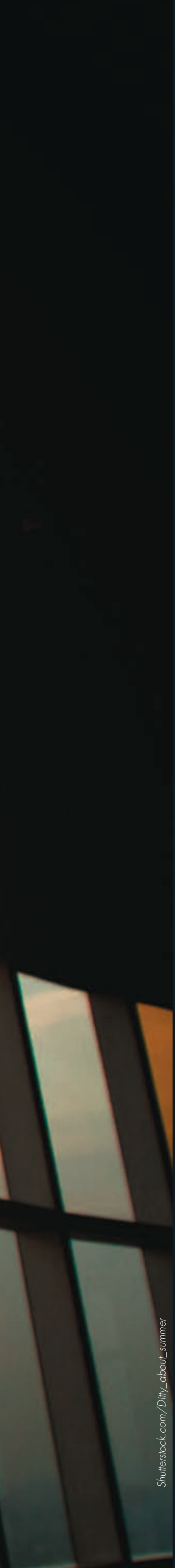


5

TERMINOLOGY

Bernoulli distribution
Bernoulli trial
Bernoulli variable
binomial distribution
binomial variable
expected value
failure
geometric distribution
geometric variable
independent
mean
negatively skewed
positively skewed
skewed
standard deviation
success
symmetrical distribution
variance

Prior learning

DISCRETE RANDOM VARIABLES

BINOMIAL DISTRIBUTIONS

BERNOULLI DISTRIBUTIONS

- use a **Bernoulli random variable** as a model for two-outcome situations (ACMMM143)
- identify contexts suitable for modelling by **Bernoulli random variables** (ACMMM144)
- recognise the mean p and variance $p(1 - p)$ of the Bernoulli distribution with parameter p (ACMMM145)
- use Bernoulli **random variables** and associated probabilities to model data and solve practical problems. (ACMMM146)

BINOMIAL DISTRIBUTIONS

- understand the concepts of **Bernoulli trials** and the concept of a binomial **random variable** as the number of 'successes' in n independent **Bernoulli trials**, with the same probability of success p in each trial (ACMMM147)
- identify contexts suitable for modelling by binomial **random variables** (ACMMM148)
- determine and use the probabilities $P(X = r) = \binom{n}{r} p^r (1 - p)^{n-r}$ associated with the **binomial distribution** with parameters n and p; note the mean np and variance $np(1 - p)$ of a **binomial distribution** (ACMMM149)
- use binomial distributions and associated probabilities to solve practical problems. (ACMMM150) AC

5.01 THE BERNOULLI DISTRIBUTION

There are many probability situations with only two possible outcomes, such as head/tail, boy/girl, etc. In other situations, we can classify the outcomes so that there are only two possibilities, such as drawing an ace or not drawing an ace from a pack of cards.

A distribution in which there are only two possible outcomes is known as a **Bernoulli distribution**, after Jacob Bernoulli (1654–1705).

IMPORTANT

A **Bernoulli distribution** is a distribution with two possible outcomes, called **success** and **failure**, with a fixed probability of success. The probability of success is normally written as p and the probability of failure as q, such that $p + q = 1$.

The **Bernoulli random variable** X is given the values 1 for success and 0 for failure.

Example 1

Determine whether each of the following situations can be considered as Bernoulli distributions.

a Rolling a six-sided die and recording a success when a 3 occurs.

b Rolling a six-sided die and recording the number that occurs.

c Drawing a card from a deck 6 times with replacement and recording the number of hearts drawn.

d Drawing a marble from a bag containing 6 red and 4 green marbles with replacement and recording the colour of the marble drawn.

Solution

a How many outcomes are there?

There are two outcomes – rolling a 3 and rolling a number other than 3. $p = \frac{1}{6}$ and $q = \frac{5}{6}$, so $p + q = 1$.

Write the answer.

This is a Bernoulli distribution.

9780170250290

b	How many outcomes are there?	There are six outcomes.
	Write the answer.	This is not a Bernoulli distribution.
c	How many outcomes are there?	There are seven outcomes (0 to 6 hearts).
	Write the answer.	This is not a Bernoulli distribution.
d	How many outcomes are there?	There are two outcomes, red and green. $p = \frac{1}{2}$ and $q = \frac{1}{2}$, so $p + q = 1$.
	Write the answer.	This is a Bernoulli distribution

The trivial case of the hypergeometric distribution with a sample of size $n = 1$ is a Bernoulli distribution. Remember that a hypergeometric distribution is one where a sample of n items is taken from a population of N items, k of which are considered successes. The hypergeometric random variable is the number of successes.

