Chapter 6 Normal Probability Distributions
Normal as Approximation to Binomial (6.6)

Objectives

In this lesson, we will learn how to

- decide when a normal distribution can be used to approximate a binomial distribution;
- find the mean and standard deviation for the normal approximation;
- use a continuity correction when changing from discrete values to a continuous normal curve;
- approximate binomial probabilities such as ``at least,'' ``more than,'' ``at most,'' ``fewer than,'' and ``exactly'';
- interpret the probability result in context.

1 Key Idea

A binomial probability distribution is discrete because the random variable x counts whole-number successes.

A normal distribution is continuous because values can fall anywhere on a continuous scale.

Even though these are different types of distributions, a binomial distribution can often be approximated by a normal distribution when the sample size is large enough.

When Can We Use a Normal Approximation?

Given a binomial distribution with parameters n, p, and $q = 1 - p$, we can use a normal approximation if

$$np \geq 5 \qquad \text{and} \qquad nq \geq 5.$$

2 Review of Binomial Notation

Binomial Notation

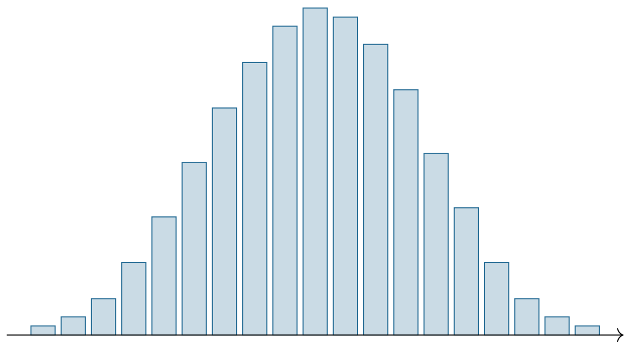
- n: fixed number of trials
- x: number of successes in n trials
- p: probability of success on one trial
- q: probability of failure on one trial, so $q = 1 - p$

A binomial distribution is based on trials that are independent, have two possible categories, and have the same probability of success for every trial.

3 Why a Normal Approximation Works

When n is large enough and p is not too close to 0 or 1, a binomial histogram often becomes approximately bell-shaped.

Binomial histogram can look bell-shaped when conditions are met



This bell shape suggests that a normal curve can be used to approximate areas from the binomial distribution.

4 Mean and Standard Deviation

Normal Approximation to a Binomial Distribution

If $np \geq 5$ and $nq \geq 5$, then the binomial random variable x can be approximated by a normal distribution with

$$\mu = np$$

and

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

Important Reminder

The standard deviation is $\sqrt{npq}$, not npq.

5 Continuity Correction

Because the binomial distribution uses whole numbers and the normal distribution is continuous, we need a continuity correction.

Continuity Correction

When using the normal distribution to approximate a binomial probability, represent a whole number x by the interval

$$x - 0.5 \quad \text{to} \quad x + 0.5.$$

Binomial Statement	Normal Approximation Region
At least x	To the right of $x - 0.5$
More than x	To the right of $x + 0.5$
At most x	To the left of $x + 0.5$
Fewer than x	To the left of $x - 0.5$
Exactly x	Between $x - 0.5$ and $x + 0.5$

6 General Procedure

Using a Normal Distribution to Approximate a Binomial Distribution

1. Check that $np \geq 5$ and $nq \geq 5$.
2. Find
$$\mu = np, \qquad \sigma = \sqrt{npq}.$$
3. Identify the relevant whole number x.
4. Apply the correct continuity correction.
5. Convert the corrected boundary value to a z score:
$$z = \frac{x - \mu}{\sigma}.$$
6. Use technology or Table A-2 to find the desired area.

Example 1: In one of Mendel's famous hybridization experiments, he expected that among 580 offspring peas, 25% would be yellow. He actually got 152 yellow peas.
Assume $n = 580$ and $p = 0.25$. Find

$$P(\text{at least } 152 \text{ yellow peas}).$$

Is 152 yellow peas significantly high?

