5.2 The Binomial Distribution

Dr. Lauren Perry

Goals

1. Express cumulative probabilities using probability notation.
2. Calculate binomial probabilities.

Trials

- Think back to replication in an experiment.
- Each replication is what we call a **trial**.
- We will consider a setting where each trial has two possible outcomes.

Example: suppose you want to know if a coin is fair (both sides equally likely).

- You might flip the coin 100 times (thus running 100 trials).
- Each trial is a flip of the coin with two possible outcomes: heads or tails.

Factorials

The product of the first k positive integers $(1, 2, 3, \dots)$ is called **k-factorial**, denoted $k!$:

$$k! = k \times (k-1) \times \cdots \times 3 \times 2 \times 1$$

We define $0! = 1$.

Example: $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$

The Binomial Coefficient

If n is a positive integer $(1, 2, 3, \ldots)$ and x is a nonnegative integer $(0, 1, 2, \ldots)$ with $x \leq n$, the **binomial coefficient** is

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

Example:

$$\binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{(2 \times 1)(3 \times 2 \times 1)}$$

Simplifying Binomial Coefficients

What?

$$20! = 2,432,902,008,176,640,000$$

Your calculator cannot.

Simplifying Binomial Coefficients

Example:

$$\binom{20}{17} = \frac{20 \times 19 \times 18 \times 17 \times 16 \times \cdots \times 3 \times 2 \times 1}{(17 \times 16 \times \cdots \times 3 \times 2 \times 1)(3 \times 2 \times 1)}$$

Bernoulli Trials

Bernoulli trials are repeated trials of an experiment where:

1. Each trial has two possible outcomes: success and failure.
2. Trials are independent.
3. The probability of success (the **success probability**) p remains the same from one trial to the next:

$$P(X = \text{success}) = p$$

The Binomial Distribution

The **binomial distribution** is the probability distribution for the number of successes in a sequence of Bernoulli trials.

The Binomial Probability Formula

Let x denote the total number of successes in n Bernoulli trials with success probability p. The probability distribution of the random variable X is given by

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, 2, \ldots, n$$

The random variable X is called a **binomial random variable** and is said to have the **binomial distribution**. Because n and p fully define this distribution, they are called the distribution's **parameters**.

To find a binomial probability formula:

Check assumptions.

Exactly n trials to be performed.

Two possible outcomes for each trial.

Trials are independent (each trial does not impact the result of the next)

Success probability p remains the same from trial to trial.

Identify a "success". Generally, this is whichever of the two possible outcomes we are most interested in.

Determine the success probability p.

Determine n, the number of trials.

Plug n and p into the binomial distribution formula.

Probabilities with Inequalities

Consider $P(X \leq x)$.

- We can rewrite this using concepts from the previous chapter

$$P(X \leq k) = P(X = k \text{ or } X = k-1 \text{ or } \dots \text{ or } X = 1 \text{ or } X = 0)$$

- Since X is a discrete random variable, each possible value is *disjoint*.
- We can use this!

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

Example Rewrite $P(X \leq 3)$.

Probabilities with Inequalities

We can also extend this concept to work with probabilities like $P(a < X \leq b)$.

Example: $P(2 < X \leq 5)$ What values make up the event of interest?

The Shape of a Binomial Distribution

The shape of a binomial distribution is determined by the success probability:

- If $p \approx 0.5$, the distribution is approximately symmetric.
- If $p < 0.5$, the distribution is right-skewed.
- If $p > 0.5$, the distribution is left-skewed.

Mean and Variance

- The mean is $\mu = np$.
- The variance is $\sigma^2 = np(1 - p)$.

Checkpoint

Suppose 38.4% of people have dogs. We will take a random sample of 10 people and ask whether they have a dog.

1. Use the binomial distribution to find the probability that exactly 2 of the people in our sample have a dog.
2. Find the probability that between 3 and 5 people (inclusive) have a dog.