Example 2

What is the probability of success if a watch is chosen from a batch of 100, of which 4 are faulty, and success is considered as getting a faulty watch?

Solution

List the hypergeometric parameters.	$x = 1,\ n = 1,\ N = 100,\ k = 4$
Find the probability.	$P(\text{success}) = \frac{4}{100} = \frac{1}{25}$

You could calculate the probability of success in Example **2** by using the formula for the hypergeometric distribution,

$$P(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

with $x = 1$, $n = 1$, $N = 100$ and $k = 4$, but it would be unnecessarily complicated to do it that way.

We saw in Chapter 2 of this book that the mean (expected value) and variance of a discrete probability distribution are defined as the following.

WS
The Bernoulli distribution

$$\mu = E(X) = \sum x \cdot p(x)$$

$$\begin{aligned} Var(X) &= E\left[(X-\mu)^2\right] \\ &= E\left(X^2\right) - \mu^2 \\ &= \sum x^2 \cdot p(x) - \left[\sum x \cdot p(x)\right]^2 \end{aligned}$$

Consider a Bernoulli distribution with probability of success p and random variable X. The calculation of the mean is easy because there are only two values of X, 0 and 1.

Then $E(X) = 0 \times q + 1 \times p = p$

$$\begin{aligned} Var(X) &= E(X^2) - \mu^2 \\ &= \sum x^2 \cdot p(x) - p^2 \\ &= [0^2 \times q + 1^2 \times p] - p^2 \\ &= p - p^2 \\ &= p(1-p) \end{aligned}$$

Since $SD(X) = \sqrt{Var(X)}$, $SD(X) = \sqrt{p(1-p)}$

IMPORTANT

The expected value, variance and standard deviation of a Bernoulli distribution are:

$E(X) = \mu = p$, $Var(X) = p(1-p)$ and $SD(X) = \sigma = \sqrt{p(1-p)}$

Example 3

What is the standard deviation of the random variable X arising from the selection of a card from a normal pack if you want it to be hearts?

Solution

Write the values of X.	$X = 1$ for hearts and 0 for other suits.
Write the value of p.	$p = \frac{1}{4}$
Write the formula.	$\sigma = \sqrt{p(1-p)}$
Substitute in the values.	$= \sqrt{\frac{1}{4}\left(1 - \frac{1}{4}\right)}$
Calculate the answer.	$= \frac{\sqrt{3}}{4}$

EXERCISE 5.01 The Bernoulli distribution

Concepts and techniques

1 Example 1 Which of the following are examples of a Bernoulli distribution?

- a Selecting a marble from a bag containing red, green and blue marbles
- b Selecting people at random and noting if they are over 18
- c Drawing a card from a deck and recording its colour
- d Tossing a coin 3 times and recording the number of tails
- e Playing roulette and noting whether the ball settles on an odd number

2 Example 2 What is the probability of randomly getting an orange that you can eat from a box that contains 30 oranges, of which 3 are rotten?

A $\frac{1}{30}$ B $\frac{3}{10}$ C $\frac{1}{3}$ D $\frac{1}{2}$ E $\frac{9}{10}$

3 Example 3 What is the standard deviation of the random variable arising from picking the multiple choice answer from this question at random?

A $\frac{\sqrt{5}}{4}$ B $\frac{1}{5}$ C $\frac{4}{5}$ D $\frac{2}{\sqrt{5}}$ E $\frac{2}{5}$

4 State whether or not each of the following are Bernoulli distributions.
- a Randomly selecting a person and recording if the person is female or male.
- b Tossing a coin 20 times and recording the number of heads.
- c Spinning a spinner with equal sized segments numbered 1 to 12 and recording the number where the spinner stops.
- d Drawing a marble from a bag containing 8 green and 5 orange marbles with replacement and recording if the marble is orange or not.
- e Drawing a card from a deck with replacement and recording if it is black or red.
- f Drawing a card from a deck without replacement and recording the number of aces drawn.
- g Rolling a six-sided die 20 times and recording the number that comes up.

