Chapter 13 Concepts

- The Binomial Setting and Binomial Distributions
- Binomial Distributions in Statistical Sampling
- Binomial Probabilities
- Binomial Mean and Standard Deviation
- The Normal Approximation to Binomial Distributions
Chapter 13 Objectives

- Describe the binomial setting
- Describe binomial distributions
- Calculate binomial probabilities
- Calculate and interpret the binomial mean and standard deviation
- Calculate the Normal approximation to binomial distributions
The Binomial Setting

When the same chance process is repeated several times, we are often interested in whether a particular outcome does or doesn’t happen on each repetition. In some cases, the number of repeated trials is fixed in advance and we are interested in the number of times a particular event (called a “success”) occurs.

A binomial setting arises when we perform several independent trials of the same chance process and record the number of times that a particular outcome occurs. The four conditions for a binomial setting are:

- **Binary?** The possible outcomes of each trial can be classified as “success” or “failure.”
- **Independent?** Trials must be independent; that is, knowing the result of one trial must not have any effect on the result of any other trial.
- **Number?** The number of trials $n$ of the chance process must be fixed in advance.
- **Success?** On each trial, the probability $p$ of success must be the same.
Consider tossing a coin \( n \) times. Each toss gives either heads or tails. Knowing the outcome of one toss does not change the probability of an outcome on any other toss. If we define heads as a success, then \( p \) is the probability of a head and is 0.5 on any toss.

The number of heads in \( n \) tosses is a binomial random variable \( X \). The probability distribution of \( X \) is called a **binomial distribution**.

**Binomial Distribution**
The count \( X \) of successes in a binomial setting has the binomial distribution with parameters \( n \) and \( p \), where \( n \) is the number of trials of the chance process and \( p \) is the probability of a success on any one trial. The possible values of \( X \) are the whole numbers from 0 to \( n \).

**Note:** Not all counts have binomial distributions; be sure to check the conditions for a binomial setting and make sure you’re being asked to count the number of successes in a certain number of trials!
Binomial Distributions in Statistical Sampling

The binomial distributions are important in statistics when we want to make inferences about the proportion $p$ of successes in a population.

Suppose 10% of CDs have defective copy-protection schemes that can harm computers. A music distributor inspects an SRS of 10 CDs from a shipment of 10,000. Let $X = \text{number of defective CDs}$. **What is $P(X = 0)$? Note, this is not quite a binomial setting.** Why?

The actual probability is

$$P(\text{no defectives}) = \frac{9000 \times 8999 \times 8998 \times \ldots \times 8991}{10000 \times 9999 \times 9998 \times \ldots \times 9991} = 0.3485$$

**Sampling Distribution of a Count**

Choose an SRS of size $n$ from a population with proportion $p$ of successes. When the population is much larger than the sample, the count $X$ of successes in the sample has approximately the binomial distribution with parameters $n$ and $p$.

Using the binomial distribution,

$$P(X = 0) = \binom{10}{0} \frac{(0.10)^0(0.90)^{10}}{0} = 0.3487$$
Binomial Probability

We can find a formula for the probability that a binomial random variable takes any value by adding probabilities for the different ways of getting exactly that many successes in \(n\) observations.

The number of ways of arranging \(k\) successes among \(n\) observations is given by the binomial coefficient

\[
\binom{n}{k} = \frac{n!}{k!(n-k)!}
\]

for \(k = 0, 1, 2, \ldots, n\).

**Note:** \(n! = n(n-1)(n-2)\ldots(3)(2)(1)\) and \(0! = 1\).
Binomial Probability

The binomial coefficient counts the number of different ways in which \( k \) successes can be arranged among \( n \) trials. The binomial probability \( P(X = k) \) is this count multiplied by the probability of any one specific arrangement of the \( k \) successes.

