# Chapter 6: Discrete Random Variables
## STAT 2601 – Business Statistics

Esam Mahdi

School of Mathematics and Statistics
Carleton University

January 29, 2026

# Table of Contents

# What is a Discrete Random Variable?

## Definition

A **discrete random variable** is a variable whose values are countable and typically result from counting.

## Examples

- Number of customers arriving in an hour
- Number of defective items in a batch
- Number of emails received per day
- Number of successes in $n$ trials

## Random Variables

**Discrete:** Possible values can be counted or listed with discrete gaps in between.
**Continuous:** May assume any numerical value in one or more intervals. We cannot count or list the numbers in such an interval or continuum because they are infinitesimally close together.

# Probability Distribution Function (PDF)

## Definition

The **probability distribution** of a discrete random variable $X$ lists all possible values of $X$ and their corresponding probabilities (in a table, graph, or formula).

**Table**

| $x$ | $P(X = x)$ |
|---|---|
| $x_1$ | $P(X = x_1)$ |
| $x_2$ | $P(X = x_2)$ |
| $\vdots$ | $\vdots$ |
| $x_k$ | $P(X = x_k)$ |

**Formula**

$$P(X = x) = f(x)$$

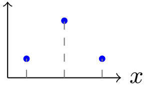
Example:

$$P(X = x) = C_x^n p^x (1-p)^{n-x}$$

(Binomial distribution)

**Graph**

$P(X = x)$




## Properties

For a valid probability distribution:

1. $0 \leq p(x_i) \leq 1$ for all $x_i$
2. $\sum p(x_i) = 1$

# Example: Probability Distribution Function

## Example

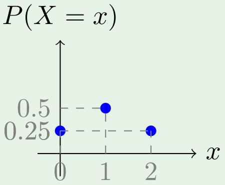
Let $X$ = number of heads in 2 coin tosses $\Rightarrow x = 0, 1, 2$. Note $S = \{HH, HT, TH, TT\}$.

**Table:**

| x | 0 | 1 | 2 |
|---|---|---|---|
| P(X=x) | 0.25 | 0.50 | 0.25 |

**Formula:** $P(X = x) = \frac{\binom{2}{x}}{4}$ or $P(X = x) = C_x^2 (0.5)^x (0.5)^{2-x}$

**Graph:**



# Expected Value (Mean)

**Definition**

The **expected value** $E(X)$ or $\mu$ is the weighted average of all possible values of $X$, weighted by their probabilities.

$$E(X) = \mu = \sum_i x_i \cdot p(x_i)$$

**Example**

Let $X$ = daily sales (in units) that has the following probability distribution:

| $x$ | 10 | 20 | 30 |
|---|---|---|---|
| $p(x)$ | 0.3 | 0.5 | 0.2 |

$E(X) = (10 \times 0.3) + (20 \times 0.5) + (30 \times 0.2) = 3 + 10 + 6 = 19$ units

# Variance and Standard Deviation

## Definition

**Variance** measures the spread or dispersion of a random variable around its mean.

$$\text{Variance: } V(X) = \sigma^2 = \sum (x_i - \mu)^2 \cdot p(x_i) = \sum_{\text{all } x} x^2 p(x) - \mu^2$$

$$\text{Standard Deviation: } SD(X) = \sigma = \sqrt{\sigma^2} = \sqrt{V(X)}$$

## Example

Using previous sales example ($\mu = 19$):

$$\begin{aligned}
\sigma^2 &= (10-19)^2(0.3) + (20-19)^2(0.5) + (30-19)^2(0.2) \\
&= 81(0.3) + 1(0.5) + 121(0.2) \\
&= 24.3 + 0.5 + 24.2 = 49 \\
\sigma &= \sqrt{49} = 7
\end{aligned}$$

## Example: Calculating Variance of Sales

A coffee shop tracks the number of specialty drinks sold per hour (X). The probability distribution is:

| **Drinks (x)** | 10 | 15 | 20 | 25 | 30 |
|---|---|---|---|---|---|
| **P(X = x)** | 0.1 | 0.2 | 0.4 | 0.2 | 0.1 |

Calculate the expected value ($\mu$) , the variance ($\sigma^2$), and the standard deviation ($\sigma$)