5 Calculate the probability of success in each of the following Bernoulli distributions.
- a Choosing a red onion at random from a bag with an equal mixture of red and brown onions.
- b Getting a six from rolling an eight sided-die.
- c Picking a sherbet at random from a bag containing 3 sherbets, 4 jubes and 5 toffees.

6 Calculate the variances of the random variables in Bernoulli distributions with the following probabilities.

a $p = \frac{2}{5}$ b $q = \frac{1}{3}$ c $p = 0.6$ d $q = 0.3$ e $p = \frac{3}{4}$

7 Calculate the standard deviation of the random variables in Bernoulli distributions with the following probabilities.

a $q = 0.4$ b $p = \frac{2}{3}$ c $q = \frac{24}{25}$ d $p = 0.2$ e $p = \frac{7}{9}$

Reasoning and communication

8 It is known that 8% of all items produced on a particular assembly line are defective. They are packed in boxes of 5 items. Boxes pass quality control if none are defective. What is the probability of a box passing quality control?

9 Peter picks a marble from a bag containing two red, three green and one blue marble. He wants a red marble. What is the standard deviation of the random variable for success?

10 A card is drawn from a well-shuffled deck, checked to see whether it is an ace and then returned to the deck. What is the variance of the random variable arising from this situation?

11 On average, Simone is late for school 20% of the time. In a normal school week, what is the probability that she will be successful in getting to school every day?

12 A manufacturer will only accept a batch of 200 bags of plastic beads if a sample of 10 bags has at most 1 bag that has obvious misshaped beads, because these occasionally cause problems in the plant. What is the probability of acceptance, given that the supplier says that only 10% of the bags have misshaped beads?

5.02 THE GEOMETRIC DISTRIBUTION

The geometric distribution is a distribution that arises from multiple trials, each of which is the same Bernoulli distribution. These are called **Bernoulli trials**.

IMPORTANT

Bernoulli trials are independent trials that have only two possible outcomes, called **success** and **failure**, where all the trials have the same probability of success. The probability of success is normally written as p and the probability of failure as q, so $p + q = 1$.

In many cases involving Bernoulli trials, you may continue until you get a success. For example, if you wanted to grow an oak tree from an acorn, you would keep planting seeds until you succeeded in growing the tree.

The pattern is a sequence of failures (F) before the first success (S), symbolised as:

$$F, F, F, \ldots, S$$

Example 4

A fair eight-sided die is rolled repeatedly until a multiple of 3 (a success) is observed. What is the probability that exactly five rolls are required?

Solution

Identify the situation.	The trials are Bernoulli trials with a sequence *FFFFS*.
Calculate p.	$p = P(3 \text{ or } 6) = \frac{2}{8} = \frac{1}{4}$
Calculate q.	$q = 1 - \frac{1}{4} = \frac{3}{4}$
Use independence of trials.	$P(FFFFS) = \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} \times \frac{1}{4}$
	$= \frac{3^4}{4^5}$
Calculate the result.	$= \frac{81}{1024}$
	$= 0.079\,101\ldots$
Write the answer.	The probability of 5 rolls is about 0.079 or 7.9%.

The pattern in Example 4 is called a **geometric distribution**. It is a discrete distribution, but unlike the ones you have previously met, it is an infinite distribution. This means that there is a finite (although very small) probability that the number of failures before you get a success is any large number that you choose.

9780170250290

IMPORTANT

The **geometric distribution** is the probability distribution of the number of failures of Bernoulli trials before the first success.

The **geometric random variable** X has the values 0, 1, 2, 3, …: $X \in \boldsymbol{J}^+$.

The probability function is $P(X = x) = q^x p$.

The geometric distribution is sometimes defined as the number of trials Y needed to get a success, in which case the random variable takes the values 1, 2, … and $P(Y = y) = q^{y-1}p$.

The probability function is just a consequence of the fact that the trials are independent, so the overall probability is the product of the probabilities of x failures followed by 1 success.

The probability function is a geometric sequence.

You may have seen last year that the infinite sum of a GP with common ratio r is given by $\frac{a}{1-r}$, provided that $-1 \le r \le 1$. For the geometric sequence, this gives

$$\sum q^x p = \frac{p}{1-q} = \frac{1-q}{1-q} = 1$$

as expected.

