Chapter 5 Discrete Probability Distributions
Binomial Probability Distributions (5.2)

Objectives

In this lesson, we will learn how to

- identify whether a situation has a binomial probability distribution;
- use the notation n, x, p, q, and $P(x)$ correctly;
- find binomial probabilities using the formula, technology, or a table;
- find the mean, variance, and standard deviation of a binomial distribution;
- decide whether a result is significantly low or significantly high.

1 Key Concept

The focus of this lesson is the binomial probability distribution.

A binomial distribution is used when we repeat the same type of trial several times and count how many times a particular outcome occurs.

Main Ideas

- A binomial experiment has a fixed number of trials.
- Each trial has only two categories: success and failure.
- We can find probabilities for exactly x successes.
- We can use probabilities or the range rule of thumb to decide whether a result is unusual.

2 Binomial Probability Distribution

A binomial probability distribution results from a procedure that satisfies all four requirements below.

Requirements for a Binomial Distribution

1. The procedure has a fixed number of trials.
2. The trials are independent.
3. Each trial has exactly two possible categories: success and failure.
4. The probability of success remains the same for every trial.

Important Note

The word success does not necessarily mean something good. It simply means the outcome we are counting.
For example, if we are counting people who are cashless, then ``cashless'' is called success.

3 Notation for Binomial Distributions

Notation

Symbol	Meaning
S	success
F	failure
p	P(S), probability of success
q	P(F)=1-p, probability of failure
n	fixed number of trials
x	specific number of successes among the n trials
$P(x)$	probability of getting exactly x successes

Consistency Check

Always make sure that x and p refer to the same definition of success. If x counts cashless adults, then p must be the probability that one selected adult is cashless.

Example 1: When an adult smartphone owner is randomly selected, suppose there is a 0.05 probability that this person is cashless, meaning that the person never carries cash.
We randomly select 10 adults.

(a) Does this procedure result in a binomial distribution?

(b) If it is binomial, identify n, x, p, and q for finding the probability that exactly 2 adults are cashless.

Solution

Part (a): Yes, this is a binomial distribution.

- The number of trials is fixed: $n = 10$.
- The trials are independent because each adult is randomly selected with replacement or from a very large population.
- Each adult is either cashless or not cashless.

- The probability of being cashless remains 0.05 for each selected adult.

Part (b):

Symbol	Meaning in the cashless example
n	10 adults
x	2 cashless adults
p	0.05 probability of cashless
q	0.95 probability of not cashless

Therefore,

$$n = 10, \qquad x = 2, \qquad p = 0.05, \qquad q = 0.95.$$

4 Treating Dependent Events as Independent

Sometimes sampling is done without replacement, so the selections are technically dependent.

However, if the sample is small compared with the population, we often treat the selections as independent.

5% Guideline for Cumbersome Calculations

When sampling without replacement, if the sample size is no more than 5% of the population size, we may treat the selections as independent.

5 Binomial Probability Formula

The binomial probability formula gives the probability of exactly x successes in n trials.

Binomial Probability Formula

$$P(x) = \frac{n!}{(n-x)!x!} p^x q^{n-x}$$

where

- n is the number of trials,
- x is the number of successes,
- p is the probability of success,
- $q = 1 - p$ is the probability of failure.

The expression

$$\frac{n!}{(n-x)!x!}$$

counts the number of different ways to get exactly x successes among n trials.

Example 2: Given that there is a 0.05 probability that a randomly selected adult smartphone owner is cashless, find the probability that when 10 adults are randomly selected, exactly 2 are cashless.
Use

$$n = 10, \qquad x = 2, \qquad p = 0.05, \qquad q = 0.95.$$

Solution

Use the binomial probability formula:

$$P(x) = \frac{n!}{(n-x)!x!} p^x q^{n-x}.$$

Substitute the given values:

$$P(2) = \frac{10!}{(10-2)!2!} (0.05)^2 (0.95)^{10-2}.$$

$$P(2) = \frac{10!}{8!2!}(0.05)^2(0.95)^8.$$

$$P(2) = 45(0.0025)(0.6634).$$

$$P(2) \approx 0.0746.$$

Therefore, the probability that exactly 2 of the 10 adults are cashless is

$$\boxed{0.0746}.$$

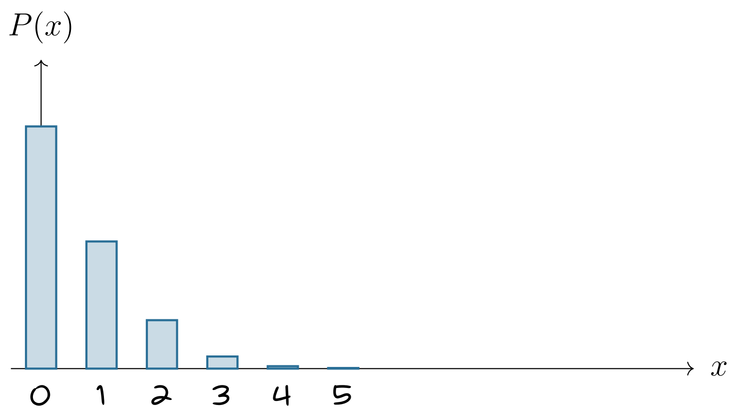
6 Methods for Finding Binomial Probabilities

Three Common Methods

1. Formula: Use the binomial probability formula.
2. Technology: Use StatCrunch, calculator, Excel, or other software.
3. Table: Use a binomial probability table when available.

