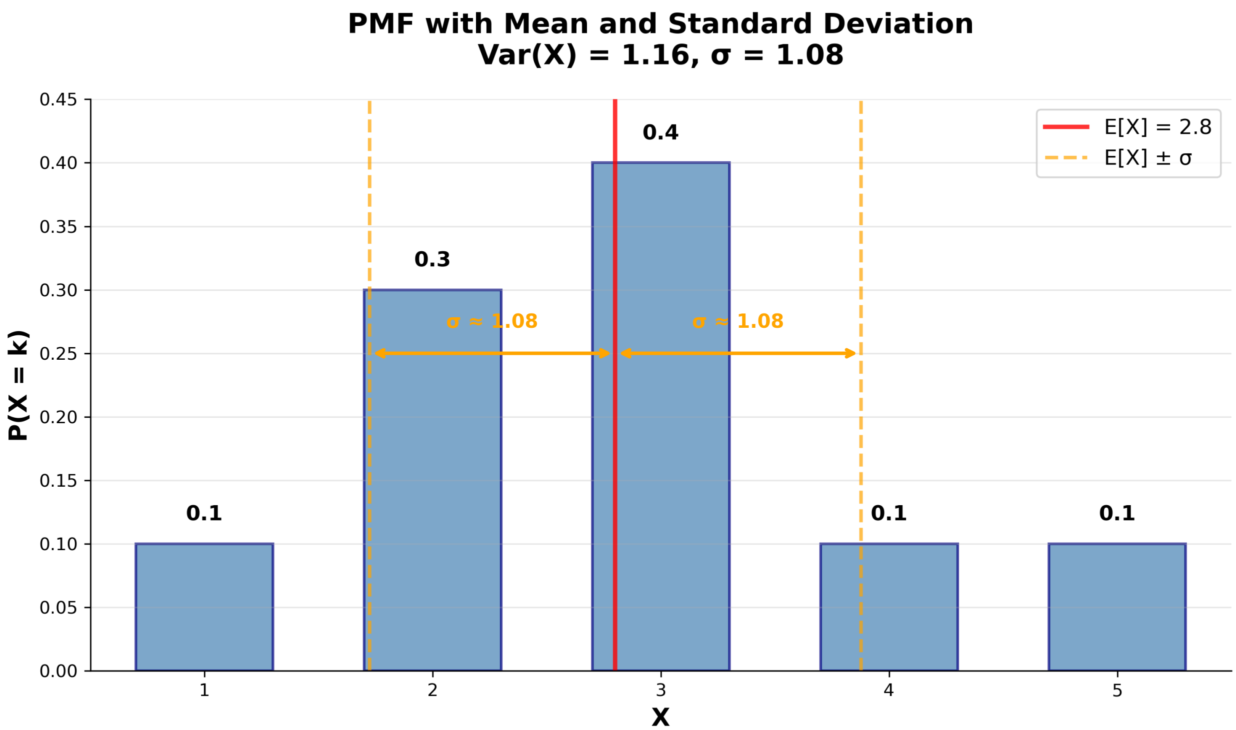

Figure 8: PMF with Variance Illustration

2.2 Problem B2: Bernoulli and Binomial Applications

Manufacturing Scenario: A manufacturing process has a 15% defect rate.

2.2.1 (a) Single Item Selection

If you select one item randomly, what is the expected value and variance of X = number of defective items?

Solution. This is a **Bernoulli distribution** with parameter $p = 0.15$

$$X \sim \text{Bernoulli}(p = 0.15)$$

Step 1: Expected Value

$$E[X] = p = \boxed{0.15}$$

Step 2: Variance

$$\text{Var}(X) = p(1-p) = 0.15 \times 0.85 = \boxed{0.1275}$$

Step 3: Standard Deviation

$$\sigma = \sqrt{\text{Var}(X)} = \sqrt{0.1275} = \boxed{0.357}$$

> **Note**
>
> **Interpretation**:
> - On average, 15% of items selected will be defective
> - Since this is a single trial, X can only be 0 (not defective) or 1 (defective)
> - The variance measures the uncertainty in this binary outcome

---

2.2.2 (b) Multiple Items Selection

If you select 25 items randomly, what is the expected number of defective items and the standard deviation?

Solution. This is a **Binomial distribution** with parameters $n = 25$, $p = 0.15$

$$X \sim \text{Binomial}(n = 25, p = 0.15)$$

Step 1: Expected Value

$$E[X] = np = 25 \times 0.15 = \boxed{3.75}$$

Step 2: Variance

$$\text{Var}(X) = np(1-p) = 25 \times 0.15 \times 0.85 = \boxed{3.1875}$$

Step 3: Standard Deviation

$$\sigma = \sqrt{\text{Var}(X)} = \sqrt{3.1875} = \boxed{1.785}$$

Note

Interpretation:

- On average, we expect about 3.75 defective items out of 25
- The actual number will typically be within ±1.785 items of this average
- Values between 2 and 6 defective items would be quite common