You can calculate the expected value as shown below:

$$E(X) = \sum xq^x p = 0 \times q^0 \times p + 1 \times q^1 \times p + 2 \times q^2 \times p + 3 \times q^3 \times p + 4 \times q^4 \times p + \cdots$$

Thus $E(X) = 1 \times q^1 \times p + 2 \times q^2 \times p + 3 \times q^3 \times p + 4 \times q^4 \times p + 5 \times q^5 \times p + \cdots$

so $qE(X) = 1 \times q^2 \times p + 2 \times q^3 \times p + 3 \times q^4 \times p + 4 \times q^5 \times p + \cdots$

Subtracting the two expressions gives the following.

$$\begin{aligned} E(X) - qE(X) &= q^1 \times p + q^2 \times p + q^3 \times p + q^4 \times p + q^5 \times p + \cdots \\ &= \frac{pq}{1-q} \text{ (using the infinite sum of a GP again)} \\ &= \frac{(1-q)q}{1-q} = q \end{aligned}$$

Factorising the LHS, $(1 - q)E(X) = q$, so $E(X) = \frac{q}{1-q} = \frac{(1-p)}{1-(1-p)} = \frac{1-p}{p}$

IMPORTANT

The expected value of the geometric distribution is $E(X) = \frac{q}{1-q} = \frac{1-p}{p}$

Example 5

In a bottling plant, a problem has been discovered with a labelling machine. Over time, it has been determined that 25% of all labels are not properly fixed to bottles. If X denotes the number of labels that are fixed correctly before the first one that is not fixed properly, calculate:

a $P(X = 5)$
b $P(X \le 5)$
c $E(X)$ and explain the result.


123RF/gemenacom

Solution

a Calculate p and q.

$p = 25\% = \frac{1}{4}$

$q = 1 - \frac{1}{4} = \frac{3}{4}$

Write the geometric function.

$P(X = x) = \left(\frac{3}{4}\right)^x \left(\frac{1}{4}\right)$

Let $X = 5$.

$$P(X = 5) = \left(\frac{3}{4}\right)^5 \left(\frac{1}{4}\right) = \frac{243}{4096} \approx 0.0593$$

b Write an expression for $P(X \le 5)$.

$P(X \le 5) = P(X = 0) + P(X = 1) + P(X = 2) + \cdots + P(X = 5)$

Substitute.

$= \left(\frac{1}{4}\right) + \left(\frac{3}{4}\right)^1\left(\frac{1}{4}\right) + \left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right) + \cdots + \left(\frac{3}{4}\right)^5\left(\frac{1}{4}\right)$

State the kind of sum.

This is a GP with $a = \frac{1}{4}$, $r = \frac{3}{4}$ and $n = 6$.

Use the rule for the sum of n terms.

$S_n = \frac{a(1 - r^6)}{1 - r}$

Substitute for a, r and n.

$= \frac{\frac{1}{4}\left[1 - \left(\frac{3}{4}\right)^6\right]}{1 - \frac{3}{4}}$

Evaluate and round off.

$= 1 - \left(\frac{3}{4}\right)^6$

≈ 0.8220

9780170250290

c	Write the rule for $E(X)$.	$E(X)=\dfrac{1-p}{p}$
	Substitute for p.	$=\dfrac{1-\frac{1}{4}}{\frac{1}{4}}$
	Evaluate.	$=\dfrac{3}{4}\times\dfrac{4}{1}=3$
	State the result.	$E(X)=3$. There will be an average of 3 bottles labelled correctly before the first incorrectly labelled bottle occurs.

Even though the average number of bottles before a dud label is found is 3, this does not mean that there couldn't be a run of 100, or 1000, or any number you like to think of. Remember that the geometric distribution is infinite.

