Binomial Probability Distributions

Colby Community College

Types of Discrete Probability Distributions

In this section we will talk about the Binomial Probability Distribution, which is a discrete probability distribution.

Other types of discrete probability distributions include:

- Poisson distributions
- Geometric distributions
- Hypergeometric distributions

Types of Discrete Probability Distributions

In this section we will talk about the Binomial Probability Distribution, which is a discrete probability distribution.

Other types of discrete probability distributions include:

- Poisson distributions
- Geometric distributions
- Hypergeometric distributions

Definition

A **binomial probability distribution** results from a procedure that meets these four requirements:

Types of Discrete Probability Distributions

In this section we will talk about the Binomial Probability Distribution, which is a discrete probability distribution.

Other types of discrete probability distributions include:

- Poisson distributions
- Geometric distributions
- Hypergeometric distributions

Definition

A **binomial probability distribution** results from a procedure that meets these four requirements:

- The procedure has a fixed number of trials.

Types of Discrete Probability Distributions

In this section we will talk about the Binomial Probability Distribution, which is a discrete probability distribution.

Other types of discrete probability distributions include:

- Poisson distributions
- Geometric distributions
- Hypergeometric distributions

Definition

A **binomial probability distribution** results from a procedure that meets these four requirements:

- The procedure has a fixed number of trials.
- The trials must be independent.

Types of Discrete Probability Distributions

In this section we will talk about the Binomial Probability Distribution, which is a discrete probability distribution.

Other types of discrete probability distributions include:

- Poisson distributions
- Geometric distributions
- Hypergeometric distributions

Definition

A **binomial probability distribution** results from a procedure that meets these four requirements:

- The procedure has a fixed number of trials.
- The trials must be independent.
- Each trial must have all outcomes classified into exactly two categories, commonly referred to as **success** and failure.

Types of Discrete Probability Distributions

In this section we will talk about the Binomial Probability Distribution, which is a discrete probability distribution.

Other types of discrete probability distributions include:

- Poisson distributions
- Geometric distributions
- Hypergeometric distributions

Definition

A **binomial probability distribution** results from a procedure that meets these four requirements:

- The procedure has a fixed number of trials.
- The trials must be independent.
- Each trial must have all outcomes classified into exactly two categories, commonly referred to as **success** and failure.
- The probability of a success remains the same in all trails.

Notation

S (success) and F (failure) denote the two possible categories.

Notation

S (success) and F (failure) denote the two possible categories.
$P(S) = p \quad$ (p is the probability of a success)

Notation

S (success) and F (failure) denote the two possible categories.
$P(S) = p \quad$ (p is the probability of a success)
$P(F) = q \quad$ ($q = 1 - p$ is the probability of a failure)

Notation

S (success) and F (failure) denote the two possible categories.

$P(S) = p$	(p is the probability of a success)
$P(F) = q$	($q = 1 - p$ is the probability of a failure)
n	the fixed number of trails

Notation

S (success) and F (failure) denote the two possible categories.

$P(S) = p$	(p is the probability of a success)
$P(F) = q$	($q = 1 - p$ is the probability of a failure)
n	the fixed number of trails
x	a specific number of successes in n trails, so x is a whole number between 0 and n, inclusive.

Notation

S (success) and F (failure) denote the two possible categories.

$P(S) = p$	(p is the probability of a success)
$P(F) = q$	($q = 1 - p$ is the probability of a failure)
n	the fixed number of trails
x	a specific number of successes in n trails, so x is a whole number between 0 and n, inclusive.
$P(x)$	probability of getting exactly x successes among the n trials.

Notation

S (success) and F (failure) denote the two possible categories.

$P(S) = p$	(p is the probability of a success)
$P(F) = q$	($q = 1 - p$ is the probability of a failure)
n	the fixed number of trails
x	a specific number of successes in n trails, so x is a whole number between 0 and n, inclusive.
$P(x)$	probability of getting exactly x successes among the n trials.

Caution

When using a binomial probability distribution, *always* be sure that x and p are consistent in the sense that they both refer to the *same* category being called a success.

Notation

S (success) and F (failure) denote the two possible categories.	
$P(S) = p$	(p is the probability of a success)
$P(F) = q$	($q = 1 - p$ is the probability of a failure)
n	the fixed number of trails
x	a specific number of successes in n trails, so x is a whole number between 0 and n, inclusive.
$P(x)$	probability of getting exactly x successes among the n trials.

Caution

When using a binomial probability distribution, *always* be sure that x and p are consistent in the sense that they both refer to the *same* category being called a success.

Caution

The probability p is the probability of getting a success on just *one* individual trail.