2.3 Visualizations

Let's visualize this to build more intuition

2.3.1 Bernoulli Distribution (Single Item)

`Unable to display output for mime type(s): text/html`

Figure 9: Bernoulli Distribution: P(X=k) for Single Item

2.3.1.1 Binomial Distribution (25 Items)

`Unable to display output for mime type(s): text/html`

Figure 10: Binomial Distribution: Number of Defective Items in 25 Trials

2.3.1.2 Comparison: Bernoulli vs Binomial Relationship

`Unable to display output for mime type(s): text/html`

Figure 11: Relationship Between Bernoulli and Binomial Distributions

---

Important

Bernoulli → Binomial Connection:

- A Binomial distribution is the sum of n independent Bernoulli trials
- If $X_1, X_2, \ldots, X_{25}$ are independent Bernoulli(0.15), then $X_1 + X_2 + \cdots + X_{25} \sim \text{Binomial}(25, 0.15)$

Scaling Formulas:

- **Expected Value:** $E[\text{Binomial}] = n \times E[\text{Bernoulli}] = 25 \times 0.15 = 3.75$
- **Variance:** $\text{Var}(\text{Binomial}) = n \times \text{Var}(\text{Bernoulli}) = 25 \times 0.1275 = 3.1875$

Tip

- Single inspection: 15% chance of finding a defect
- Batch inspection (25 items): Expect $3 - 4$ defective items typically Acceptable range: $2 - 6$ defective items would be within 1 standard deviation
- Red flag: Finding 7+ defective items might indicate process issues (beyond 2 σ)

Table 4: Summary of Bernoulli vs Binomial Distributions

Unable to display output for mime type(s): text/html

3 Optional: Conceptual Understanding - SOLUTIONS

❗ Important

Objective: Deepen understanding of key differences between probability distributions and their applications.

3.0.1 (a) Binomial vs. Geometric Distributions

Explain the key difference between a Binomial distribution and a Geometric distribution in terms of what they count.

Solution:

Key Difference: What We Count

Distribution	What We Count	Fixed Parameter	Variable
Binomial	Number of **successes**	Number of trials (n)	Number of successes
Geometric	Number of **trials**	Until first success	Number of trials

- **Binomial Distribution**: Counts the **number of successes** in a **fixed number** of trials
  - Example: "How many heads in 10 coin flips?"
  - We know we'll flip exactly 10 times, but don't know how many heads
- **Geometric Distribution**: Counts the **number of trials needed** to get the **first success**
  - Example: "How many coin flips until the first head?"
  - We know we'll get exactly 1 head, but don't know how many flips it takes

3.0.2 Visual Comparison : Binomial vs Geometric - Fundamental Difference

Unable to display output for mime type(s): text/html

Figure 12: Binomial vs Geometric: What They Count

Tip

Binomial: "How many successes in a fixed box of trials?"
Fixed trials, variable successes
Geometric: "How many attempts until first success?"
Fixed successes (1), variable trials

3.0.3 (b) Poisson vs. Binomial: When to Use Each

When would you use a Poisson distribution instead of a Binomial distribution?

Solution. Use Poisson when:

- Events occur over time or space at a constant rate
- The number of possible events is very large but the probability of each is very small
- We don't have a fixed number of trials Examples: arrivals, defects per unit area, accidents per day

Decision Framework: Poisson vs. Binomial

Criterion	Use Binomial	Use Poisson
Trials	Fixed number (n)	No fixed limit
Time/Space	Not the focus	Events over time/space
Probability	Moderate p	Very small p
Rate	Not applicable	Constant rate (λ)
Examples	Coin flips, surveys	Phone calls, defects

3.1 Comparative Examples: Poisson vs Binomial - Choosing the Right Distribution

`Unable to display output for mime type(s): text/html`

Figure 13: Poisson vs Binomial: When to Use Each

⚠ Warning

Common Mistake
- Don't use Poisson just because events are "rare." The key criteria are:
- No fixed number of trials
- Events over time/space
- Constant rate (λ)

A rare event in a fixed number of trials is still Binomial!

3.1.1 (c) Variance Maximization in Binomial Distribution

If $X \sim \text{Binomial}(n, p)$, under what conditions would the variance be maximized?

Solution. For a Binomial distribution: $\text{Var}(X) = np(1-p)$

For **fixed** n, variance is maximized when $p(1-p)$ is maximized.

Approach:

Taking the derivative with respect to p:

$\frac{d}{dp}[p(1-p)] = \frac{d}{dp}[p - p^2] = 1 - 2p$

Setting equal to zero:

$1 - 2p = 0 \quad \Longrightarrow \boxed{p = 0.5}$

Second derivative $= -2 < 0$, confirming this is a **maximum**.

The variance is maximized when $p = 0.5$ (fair coin scenario).

3.1.2 Visualization of Variance vs. Probability

`Unable to display output for mime type(s): text/html`

Figure 14: Binomial Variance Maximization: Effect of p

3.1.3 Intuitive Understanding of Variance Maximization

`Unable to display output for mime type(s): text/html`

Figure 15: Why p = 0.5 Maximizes Variance: Mathematical Intuition

> **!** Important
>
> **Key Insights**
> **Maximum:** $p = 0.5$ maximizes $p(1 - p)$ for any fixed n
> **Intuitive Explanation:** Maximum uncertainty occurs when success and failure are equally likely
> **Practical Meaning:** A fair coin (50-50) has the highest variability in outcomes
> Extremes: When p approaches 0 or 1, outcomes become predictable (low variance)