Example 6

A basketball player knows from experience that she will make 60% of all attempted baskets from the free throw line. If Y denotes the number of free throws she will make to get her first success, find:

a $P(Y=8)$ **b** $P(2\le Y\le 7)$

Solution

a	Identify the situation.	This is the alternate geometric distribution.
	Calculate p and q.	$p=60\%=0.6$ and $q=1-0.6=0.4$
	Write the function.	$P(Y=y)=(0.4)^{y-1}(0.6)$
	Substitute for $Y=8$.	$P(Y=8)=(0.4)^{8-1}(0.6)$ $=(0.4)^7(0.6)$ $\approx 0.000\,98$
b	Define $P(2\le Y\le 7)$.	$P(2\le Y\le 7)=P(Y=2)+P(Y=3)+P(Y=4)+\cdots+P(Y=7)$
	Substitute.	$=(0.4)(0.6)+(0.4)^2(0.6)+(0.4)^3(0.6)+\cdots+(0.4)^6(0.6)$
	State the kind of sum.	This is a GP with $a=(0.4)(0.6)$, $r=0.4$ and $n=6$
	Use the rule for the sum.	$S_n=\dfrac{a(1-r^6)}{1-r}$
	Substitute.	$=\dfrac{(0.4)(0.6)\left[1-(0.4)^6\right]}{1-0.4}$
	Evaluate.	≈ 0.3984

EXERCISE 5.02 The geometric distribution

Concepts and techniques

Geometric probability distributions

1 Which of the following variables will have a geometric probability distribution?

A the number of phone calls received at a call centre in successive 10-minute periods

B the number of cards I need to deal from a well-shuffled pack of 52 cards until at least two of the cards are hearts

C the number of marbles drawn from a bag of coloured marbles until a red and a green marble have been selected

D the number of digits I read beginning at a randomly selected point in a table of random digits until I find a 9

E the number of times a coin must be flipped until 2 heads or 3 tails have been flipped

2 Example 4 A basketball player has an 80% chance of making a free throw. If this probability is the same for each free throw attempted, what is the probability that the player doesn't make a free throw in a game until the fifth attempt?

A 0.000 31 B 0.001 28 C 0.002 57 D 0.005 26 E 0.081 92

3 Example 5 X has a geometric distribution where X is number of failures before the first success. If the probability of success is 0.3. Then $P(X = 4)$ is equal to:

A 0.0057 B 0.0187 C 0.0519 D 0.07203 E 0.1029

4 Example 6 For the basketball player mentioned in question **2**, what is the probability that it takes more than 3 free throws before the player makes the first free throw?

A 0.0008 B 0.0016 C 0.0032 D 0.0064 E 0.1024

123RF/Konstantin Pukhov

5 For the basketball player mentioned in question **2**, what is the expected number of throws required *until the player makes the first free throw* in a game?

A 0.25 B 0.8 C 1.25 D 2 E 4

9780170250290

6 Which of the following situations are examples of geometric distributions?

a If both parents carry genes for a particular condition, each child has a 25% chance of getting two genes (one from each parent) for that condition. Children inherit genes independently of each other. We wish to find the probability that the first child these parents have with this condition is their third child.

b A card is drawn from a well-shuffled deck, its suit is noted and the card is set aside before the next card is drawn. We wish to find the number of cards drawn before the first heart is selected.

c Coloured marbles are randomly selected from a bag, the colour noted and the marble returned before the next selection. The bag contains 5 red, 7 white and 2 green marbles. We wish to find the number of marbles drawn before the first green marble is selected.

d At the maternity section of a large hospital, a study is undertaken to determine the number of children born to a couple before they have a girl.

e The pool of potential jurors for a murder trial contains 500 people chosen randomly from the adult population of a large city. Each person in the pool is asked if he or she believes that judges are too lenient in sentencing people found guilty of major crimes. We are interested in the number of potential jurors interviewed before the first 'yes' response.

7 A six-sided die is rolled until the first time that a number greater than 4 is observed. What is the probability that exactly six rolls are required?

8 A baseball batter knows that he hits only one pitch out of every three. What is the probability that he misses the ball less than three times?

9 A gambler keeps a careful record of his performance and knows that, in the long run, he wins once in every five bets. If X represents the number of losses before the first win on a particular day, find:

a $P(X = 5)$ b $P(X \leq 5)$ c $E(X)$

10 In a batch of graded eggs, three-quarters of the eggs are underweight. If X represents the number of underweight eggs before an egg with the correct weight is found, calculate:

a $P(X \geq 1)$ b $P(1 \leq X \leq 5)$

11 A beginner golfer knows that, on average, he is able to hit his drive straight once in every 10 attempts. If X represents the number of bad shots before he hits a straight one, find:

a $P(4 \leq X \leq 14)$ b the expected value of X.