Example 1

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

Example 1

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

Does this procedure result in a binomial distribution?

Example 1

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

Does this procedure result in a binomial distribution?

- The procedure has 5 trials.

Example 1

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

Does this procedure result in a binomial distribution?

- The procedure has 5 trials.
- The 5 trials are independent because the probability of any adult knowing Twitter is not affected by results from other selected adults.

Example 1

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

Does this procedure result in a binomial distribution?

- The procedure has 5 trials.
- The 5 trials are independent because the probability of any adult knowing Twitter is not affected by results from other selected adults.
- The two categories are "Knows Twitter" and "Doesn't know Twitter."

Example 1

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

Does this procedure result in a binomial distribution?

- The procedure has 5 trials.
- The 5 trials are independent because the probability of any adult knowing Twitter is not affected by results from other selected adults.
- The two categories are "Knows Twitter" and "Doesn't know Twitter."
- For each randomly selected adult, the probability they know Twitter is 0.85 and that probability remains the same for each selected person.

Example 1

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

Does this procedure result in a binomial distribution?

- The procedure has 5 trials.
- The 5 trials are independent because the probability of any adult knowing Twitter is not affected by results from other selected adults.
- The two categories are "Knows Twitter" and "Doesn't know Twitter."
- For each randomly selected adult, the probability they know Twitter is 0.85 and that probability remains the same for each selected person.

We see that this is a binomial distribution.

Example 1

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

Does this procedure result in a binomial distribution?

- The procedure has 5 trials.
- The 5 trials are independent because the probability of any adult knowing Twitter is not affected by results from other selected adults.
- The two categories are "Knows Twitter" and "Doesn't know Twitter."
- For each randomly selected adult, the probability they know Twitter is 0.85 and that probability remains the same for each selected person.

We see that this is a binomial distribution.

In this example we have

$$\begin{aligned} n &= 5 & x &= 3 \\ p &= 0.85 & q &= 1 - 0.85 = 0.15 \end{aligned}$$

Note

Sampling without replacement violates the requirements for a binomial distribution.

Note

Sampling without replacement violates the requirements for a binomial distribution.

5% Guideline for Cumbersome Calculations

When sampling without replacement and the same size is no more than 5% of the size of the population, treat the selections as being independent (even though they are actually dependent).

Note

Sampling without replacement violates the requirements for a binomial distribution.

5% Guideline for Cumbersome Calculations

When sampling without replacement and the same size is no more than 5% of the size of the population, treat the selections as being independent (even though they are actually dependent).

Methods of Finding Binomial Probabilities

1. Using the Binomial Probability Formula.
2. Using technology.
3. Using a table.

Note

Sampling without replacement violates the requirements for a binomial distribution.

5% Guideline for Cumbersome Calculations

When sampling without replacement and the same size is no more than 5% of the size of the population, treat the selections as being independent (even though they are actually dependent).

Methods of Finding Binomial Probabilities

1. Using the Binomial Probability Formula.
2. Using technology.
3. Using a table.

Note

Technology is most often used to calculate binomial probabilities.

Binomial Probability Formula

$$P(x) = \frac{n!}{(n-x)!x!} \cdot p^x \cdot q^{n-x} \qquad \text{for } x = 0, 1, \ldots, n$$

where

n	is the number of trails.
x	is the number of successes among n trials.
p	is the probability of success in any one trial.
q	is the probability of failure in any one trial. $(q = 1 - p)$

Binomial Probability Formula

$$P(x) = \frac{n!}{(n-x)!x!} \cdot p^x \cdot q^{n-x} \qquad \text{for } x = 0, 1, \ldots, n$$

where

n	is the number of trails.
x	is the number of successes among n trials.
p	is the probability of success in any one trial.
q	is the probability of failure in any one trial. $(q = 1 - p)$

Note

We can also write this formula as

$$P(x) = {}_nC_x \cdot p^x \cdot q^{n-x} \qquad \text{for } x = 0, 1, \ldots, n$$

where x items identical to themselves, and $n - x$ other items identical to themselves, the number of permutations is ${}_nC_x$.