If \( X \) has the binomial distribution with \( n \) trials and probability \( p \) of success on each trial, the possible values of \( X \) are 0, 1, 2, …, \( n \). If \( k \) is any one of these values,

\[
P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}
\]
Each child of a particular pair of parents has probability 0.25 of having blood type O. Suppose the parents have five children.

(a) **Find the probability that exactly three of the children have type O blood.**

Let $X =$ the number of children with type O blood. We know $X$ has a binomial distribution with $n = 5$ and $p = 0.25$.

$$P(X = 3) = \frac{5}{3} \frac{(0.25)^3(0.75)^2}{(0.25)^3(0.75)^2} = 10(0.25)^3(0.75)^2 = 0.08789$$

(b) **Should the parents be surprised if more than three of their children have type O blood?**

$$P(X > 3) = P(X = 4) + P(X = 5)$$

$$= \frac{5}{4} (0.25)^4 (0.75)^1 + \frac{5}{5} (0.25)^5 (0.75)^0$$

$$= 5(0.25)^4 (0.75)^1 + 1(0.25)^5 (0.75)^0$$

$$= 0.01465 + 0.00098 = 0.01563$$

Since there is only a 1.5% chance that more than three children out of five would have Type O blood, the parents should be surprised!
Binomial Mean and Standard Deviation

If a count $X$ has the binomial distribution based on $n$ observations with probability $p$ of success, what is its mean $\mu$? In general, the mean of a binomial distribution should be $\mu = np$. Here are the facts:

**Mean and Standard Deviation of a Binomial Random Variable**

If a count $X$ has the binomial distribution with number of trials $n$ and probability of success $p$, the **mean** and **standard deviation** of $X$ are:

$$
\mu_X = np
$$

$$
\sigma_X = \sqrt{np(1-p)}
$$

**Note:** These formulas work ONLY for binomial distributions. They can’t be used for other distributions!
As $n$ gets larger, something interesting happens to the shape of a binomial distribution.

Suppose that $X$ has the binomial distribution with $n$ trials and success probability $p$. When $n$ is large, the distribution of $X$ is approximately Normal with mean and standard deviation

$$\mu_X = np$$

$$\sigma_X = \sqrt{np(1 - p)}$$

As a rule of thumb, we will use the Normal approximation when $n$ is so large that $np \geq 10$ and $n(1 - p) \geq 10$. 

Normal Approximation for Binomial Distributions
Example

Sample surveys show that fewer people enjoy shopping than in the past. A survey asked a nationwide random sample of 2500 adults if they agreed or disagreed that “I like buying new clothes, but shopping is often frustrating and time-consuming.” Suppose that exactly 60% of all adult U.S. residents would say “Agree” if asked the same question. Let $X$ = the number in the sample who agree. Estimate the probability that 1520 or more of the sample agree.

1) Verify that $X$ is approximately a binomial random variable.

B: Success = agree, Failure = don’t agree
I: Because the population of U.S. adults is greater than 25,000, it is reasonable to assume the sampling without replacement condition is met.
N: $n = 2500$ trials of the chance process
S: The probability of selecting an adult who agrees is $p = 0.60$.

2) Check the conditions for using a Normal approximation.

Since $np = 2500(0.60) = 1500$ and $n(1 - p) = 2500(0.40) = 1000$ are both at least 10, we may use the Normal approximation.

3) Calculate $P(X \geq 1520)$ using a Normal approximation.

$$
\mu = np = 2500(0.60) = 1500
\sigma = \sqrt{np(1 - p)} = \sqrt{2500(0.60)(0.40)} = 24.49
$$

$$
z = \frac{1520 - 1500}{24.49} = 0.82
$$

$$
P(X \geq 1520) = P(Z \geq 0.82) \approx 0.7939 \approx 0.2061
$$
Chapter 13 Objectives Review

- Describe the binomial setting
- Describe binomial distributions
- Calculate binomial probabilities
- Calculate and interpret the binomial mean and standard deviation
- Calculate the Normal approximation to binomial distributions