12 A student is sitting for a multiple choice test in which each question has 10 answer choices. The student randomly selects the answer for each question. If Y denotes the number of questions answered for the first correct answer, find:

a $P(Y = 7)$ b $P(2 \leq Y \leq 8)$

Reasoning and communication

13 If a fair six-sided die is rolled until an even number appears, what is the probability that the number of times it is rolled is:

a exactly 2? b at least 2? c no more than 2?

14 A golfer knows from experience that, if she putts within 5 m of the hole, she sinks the putt three times out of four. Find the probability that, in a particular game, she fails to sink a putt within 5 m four times before finally sinking one.

15 Police know from long experience that, on a particular stretch of road, 1 car in every 10 will exceed the speed limit. If a speed gun is used on this stretch of road, find the probability that the police will find that the first five cars will be within the speed limit and the sixth car will be speeding.

16 Assume that the Australian Tax Office (ATO) catches about 25% of all fraudulent returns each year. If you submit a fraudulent return every year:

a what is the probability that you could get away with it five years in a row?

b what is the expected number of fraudulent returns you could submit before being caught for the first time?

17 It is known that 75% of students hold jobs while attending a particular university. Suppose that students are selected at random from the student body. What is the probability that the first student selected who does not hold a job is the fourth student selected?

5.03 THE BINOMIAL DISTRIBUTION

You have already looked at Bernoulli distributions and Bernoulli trials. In this section you will study the most important distribution with Bernoulli trials. The geometric distribution arises from Bernoulli trials where the random variable is the number of failures before the first success. The **binomial distribution** arises from a finite number of Bernoulli trials where the random variable is the number of successes in a fixed number of trials.

IMPORTANT

The **binomial distribution** is the probability distribution of the number of successes arising from a fixed number of Bernoulli trials. A situation that produces a binomial distribution may be called a **binomial experiment**.

The properties of the binomial distribution are:

- there are a fixed number, n, of repeated trials
- each trial has only two outcomes – success and failure with probabilities p and q
- the probability of success is the same for each trial
- the trials are independent

The **binomial random variable** is the number of successes from n trials.

9780170250290

Example 7

Determine which of the following situations are examples of binomial experiments.

a Tossing a coin 20 times and recording the number of heads that occur.

b Rolling a six-sided die until the number 6 occurs.

c Drawing a card from a deck without replacement 10 times and recording the number of spades drawn.

d Drawing a marble from bag containing 8 black, 3 white and 5 pink marbles with replacement 12 times and recording the number of pink marbles drawn.

Solution

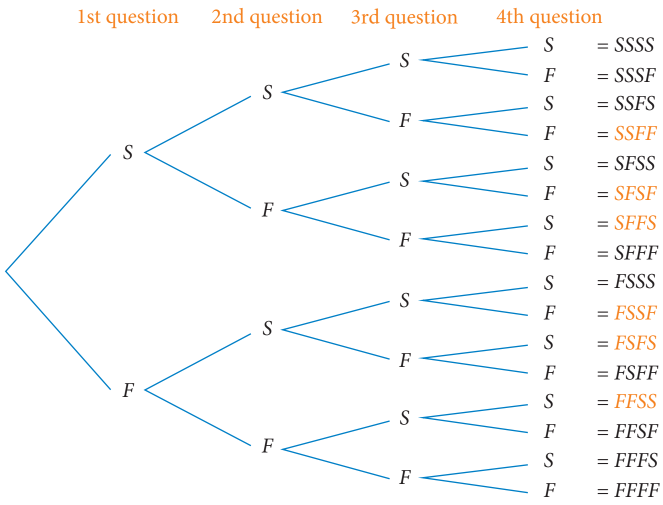
a	Are there a fixed number of trials?	There are 20 trials.
	How many outcomes are there?	There are 2 outcomes, with heads = success.
	Are the probabilities fixed?	$p = 0.5$ for every trial.
	Are the trials independent?	The trials are independent.
	What is the variable?	The variable is the number of successes.
	Write the conclusion.	This is a binomial experiment.
b	Are there a fixed number of trials?	The number of trials will vary, so it is not binomial.
c	Are there a fixed number of trials?	There are 10 trials.
	How many outcomes are there?	There are 2 outcomes, with spades = success.
	Are the probabilities fixed?	Cards are not replaced so p changes.
	Write the conclusion.	This is not binomial.
d	Are there a fixed number of trials?	There are 12 trials.
	How many outcomes are there?	There are 2 outcomes, with success = pink.
	Are the probabilities fixed?	$p = \frac{5}{16}$ for every trial
	Are the trials independent?	Replacement means the trials are independent.
	What is the variable?	The variable is the number of successes.
	Write the conclusion.	This is a binomial experiment.