## Option 1: Direct variance formula using $\sigma^2 = \sum(x_i - \mu)^2 \cdot p(x_i)$

| $x$ | $p(x)$ | $x \cdot p(x)$ | $x - \mu$ | $(x-\mu)^2$ | $(x-\mu)^2 \cdot p(x)$ |
|---|---|---|---|---|---|
| 10 | 0.1 | $10 \times 0.1 = 1$ | $(10 - 20) = -10$ | 100 | $100 \times 0.1 = 10$ |
| 15 | 0.2 | $15 \times 0.2 = 3$ | $(15 - 20) = -5$ | 25 | $25 \times 0.2 = 5$ |
| 20 | 0.4 | $20 \times 0.4 = 8$ | $(20 - 20) = 0$ | 0 | $0 \times 0.4 = 0$ |
| 25 | 0.2 | $25 \times 0.2 = 5$ | $(25 - 20) = 5$ | 25 | $25 \times 0.2 = 5$ |
| 30 | 0.1 | $30 \times 0.1 = 3$ | $(30 - 20) = 10$ | 100 | $100 \times 0.1 = 10$ |
| Sum | | $\mu = \mathbf{20}$ | | | $\sigma^2 = \mathbf{30}$ |

$$\mu = 20, \quad , \sigma^2 = 30 \quad \Rightarrow \quad \sigma = \sqrt{30} = \mathbf{5.477} \text{ drinks}$$

## Option 2: Shortcut variance formula using $\sigma^2 = \sum_{\text{all } x} x^2 p(x) - \mu^2$

| $x$ | $p(x)$ | $x^2$ | $x^2 \cdot p(x)$ |
|---|---|---|---|
| 10 | 0.1 | 100 | $100 \times 0.1 = 10$ |
| 15 | 0.2 | 225 | $225 \times 0.2 = 45$ |
| 20 | 0.4 | 400 | $400 \times 0.4 = 160$ |
| 25 | 0.2 | 625 | $625 \times 0.2 = 125$ |
| 30 | 0.1 | 900 | $900 \times 0.1 = 90$ |
| Total | | | **430** |

Note that the expected value is

$$\mu = E(X) = \sum x \cdot p(x)$$
$$= (10 \times 0.1) + (15 \times 0.2) + (20 \times 0.4) + (25 \times 0.2) + (30 \times 0.1) = \mathbf{20} \text{ drinks}$$

$$\sum_{\text{all } x} x^2 p(x) = 430, \quad \mu^2 = (20)^2 = 400$$

Thus,

$$\sigma^2 = 430 - 400 = \mathbf{30} \quad \Rightarrow \quad \sigma = \sqrt{30} = \mathbf{5.477} \text{ drinks}$$

# Characteristics of Binomial Experiments

A Binomial experiment must satisfy:

1. Fixed number of trials ($n$)
2. Each trial has only two outcomes: **success** or **failure**
3. Constant probability of success ($p$).
4. Trials are independent

## Notation

- $n$: number of trials
- $x$: number of successes
- $p$: probability of success
- $q = 1 - p$: probability of failure

# Binomial Probability Formula

Let $X$ be a random variable representing the number of successes in $n$ trials that satisfies the characteristics of a **Binomial experiment**.

Then the probability mass function is given by:

$$P(X = x) = C_x^n p^x q^{n-x} \quad \text{for } x = 0, 1, 2, \ldots, n$$

where

$$C_x^n = \frac{n!}{x!(n-x)!}, \quad n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1$$

Note, $1! = 0! = 1$, where ! means **factorial**.

## Mean, Variance, and Standard Deviation of Binomial Distribution

$$\mu = E(X) = np$$

$$\sigma^2 = V(X) = npq$$

$$\sigma = SD(X) = \sqrt{npq}$$

## Note

$q = 1 - p$, where $p$ is the probability of success

# Binomial Example 1: Quality Control

**Example**

A factory produces light bulbs with 5% defect rate. In a sample of 20 bulbs:

- $X$ = number of defective light bulbs.
- $n = 20$, $p = 0.05$, $q = 0.95$

1. **Probability exactly 2 are defective:**

$$P(X = 2) = C_x^n p^x q^{n-x} = \frac{20!}{2!18!}(0.05)^2(0.95)^{18}$$
$$= 190 \times 0.0025 \times 0.3972 = 0.1887$$

2. **Expected number of defectives:**

$$\mu = np = 20 \times 0.05 = 1$$

3. **Standard deviation:**

$$\sigma = \sqrt{npq} = \sqrt{20 \times 0.05 \times 0.95} = \sqrt{0.95} = 0.9747$$

## Example: Marksman Target Practice

A marksman hits a target 80% of the time. He fires 6 shots at the target. Let $X$ represent the number of shots that the marksman hits the target.

1. State the name of the probability distribution of $X$. List the values of all relevant parameters.
2. What is the probability that exactly 4 shots hit the target?
3. What is the probability that more than 4 shots hit the target?
4. What is the probability that at least 2 shots hit the target?
5. What is the probability that at most 5 shots hit the target?
6. Would it be unusual that none of the shots hits the target?