Example 2

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

Example 2

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

We have

$$\begin{array}{ll} n = 5 & x = 3 \\ p = 0.85 & q = 1 - 0.85 = 0.15 \end{array}$$

Example 2

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

We have

$$\begin{array}{ll} n = 5 & x = 3 \\ p = 0.85 & q = 1 - 0.85 = 0.15 \end{array}$$

Applying the formula gives

$$P(3) = \frac{5!}{(5-3)!3!} \cdot (0.85)^3 \cdot (0.15)^{5-3}$$

Example 2

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

We have

$$\begin{aligned} n &= 5 & x &= 3 \\ p &= 0.85 & q &= 1 - 0.85 = 0.15 \end{aligned}$$

Applying the formula gives

$$\begin{aligned} P(3) &= \frac{5!}{(5-3)!3!} \cdot (0.85)^3 \cdot (0.15)^{5-3} \\ &= 0.138 \text{ (rounded)} \end{aligned}$$

Example 2

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

We have

$$\begin{array}{ll} n = 5 & x = 3 \\ p = 0.85 & q = 1 - 0.85 = 0.15 \end{array}$$

Applying the formula gives

$$\begin{aligned} P(3) &= \frac{5!}{(5-3)!3!} \cdot (0.85)^3 \cdot (0.15)^{5-3} \\ &= 0.138 \text{ (rounded)} \end{aligned}$$

We then round to three significant digits to get the probability that exactly three out of five randomly selected adults know Twitter is 0.138.

Technology

If you have a programmable calculator that does not have a binomial distribution function, you can create a program from the formula.

Technology

If you have a programmable calculator that does not have a binomial distribution function, you can create a program from the formula.

Example 3

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

Technology

If you have a programmable calculator that does not have a binomial distribution function, you can create a program from the formula.

Example 3

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

We have

$$\begin{array}{ll} n = 5 & x = 3 \\ p = 0.85 & q = 0.15 \end{array}$$

Technology

If you have a programmable calculator that does not have a binomial distribution function, you can create a program from the formula.

Example 3

When an adult is randomly selected (with replacement), there is a 0.85 probability that this person knows what Twitter is (based on results from a Pew Research Center survey). Suppose that we want to find the probability that exactly three of five randomly selected adults know what Twitter is.

We have

$$\begin{array}{ll} n = 5 & x = 3 \\ p = 0.85 & q = 0.15 \end{array}$$

Calculator gives:	0.138178125
Statdisk gives:	0.1381781
Table gives:	0.139

Example 4

Between 1974 and 2011, there were 460 NFL football games decided in overtime, and 252 of them were won by the team that won the overtime coin toss. Is the result of 252 wins in the 460 games equivalent to random chance, or is the 252 wins significantly high?

Example 4

Between 1974 and 2011, there were 460 NFL football games decided in overtime, and 252 of them were won by the team that won the overtime coin toss. Is the result of 252 wins in the 460 games equivalent to random chance, or is the 252 wins significantly high?

Using the notation for binomial probabilities, we have $n = 460$, $p = 0.5$, and $q = 0.5$. What is x?

Example 4

Between 1974 and 2011, there were 460 NFL football games decided in overtime, and 252 of them were won by the team that won the overtime coin toss. Is the result of 252 wins in the 460 games equivalent to random chance, or is the 252 wins significantly high?

Using the notation for binomial probabilities, we have $n = 460$, $p = 0.5$, and $q = 0.5$. What is x?

Significantly high means we need to check if the probability of 252 or more wins in 460 games is less than 0.05.

Example 4

Between 1974 and 2011, there were 460 NFL football games decided in overtime, and 252 of them were won by the team that won the overtime coin toss. Is the result of 252 wins in the 460 games equivalent to random chance, or is the 252 wins significantly high?

Using the notation for binomial probabilities, we have $n = 460$, $p = 0.5$, and $q = 0.5$. What is x?

Significantly high means we need to check if the probability of 252 or more wins in 460 games is less than 0.05.

To calculate this with the formula, you would need to calculate

$$\begin{aligned} P\,(252 \text{ or more}) &= P\,(252 \text{ or } 253 \text{ or } \ldots \text{ or } 459 \text{ or } 460) \\ &= P\,(252) + P\,(253) + \cdots + P\,(459) + P\,(460) \end{aligned}$$

Example 4

Between 1974 and 2011, there were 460 NFL football games decided in overtime, and 252 of them were won by the team that won the overtime coin toss. Is the result of 252 wins in the 460 games equivalent to random chance, or is the 252 wins significantly high?

Using the notation for binomial probabilities, we have $n = 460$, $p = 0.5$, and $q = 0.5$. What is x?

Significantly high means we need to check if the probability of 252 or more wins in 460 games is less than 0.05.

To calculate this with the formula, you would need to calculate

$$\begin{aligned} P\left(252 \text{ or more}\right) &= P\left(252 \text{ or } 253 \text{ or } \ldots \text{ or } 459 \text{ or } 460\right) \\ &= P\left(252\right) + P\left(253\right) + \cdots + P\left(459\right) + P\left(460\right) \end{aligned}$$

Instead, it is better to use technology.

$$P\left(252 \text{ or more}\right) = 0.0224$$

Example 4

Between 1974 and 2011, there were 460 NFL football games decided in overtime, and 252 of them were won by the team that won the overtime coin toss. Is the result of 252 wins in the 460 games equivalent to random chance, or is the 252 wins significantly high?