Consider the experiment where a multiple choice test has four questions, each question has five alternate answers A, B, C, D and E and the answer to each is randomly selected. Only one answer is correct for any question.

This is a binomial experiment in which

- a success (S) is selecting the correct answer and $P(\text{success}) = p = \frac{1}{5}$
- a failure (F) is selecting the incorrect answer and $P(\text{failure}) = q = \frac{4}{5}$

You can construct a tree diagram to determine the various outcomes for this experiment.



This information can be summarised as follows.

Number of successes	Possible outcomes	Probability	
0	*FFFF*	$1 \times \left(\frac{4}{5}\right)^4 = \frac{256}{625}$	q^4
1	*SFFF, FSFF, FFSF, FFFS*	$4 \times \left(\frac{4}{5}\right)^3 \left(\frac{1}{5}\right) = \frac{256}{625}$	$4q^3p$
2	*SSFF, SFSF, SFFS, FSSF, FSFS, FFSS*	$6 \times \left(\frac{4}{5}\right)^2 \left(\frac{1}{5}\right)^2 = \frac{96}{625}$	$6q^2p^2$
3	*SSSF, SSFS, SFSS, FSSS*	$4 \times \left(\frac{4}{5}\right)\left(\frac{1}{5}\right)^3 = \frac{16}{625}$	$4qp^3$
4	*SSSS*	$1 \times \left(\frac{1}{5}\right)^4 = \frac{1}{625}$	p^4

The table above is the probability distribution for this binomial experiment. You can see that the values of the distribution are the expansion of $(p + q)^4$.

If you had to calculate the values of a binomial distribution using tree diagrams all the time, it would be extremely tedious. It is obvious from this example that you can use the expansion of $(p + q)^n$ instead. This means that you can use Pascal's triangle or combinations for the coefficients of the values.

Consider how many ways you can get 3 successes from 7 trials. You want to get 3 successes from the 7 that are possible, and the rest are failures. Since the order doesn't matter, there must be $^7C_3 = \binom{7}{3}$ ways of doing this. That means that the probability of 3 successes from 7 trials must be $\binom{7}{3} p^3q^4$. This will obviously work for any value of the binomial distribution.

9780170250290

IMPORTANT

Binomial distribution

If X is a binomial random variable, then the probability of x successes from n trials is given by

$$P(X=x)=\binom{n}{x}p^x q^{n-x} \quad \text{or} \quad P(X=x)=\binom{n}{x}p^x(1-p)^{n-x}$$

Example 8

a What is the probability of guessing which day of the week someone's birthday is this year?

b What is the probability of correctly guessing the day of the week of birthdays this year for 2 people from a group of 9 people?

c What is the probability of correctly guessing the day of the week of birthdays this year for 3 people from a group of 21 people?

Solution

a There are 7 possible days.

Probability of correct guess $=\frac{1}{7}$

b Write the details.

This is binomial with $n=9$, $p=\frac{1}{7}$ and $x=2$

Write the formula.

$$P(X=x)=\binom{n}{x}p^x q^{n-x}$$

Substitute in the values.

$$P(X=2)=\binom{9}{2}\left(\frac{1}{7}\right)^2\left(\frac{6}{7}\right)^{9-2}$$

Evaluate.

$$=\frac{36\times 6^7}{7^9}$$
$$=0.249\,734\ldots$$

State the answer.

The probability of guessing 2 out of 9 is about 0.2497.

c Write the details.