**1. Binomial Distribution Parameters**

- $n = 6$ (number of trials)
- $p = 0.8$ (probability of success - hitting target)
- $q = 1 - p = 0.2$ (probability of failure - missing target)
- $X$ = number of shots that hit the target

$$X \text{ follows a Binomial distribution with parameters } n = 6, \text{ and } p = 0.8$$

$$\text{or} \qquad X \sim B(n = 6, p = 0.8)$$

# Example of Marksman Target Practice (cont'd):

## Binomial Probability Formula

$$P(X = x) = C_x^n p^x q^{n-x} \quad \text{for } x = 0, 1, 2, \ldots, n$$

**2. Probability of Exactly 4 Hits:**

$$P(X = 4) = \frac{6!}{4!(6-4)!}(0.8)^4(0.2)^2 = 15 \times 0.4096 \times 0.04 = \mathbf{0.24576}$$

**3. Probability of More Than 4 Hits:**

$$P(X > 4) = P(X = 5) + P(X = 6)$$

## $P(X = 5)$

$$\begin{aligned} P(X = 5) &= \frac{6!}{5!1!}(0.8)^5(0.2)^1 \\ &= 6 \times 0.32768 \times 0.2 \\ &= 0.393216 \end{aligned}$$

## $P(X = 6)$

$$\begin{aligned} P(X = 6) &= \frac{6!}{6!0!}(0.8)^6(0.2)^0 \\ &= 1 \times 0.262144 \times 1 \\ &= 0.262144 \end{aligned}$$

$$P(X > 4) = 0.393216 + 0.262144 \approx \mathbf{0.655}$$

# Example of Marksman Target Practice (cont'd):

**4. Probability of At Least 2 Hits:**

## What We Need to Calculate

At least 2 hits means: $X \geq 2$ or $X = 2, 3, 4, 5, 6$
Easier to use complement rule:

$$P(X \geq 2) = 1 - P(X < 2) = 1 - [P(X = 0) + P(X = 1)]$$

## $P(X = 0)$

$$\begin{aligned} P(X = 0) &= \frac{6!}{0!6!}(0.8)^0(0.2)^6 \\ &= 1 \times 1 \times 0.000064 \\ &= 0.000064 \end{aligned}$$

## $P(X = 1)$

$$\begin{aligned} P(X = 1) &= \frac{6!}{1!5!}(0.8)^1(0.2)^5 \\ &= 6 \times 0.80 \times 0.00032 \\ &= 0.001536 \end{aligned}$$

$$P(X \geq 2) = 1 - (0.000064 + 0.001536) = 1 - 0.0016 = \mathbf{0.9984}$$

# Example of Marksman Target Practice (cont'd):

**5. Probability of At Most 5 Hits:**

## What We Need to Calculate

At most 5 hits means: $X \leq 5$ or $X = 0, 1, 2, 3, 4, 5$
Easier to use complement rule:

$$P(X \leq 5) = 1 - P(X = 6)$$

## Recall $P(X = 6)$ from Part 2

$$P(X = 6) = 0.262144$$

Thus,

$$P(X \leq 5) = 1 - 0.262144 = \mathbf{0.737856}$$

## Alternative Method (not recommended!)

$$\begin{aligned} P(X \leq 5) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) \\ &= 0.000064 + 0.001536 + 0.01536 + 0.08192 + 0.24576 + 0.393216 \\ &= \mathbf{0.737856} \end{aligned}$$

# Example of Marksman Target Practice (cont'd):

**6. Check whether it would be unusual for none of the shots to hit the target:**

**Method 1:** Use the z-score

$$z = \frac{X - \mu_x}{\sigma_x},$$

- Expected value: $\mu = np = 6 \times 0.8 = 4.8$ hits
- Standard deviation: $\sigma = \sqrt{npq} = \sqrt{6 \times 0.8 \times 0.2} = \sqrt{0.96} = 0.9798$

z-score for 0 hits:

$$z = \frac{0 - 4.8}{0.9798} = -4.9$$

Since $z = -4.9$ is beyond the interval $(-3, 3)$, it would be unusual that no shot hits the target.

**Method 2:**

$$P(X = 0) = 0.000064 = 0.0064\%$$

Since the probability $0.000064$ is very tiny (i.e., chance is almost zero), it would be extremely unusual that no shot hits the target.

# Poisson Distribution

**Characteristics of Poisson Distribution:** Poisson distributions are used to model the events occurring over time or space when:

- Events occur independently
- Average rate ($\mu$) is constant
- Two events cannot occur at exactly the same instant/position

## Common Applications

- Number of text messages a person receives on their cell phone per 30 minutes
- Number of patients arrive at a hospital emergency room per hour
- Number of insurance claims per day
- Number of cars served at the gas station in 24-hour period
- Number of defects per square meter
- Number of typographical errors in a given book per 10 pages

**Poisson Probability Formula**

$$P(X = x) = \frac{e^{-\mu}\mu^x}{x!} \quad \text{for } x = 0, 1, 2, \ldots$$

## Mean, Variance, and Standard Deviation

$$E(X) = \mu, \qquad V(X) = \mu, \qquad \sigma = \sqrt{\mu}$$

# Poisson Example: Customer Arrivals

## Example

A bank observes an average of 3 customers arrive every 15 minutes during lunch hour.