Using the notation for binomial probabilities, we have $n = 460$, $p = 0.5$, and $q = 0.5$. What is x?

Significantly high means we need to check if the probability of 252 or more wins in 460 games is less than 0.05.

To calculate this with the formula, you would need to calculate

$$\begin{aligned} P(252 \text{ or more}) &= P(252 \text{ or } 253 \text{ or } \ldots \text{ or } 459 \text{ or } 460) \\ &= P(252) + P(253) + \cdots + P(459) + P(460) \end{aligned}$$

Instead, it is better to use technology.

$$P(252 \text{ or more}) = 0.0224$$

Since $0.0224 < 0.05$ we see that it is unlikely we would get 252 or more wins by chance.

Mean and Standard Deviation

For a binomial distribution, the formulas for mean and standard deviation can be rewritten as:

$$\begin{aligned} \text{Mean:} \quad & \mu &= np \\ \text{Variance:} \quad & \sigma^2 &= npq \\ \text{Standard Deviation:} \quad & \sigma &= \sqrt{npq} \end{aligned}$$

Mean and Standard Deviation

For a binomial distribution, the formulas for mean and standard deviation can be rewritten as:

$$\begin{aligned} \text{Mean:} \quad & \mu &= np \\ \text{Variance:} \quad & \sigma^2 &= npq \\ \text{Standard Deviation:} \quad & \sigma &= \sqrt{npq} \end{aligned}$$

Range Rule of Thumb

Significantly low values $\leq (\mu - 2\sigma)$

Significantly high values $\geq (\mu + 2\sigma)$

Values not significant: Between $(\mu - 2\sigma)$ and $(\mu + 2\sigma)$

Example 5

Between 1974 and 2011, there were 460 NFL football games decided in overtime, and 252 of them were won by the team that won the overtime coin toss.

Example 5

Between 1974 and 2011, there were 460 NFL football games decided in overtime, and 252 of them were won by the team that won the overtime coin toss.

Using the notation for binomial probabilities, we have $n = 460$, $p = 0.5$, and $q = 0.5$. We then get:

$$\mu = np = (460)(0.5) = 230 \text{ games}$$

$$\sigma = \sqrt{npq} = \sqrt{(460)(0.5)(0.5)} = 10.7 \text{ games (rounded)}$$

Example 5

Between 1974 and 2011, there were 460 NFL football games decided in overtime, and 252 of them were won by the team that won the overtime coin toss.

Using the notation for binomial probabilities, we have $n = 460$, $p = 0.5$, and $q = 0.5$. We then get:

$$\mu = np = (460)(0.5) = 230 \text{ games}$$
$$\sigma = \sqrt{npq} = \sqrt{(460)(0.5)(0.5)} = 10.7 \text{ games (rounded)}$$

The number of wins that are significantly low are:

$$\mu - 2\sigma = 230 - 2(10.7) = 208.6 \text{ games or fewer}$$

Example 5

Between 1974 and 2011, there were 460 NFL football games decided in overtime, and 252 of them were won by the team that won the overtime coin toss.

Using the notation for binomial probabilities, we have $n = 460$, $p = 0.5$, and $q = 0.5$. We then get:

$$\mu = np = (460)(0.5) = 230 \text{ games}$$
$$\sigma = \sqrt{npq} = \sqrt{(460)(0.5)(0.5)} = 10.7 \text{ games (rounded)}$$

The number of wins that are significantly low are:

$$\mu - 2\sigma = 230 - 2(10.7) = 208.6 \text{ games or fewer}$$

The number of wins that are significantly high are:

$$\mu + 2\sigma = 230 + 2(10.7) = 251.4 \text{ games or more}$$

Example 5

Between 1974 and 2011, there were 460 NFL football games decided in overtime, and 252 of them were won by the team that won the overtime coin toss.

Using the notation for binomial probabilities, we have $n = 460$, $p = 0.5$, and $q = 0.5$. We then get:

$$\mu = np = (460)(0.5) = 230 \text{ games}$$
$$\sigma = \sqrt{npq} = \sqrt{(460)(0.5)(0.5)} = 10.7 \text{ games (rounded)}$$

The number of wins that are significantly low are:

$$\mu - 2\sigma = 230 - 2(10.7) = 208.6 \text{ games or fewer}$$

The number of wins that are significantly high are:

$$\mu + 2\sigma = 230 + 2(10.7) = 251.4 \text{ games or more}$$

Since $252 > 251.4$, we see that 252 wins is significantly high.