Bookmark this page

# Binomial Random Variables: Introduction

**Learning Objective: Fit the binomial model when appropriate, and use it to perform simple calculations.**

## Binomial Random Variables

So far, in our discussion about discrete random variables, we have been introduced to:

1. The probability distribution, which tells us which values a variable takes, and how often it takes them.
2. The mean of the random variable, which tells us the long-run average value that the random variable takes.
3. The standard deviation of the random variable, which tells us a typical (or long-run average) distance between the mean of the random variable and the values it takes.

We will now introduce a special class of discrete random variables that are very common, because as you'll see, they will come up in many situations—**binomial random variables.**

Here's how we'll present this material. First, we'll explain what kind of random experiments give rise to a binomial random variable, and how the binomial random variable is defined in those types of experiments.

We'll then present the probability distribution of the binomial random variable, which will be presented as a formula (which, as you remember, is one of the three ways in which a probability distribution of a discrete random variable can be presented), and explain why the formula makes sense. We'll conclude our discussion by presenting the mean and standard deviation of the binomial random variable.

As we just mentioned, we'll start by describing what kind of random experiments give rise to a binomial random variable. We'll call this type of random experiment a "binomial experiment."

## Binomial Experiment

Binomial experiments are random experiments that consist of a fixed number of repeated trials, like tossing a coin 10 times, randomly choosing 10 people, rolling a die 5 times, etc. These trials, however, need to be independent in the sense that the outcome in one trial has no effect on the outcome in other trials. In each of these repeated trials there is one outcome that is of interest to us (we call this outcome "success"), and each of the trials is identical in the sense that the probability that the trial will end in a "success" is the same in each of the trials. So for example, if our experiment is tossing a coin 10 times, and we are interested in the outcome "heads" (our "success"), then this will be a binomial experiment, since the 10 trials are independent, and the probability of success is 1/2 in each of the 10 trials. Let's summarize and give more examples.

To summarize, the requirements for a random experiment to be a binomial experiment are:

- a fixed number (n) of trials
- each trial must be independent of the others
- each trial has just two possible outcomes, called "**success**" (the outcome of interest) and "**failure**"
- there is a constant **probability (p) of success** for each trial, the complement of which is the **probability (1 - p) of failure**

In binomial random experiments, the number of successes in n trials is random. It can be as low as 0, if all the trials end up in failure, or as high as n, if all n trials end in success.

The random variable X that represents the number of successes in those n trials is called **binomial**, and is determined by the values of n and p. We say, "X is binomial with n = ... and p = ..."

### Example: Random Experiments (Binomial or Not?)

Let's consider a few random experiments.

In each of them, we'll decide whether the random variable is binomial. If it is, we'll determine the values for n and p. If it isn't, we'll explain why not.

1. A fair coin is flipped 20 times; X represents the number of heads.

**X is binomial with n = 20 and p = 0.5**.

2. You roll a fair die 50 times; X is the number of times you get a six.

   **X is binomial with n = 50 and p = 1/6**.

3. Roll a fair die repeatedly; X is the number of rolls it takes to get a six.

   **X is not binomial, because the number of trials is not fixed**.

4. Draw 3 cards at random, one after the other, **without replacement**, from a set of 4 cards consisting of one club, one diamond, one heart, and one spade; X is the number of diamonds selected.

   **X is not binomial, because the selections are not independent**. (The probability (p) of success is not constant, because it is affected by previous selections.)

5. Draw 3 cards at random, one after the other, **with replacement**, from a set of 4 cards consisting of one club, one diamond, one heart, and one spade; X is the number of diamonds selected. Sampling with replacement ensures independence.

   **X is binomial with n = 3 and p = 1/4**.

6. Approximately 1 in every 20 children has a certain disease. Let X be the number of children with the disease out of a random sample of 100 children. Although the children are sampled without replacement, it is assumed that we are sampling from such a vast population that the selections are virtually independent.

   **X is binomial with n = 100 and p = 1/20 = 0.05**.

7. The probability of having blood type B is 0.1. Choose 4 people at random; X is the number with blood type B.

   **X is binomial with n = 4 and p = 0.1**.

8. A student answers 10 quiz questions completely at random; the first five are true/false, the second five are multiple choice, with four options each. X represents the number of correct answers.

   **X is not binomial, because p changes from 1/2 to 1/4**.

---

## Comments

**Example 4** above was not binomial because sampling without replacement resulted in dependent selections. In particular, the probability of the second card being a diamond is very dependent on whether or not the first card was a diamond: the probability is 0 if the first card was a diamond, 1/3 if the first card was not a diamond.

In contrast, **Example 5** was binomial because sampling with replacement resulted in independent selections: the probability of any of the 3 cards being a diamond is 1/4 no matter what the previous selections have been.

On the other hand, when you take a relatively small random sample of subjects from a large population, even though the sampling is without replacement, we can assume independence because the mathematical effect of removing one individual from a very large population on the next selection is negligible. For example, in **Example 6**, we sampled 100 children out of the population of all children. Even though we sampled the children without replacement, whether one child has the disease or not really has no effect on whether another child has the disease or not. The same is true for **Example (7)**.

The convention is to "fudge" the requirement of independence as long as the population is at least 10 times the sample size.

### Rule of Thumb

The number (X) of successes in a sample of size n taken without replacement from a population with proportion (p) of successes is approximately binomial with n and p as long as the sample size (n) is at most 10% of the population size (N).

In symbols, this would be: **$n \le 0.10N$**.

This is the same as saying the population size is greater than or equal to 10 times the sample size. In symbols this is: **$N \ge 10n$**.