Solution

First check the requirements.

$$np = (580)(0.25) = 145$$

and

$$nq = (580)(0.75) = 435.$$

Both are at least 5, so the normal approximation is appropriate.
Now find the mean and standard deviation.

$$\mu = np = 145$$

$$\sigma = \sqrt{npq} = \sqrt{(580)(0.25)(0.75)} = 10.4283.$$

We want $P(x \geq 152)$, so we use the continuity correction:

$$\text{at least } 152 \quad \Longrightarrow \quad \text{right of } 151.5.$$

Now convert 151.5 to a z score:

$$z = \frac{151.5 - 145}{10.4283} = 0.62.$$

Using Table A-2,

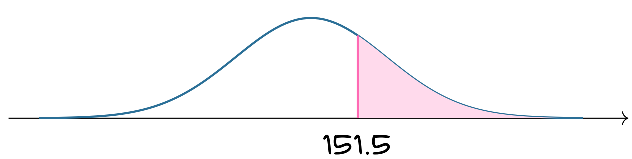
$$P(z < 0.62) = 0.7324.$$

Therefore,

$$P(x \geq 152) \approx 1 - 0.7324 = 0.2676.$$

Using technology gives approximately 0.2665.
Since this probability is much greater than 0.05, 152 yellow peas is not significantly high. Mendel's result does not contradict the expected yellow rate of 25%.

Area to the right represents 152 or more



Example 2: For a binomial distribution, explain how each statement involving $x = 152$ should be represented on a continuous normal curve.

(a) at least 152

(b) more than 152

(c) at most 152

(d) fewer than 152

(e) exactly 152

Solution

Use the interval from $x - 0.5$ to $x + 0.5$, so the whole number 152 is represented by the interval from 151.5 to 152.5.

(a) At least 152: area to the right of 151.5.

(b) More than 152: area to the right of 152.5.

(c) At most 152: area to the left of 152.5.

(d) Fewer than 152: area to the left of 151.5.

(e) Exactly 152: area between 151.5 and 152.5.

Example 3: Using the same information from the previous example, approximate the probability of getting exactly 152 yellow peas among the 580 offspring peas.
That is, find

$$P(\text{exactly } 152 \text{ yellow peas}).$$

Solution

For exactly 152, use the continuity correction:

$$151.5 < x < 152.5.$$

We already have

$$\mu = 145, \qquad \sigma = 10.4283.$$

Convert both boundary values to z scores.

$$z_1 = \frac{151.5 - 145}{10.4283} = 0.62$$

$$z_2 = \frac{152.5 - 145}{10.4283} = 0.72$$

Using Table A-2,

$$P(z < 0.62) = 0.7324$$

and

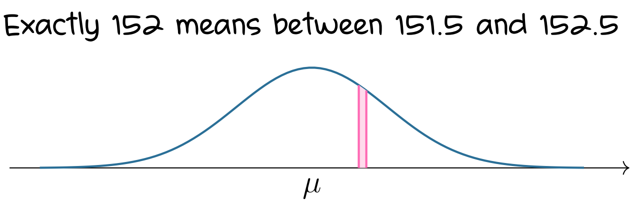
$$P(z < 0.72) = 0.7642.$$

Therefore,

$$P(151.5 < x < 152.5) = 0.7642 - 0.7324 = 0.0318.$$

So the approximate probability of exactly 152 yellow peas is

$$0.0318.$$



7 Which Probability Is Relevant?

Important Statistical Thinking

To decide whether an observed result is significantly high, we do not usually look at the probability of getting exactly that result.
Instead, we look at the probability of getting that result or something more extreme.

For Mendel's example:

- The probability of exactly 152 yellow peas is about 0.0318.
- But the relevant probability for deciding whether 152 is significantly high is

$$P(152 \text{ or more yellow peas}) \approx 0.2665.$$

Because 0.2665 is not small, the result is not significantly high.

8 Technology Note

Today, technology can often compute binomial probabilities directly.

Technology vs. Normal Approximation

The normal approximation is still important because:

- it helps us understand why binomial distributions can look normal;
- it provides a practical method when exact binomial calculations are difficult;
- it is used in later statistical methods involving proportions.

Summary

Key Ideas

- A binomial distribution may be approximated by a normal distribution when $np \geq 5$ and $nq \geq 5$.
- Use
$$\mu = np, \qquad \sigma = \sqrt{npq}.$$
- A continuity correction is needed because the binomial distribution is discrete and the normal distribution is continuous.
- ``At least x'' becomes area to the right of $x - 0.5$.
- ``Exactly x'' becomes area between $x - 0.5$ and $x + 0.5$.
- For significance, use the probability of the observed result or something more extreme.