Technology is usually the easiest method, especially when n is large or when we need a cumulative probability such as

$$P(x \geq 252) \quad \text{or} \quad P(x \leq 3).$$



Shape of a binomial distribution with small p

Example 3: Between 1974 and 2011, there were 460 NFL football games decided in overtime. Of these, 252 were won by the team that won the overtime coin toss.
Assume wins and losses are equally likely, so

$$n = 460, \qquad p = 0.5, \qquad q = 0.5.$$

Is 252 wins significantly high?

Solution

We want the probability of getting 252 or more wins:

$$P(x \geq 252).$$

This would require adding many binomial probabilities:

$$P(252) + P(253) + \cdots + P(460).$$

This is not practical by hand, so technology is recommended.
Using technology,

$$P(x \geq 252) \approx 0.0224.$$

Since

$$0.0224 \leq 0.05,$$

this result is significantly high.
Therefore, 252 wins is unlikely to occur by random chance alone. This suggests that the team winning the overtime coin toss may have an advantage.

7 Using a Binomial Probability Table

A binomial probability table can be used for selected values of n and p.

Using Table A-1

To use a binomial probability table:

1. Locate the value of n.
2. Locate the row for the desired value of x.
3. Move across to the column for the desired value of p.
4. The table entry is $P(x)$.

Example 4: Based on a Gallup poll, suppose 5% of U.S. adults are vegetarians. If we randomly select 5 adults, find the following probabilities using a binomial table.

(a) The probability that exactly 2 of the 5 adults are vegetarians.

(b) The probability that there are fewer than 3 vegetarians.

Solution

Here,
$$n = 5, \qquad p = 0.05.$$
Part (a):
From the binomial table,
$$P(2) = 0.021.$$

Part (b):
``Fewer than 3'' means 0, 1, or 2 vegetarians.

$$P(x < 3) = P(0) + P(1) + P(2).$$

Using the table values,

$$P(x < 3) \approx 0.774 + 0.204 + 0.021 = 0.999.$$

So the probability of fewer than 3 vegetarians among 5 randomly selected adults is about

$$\boxed{0.999}.$$

8 Mean, Variance, and Standard Deviation

For a binomial distribution, we can find the mean, variance, and standard deviation quickly.

Parameters of a Binomial Distribution

$$\mu = np$$

$$\sigma^2 = npq$$

$$\sigma = \sqrt{npq}$$

- μ is the mean number of successes.
- σ^2 is the variance.
- σ is the standard deviation.

9 Range Rule of Thumb

Identifying Significant Values

- Significantly low values are
$$x \le \mu - 2\sigma.$$
- Significantly high values are
$$x \ge \mu + 2\sigma.$$
- Values between $\mu - 2\sigma$ and $\mu + 2\sigma$ are not considered significant by this rule.

Example 5: For the NFL overtime example, suppose

$$n = 460, \qquad p = 0.5, \qquad q = 0.5.$$

(a) Find the mean and standard deviation for the number of wins.

(b) Use the range rule of thumb to find significantly low and significantly high values.

(c) Is 252 wins significantly high?

Solution

Part (a):

$$\mu = np = (460)(0.5) = 230.0.$$

$$\sigma = \sqrt{npq} = \sqrt{(460)(0.5)(0.5)} = \sqrt{115} \approx 10.7.$$

So the mean is 230.0 wins and the standard deviation is about 10.7 wins.
Part (b):

$$\mu - 2\sigma = 230.0 - 2(10.7) = 208.6.$$

$$\mu + 2\sigma = 230.0 + 2(10.7) = 251.4.$$

Significantly low values are 208.6 or below.
Significantly high values are 251.4 or above.
Part (c):
Since

$$252 \geq 251.4,$$

252 wins is significantly high.

10 Using Probabilities to Determine Significance

Probability Method

- x successes is significantly high if

$$P(x \text{ or more}) \leq 0.05.$$

- x successes is significantly low if

$$P(x \text{ or fewer}) \leq 0.05.$$

The cutoff 0.05 is commonly used, but it is not absolutely rigid. Sometimes 0.01 or another value may be used.

11 Rationale for the Binomial Formula

The binomial formula has two main parts.

Why the Formula Works

$$P(x) = \frac{n!}{(n-x)!x!} p^x q^{n-x}$$

- $\dfrac{n!}{(n-x)!x!}$ counts the number of ways to arrange x successes among n trials.
- $p^x q^{n-x}$ gives the probability of one particular arrangement with x successes and $n-x$ failures.

For example, if $n = 10$ and $x = 2$, then

$$\frac{10!}{8!2!} = 45.$$

This means there are 45 different ways to get exactly 2 successes in 10 trials.

Summary

Key Ideas

- A binomial distribution requires fixed trials, independence, two outcome categories, and constant probability of success.
- The probability of failure is $q = 1 - p$.
- The binomial probability formula is

$$P(x) = \frac{n!}{(n-x)!x!} p^x q^{n-x}.$$

- Technology is very helpful for cumulative probabilities and large values of n.
- For a binomial distribution,

$$\mu = np, \qquad \sigma^2 = npq, \qquad \sigma = \sqrt{npq}.$$

- A result can be judged significant using probabilities or the range rule of thumb.