---

## Scenario: On Time Flights

A Department of Transportation report about air travel found that, nationwide, 78% of all flights are on time. Suppose a random sample of 50 flights is selected from all nationwide flights that were completed in the past 30 days (over 1000 flights). Let the random variable X be defined as the number of sampled flights that arrived on time.

---

## Learn By Doing

1/1 point (graded)

What is the value of the *n* parameter of the binomial random variable X?

○ 78

● 50 ✓

○ 39

○ 30

**Answer**
Correct:
X is a binomial random variable resulting from an experiment of 50 independent trials (n), in this case the number of flights sampled.

Submit

---

## Learn By Doing

1/1 point (graded)
What is the value of the *p* parameter of the binomial random variable X?

● 0.78 ✓

○ 0.22

○ 0.05

**Answer**
Correct: X is a binomial random variable with a probability of “success” (p) = “flight on-time” = 0.78.

Submit

---

## Did I Get This

1/1 point (graded)
In the following random experiment, is the random variable X binomial or not?

There are 6 members in a family (2 parents and 4 children). Let X be the number of family members who have blue eyes.

○ binomial

● not binomial ✓

**Answer**
Correct:
The random experiment consists of 6 trials (6 family members), where the outcome of interest (success) is having blue eyes. Even though X represents the number of "successes" among the 6 "trials," X is not binomial, because eye color is hereditary, and therefore the 6 trials are not independent. Knowing, for example, that the father has blue eyes has an impact on the probability that the children have blue eyes. X, therefore, is not binomial.

Submit

---

## Did I Get This

1/1 point (graded)
In the following random experiment, is the random variable X binomial or not?

There are 30 people at a party, 12 of which are males. Out of the 30 people at the party, 5 are selected at random to participate in a game. Let X be the number of females that were selected (out of the 5).

○ binomial

● not binomial ✓

**Answer**
Correct:
Note that you are sampling 5 subjects out of 30 without replacement. Therefore the 5 "trials" (selections) are not independent. Whether the first selected person is a male or a female has an impact on the second person selected being a female. Recall that the rule of thumb says, when you are sampling without replacement, you can ignore the dependence problem, as long as the population is at least 10 times as large as the sample size. In this case, the "population" consists of the 30 people in the party, and 5 of them were selected. 30 is less than 10 * 5, and therefore the rule of thumb is not satisfied.

Submit

---

## Did I Get This

1/1 point (graded)
In the following random experiment, is the random variable X binomial or not?

It is known that roughly 8% of males have some sort of color vision deficiency (also known as color-blindness). A random sample of 1,000 males was selected. Let X be the number of males out of the sample that are color-blind.

● binomial ✓

○ not binomial

**Answer**
Correct:
This random experiment consists of 1,000 trials (1,000 males), all having the same probability of being color-blind ("success"), and X represents how many of the trials (subjects) ended in a "success" (are color-blind). Note that even though the selection here is without replacement, we can disregard the dependence problem, and regard the trials as independent, since the population (all males) is more than 10 times the sample size (10 * 1,000 = 10,000).

Submit

## Did I Get This

1/1 point (graded)
In the following random experiment, is the random variable X binomial or not?

A multiple-choice quiz has 15 problems, each with 5 possible answers, only one of which is correct. A student who does not attend lectures has no clue, and uses an independent random guess to answer each of the problems. The random variable X is the number of questions the student got right.

● binomial ✓

○ not binomial

**Answer**
Correct:
This experiment consists of 15 trials (each quiz question is a trial). Since the student uses an independent random guess to answer each of the 15 questions, the trials are independent, and all have the same probability of being correctly answered. X represents how many of the 15 questions ended in a "success" (were answered correctly).

Submit

## Did I Get This

1/1 point (graded)

In the previous question (the one about the multiple-choice quiz), the random variable X is binomial with parameters:

○ n = 1/5, p = 15

○ n = 15, p = 1/2

● n = 15, p = 1/5 ✓

○ n = 15, p = 0

**Answer**

Correct:

There are 15 trials (15 quiz questions), all having a probability of success (guessing correctly) of 1/5.

Submit

---

## Scenario: Political Affilition

According to a recent national poll, 41% of females and 32% of males identify as democrats. Suppose a random sample of 120 females and 140 males is obtained.

---

## Did I Get This

1/1 point (graded)

For the binomial random variable X, the number of democrats in the sample of 120 females, what is the value of the following parameter?

n =

● 120 ✓

○ 140

**Answer**

Correct:

X is a binomial random variable from an experiment with 120 independent trials, in this case 120 females.

Submit

---

## Did I Get This

1/1 point (graded)

For the binomial random variable X, the number of democrats in the sample of 120 females, what is the value of the following parameter?

p =

○ 0.32

● 0.41 ✓

○ 0.59

**Answer**

Correct:

X is a binomial random variable from an experiment with a probability of “success” (being “democrat”) for females of 0.41.

Submit

---

## Did I Get This

1/1 point (graded)

For the binomial random variable Y, the number of democrats in the sample of 140 males, what is the value of the following parameter?

n =

○ 120

● 140 ✓

**Answer**

Correct:

Y is a binomial random variable from an experiment with 140 independent trials, in this case 140 males.

Submit

## Did I Get This

1/1 point (graded)

For the binomial random variable Y, the number of democrats in the sample of 140 males, what is the value of the following parameter?

p =

● 0.32 ✓

○ 0.41

○ 0.59

**Answer**

Correct:

Y is a binomial random variable from an experiment with a probability of “success” (being “democrat”) for males of 0.32.

Submit