Let $X$ represent the number of customers that arrive in the next 15 minutes. Thus, $X$ follows a Poisson distribution with a mean of $\mu = 3$, or $X \sim \text{Poisson}(\mu = 3)$

1. **Probability exactly 2 customers in 15 minutes:**

$$P(X = 2) = \frac{e^{-3}3^2}{2!} = \frac{0.0498 \times 9}{2} = 0.2241$$

2. **Probability at most 2 customers:**

$$\begin{aligned} P(X \le 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \frac{e^{-3}3^0}{0!} + \frac{e^{-3}3^1}{1!} + \frac{e^{-3}3^2}{2!} \\ &= 0.0498 + 0.1494 + 0.2241 = 0.4233 \end{aligned}$$

3. **Standard deviation:**

$$\sigma = \sqrt{\mu} = \sqrt{3} = 1.732$$

# Poisson Example: Traffic Accidents

The average number of traffic accidents on a certain section of highway is 2 per week.

1. Find the probability of exactly one accident during a one-week period (Same Unit of Time)
2. Find the probability of at most three accidents during a two-week (Different Unit of Time)
3. Would it be unusual that 8 or more accidents happen per week? (Same Unit of Time)

**Part 1:**

$$\mu = 2/\text{week}$$

$$P(X = 1) = \frac{e^{-2}2^1}{1!} = 0.2707$$

**Part 2:**

$$\mu = 2/\text{week, adjusted } \mu = 4/2\text{-week} \quad \Rightarrow X \sim \text{Poisson}(\mu = 4)$$

$$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$$

$$P(X \leq 3) = \frac{e^{-4}4^0}{0!} + \frac{e^{-4}4^1}{1!} + \frac{e^{-4}4^2}{2!} + \frac{e^{-4}4^3}{3!} = 0.4335$$

**Part 3:** **Method 1:**

$$z = \frac{X - \mu_x}{\sigma_x} = \frac{8 - 2}{\sqrt{2}} = 4.24$$

Since $z = 4.24$ is beyond the interval $(-3, 3)$, it would be unusual that 8 or more accidents happen per week

**Method 2:**

$$P(X \geq 8) = 1 - P(X < 8) = 1 - P(X \leq 7) \approx 0.001$$

Since the probability 0.001 is very tiny (i.e., chance is very thin), it would be unusual that 8 or more accidents happen per week.

## Example

A statistics instructor observes that the number of typographical errors is Poisson distributed with a mean of 1.5 per **100 pages**.

(i) What is the probability that there are **no typos** in a new book of **100 pages**? (Same Unit of Space)

$$\mu = 1.5/100\text{-page}$$

$$P(X = 0) = \frac{e^{-1.5}1.5^0}{0!} = 0.2231$$

(ii) Suppose that the instructor has just received a copy of a new statistics book of **400 pages**. Find the probability that there are five or fewer typos. (Different Unit of Space)

$$\mu = 1.5/100\text{-page, adjusted } \mu = 6/400\text{-page} \quad \Rightarrow X \sim \text{Poisson}(\mu = 6)$$

$$P(X \leq 5) = P(X = 0) + P(X = 1) + \cdots + P(X = 5)$$

$$P(X \leq 5) = \frac{e^{-6}6^0}{0!} + \frac{e^{-6}6^1}{1!} + \cdots + \frac{e^{-6}6^5}{5!} = 0.45$$

## Comparison of Distributions

| Feature | Binomial | Poisson |
|---|---|---|
| Type of variable | Count of successes | Count of occurrences |
| Number of trials | Fixed ($n$) | Not fixed |
| Probability of success | Constant ($p$) | $\mu$ = average rate |
| Possible values | $0, 1, 2, \ldots, n$ | $0, 1, 2, \ldots$ |
| Mean | $\mu = np$ | $\mu$ |
| Variance | $\sigma^2 = npq$ | $\sigma^2 = \mu$ |

## Key Formulas

**General Discrete Random Variable:**

$$\mu = E(X) = \sum x \cdot p(x) \quad \sigma^2 = \sum (x-\mu)^2 p(x)$$

**Binomial Distribution:**

$$P(X = x) = C_x^n p^x q^{n-x} \quad \mu = np,\ \sigma = \sqrt{npq}$$

where $C_x^n = \dfrac{n!}{x!(n-x)!}$

**Poisson Distribution:**

$$P(X = x) = \frac{e^{-\mu}\mu^x}{x!} \quad \mu = \mu,\ \sigma = \sqrt{\mu}$$