This is binomial with $n=21$, $p=\frac{1}{7}$ and $x=3$

Write the formula.

$$P(X=x)=\binom{n}{x}p^x q^{n-x}$$

Substitute in the values.

$$P(X=3)=\binom{21}{3}\left(\frac{1}{7}\right)^3\left(\frac{6}{7}\right)^{21-3}$$

Evaluate.

$$=\frac{1330\times 6^{18}}{7^{21}}$$
$$=0.241\,832\ldots$$

State the answer.

The probability of guessing 3 out of 21 is about 0.2418 or 24.18%.

Example 9

A binomial variable, X, has the probability function

$$P(X = x) = \binom{5}{x}(0.3)^x(0.7)^{5-x}$$

a What is the number of trials?
b State the probability of success in any trial.
c Show the probability distribution in a table.

Solution

a Examine the probability function.

The number of trials is $n = 5$

b Compare with the formula.

Probability of success is $p = 0.3$

c Substitute into the probability function for $x = 0$.

$$\begin{aligned} P(X = 0) &= \binom{5}{0}(0.3)^0(0.7)^{5-0} \\ &= 1 \times 1 \times 0.7^5 \\ &\approx 0.168\ 07 \end{aligned}$$

Substitute into the probability function for $x = 1$.

$$\begin{aligned} P(X = 1) &= \binom{5}{1}(0.3)^1(0.7)^{5-1} \\ &= 5 \times 0.3 \times 0.7^4 \\ &\approx 0.360\ 15 \end{aligned}$$

Continue and put the results in a table.

x	0	1	2	3	4	5
$p(x)$	0.1681	0.3602	0.3087	0.1323	0.0284	0.0024

EXERCISE 5.03 The binomial distribution

Concepts and techniques

WS Binomial probability experiments

1 Example 7 For which of the following situations would a binomial probability model be most reasonable?
A the number of phone calls made by a call centre operator in successive 10-minute periods
B the number of spades dealt in a hand of five cards from a well-shuffled deck of 52 cards
C the number of tosses of a fair coin required before two heads are observed
D the number of 9s in a randomly selected set of 10 digits from a table of random digits
E all of the above

2 There are 10 students in a school leadership group – 5 males and 5 females. The name of each student is written on a card. The cards are shuffled, one is randomly selected and the name on the card is observed. This is done a total of six times. If X is the number cards observed with the name of a female student, then which of these best describes the probability distribution for X?
A a uniform distribution with $n = 6$ and $p = 0.5$
B a binomial distribution with $n = 6$ and $p = 0.5$
C a binomial distribution with $n = 10$ and $p = 0.5$
D a geometric distribution with $p = 0.5$
E a binomial distribution with $n = 5$ and $p = \frac{6}{10}$

9780170250290

3 Example 8 In a certain binomial experiment, a successful outcome occurs 5 out of every 6 trials. What is the probability that there will be exactly 3 successful outcomes in the next 4 trials?

A $\left(\frac{5}{6}\right)^1\left(\frac{1}{6}\right)^3$ **B** $6\left(\frac{5}{6}\right)^1\left(\frac{1}{6}\right)^3$ **C** $\left(\frac{5}{6}\right)^3\left(\frac{1}{6}\right)^1$

D $6\left(\frac{5}{6}\right)^3\left(\frac{1}{6}\right)^1$ **E** $4\left(\frac{5}{6}\right)^3\left(\frac{1}{6}\right)^1$

4 Which of the following are examples of binomial experiments?

a Rolling a die 15 times and recording the number that comes up.

b Tossing a coin 10 times and recording the number of tails that occur.

c Rolling a six-sided die 20 times and counting the number of 6s that come up.

d Tossing a coin and counting the number of tosses required before two heads occur.

e Drawing a card from a deck with replacement 15 times and recording the number of hearts drawn.

f Drawing a marble from a bag containing 10 red, 4 green and 6 blue marbles without replacement 5 times and recording the number of red marbles drawn.

g Spinning a spinner numbered 1 to 8 ten times and counting the number of odd numbers that occur.

5 Evaluate the following, correct to 4 decimal places if necessary.

a $\binom{6}{4}(0.7)^4(0.3)^2$ **b** $\binom{9}{3}(0.38)^3(0.62)^6$ **c** $\binom{5}{2}\left(\frac{1}{5}\right)^2\left(\frac{4}{5}\right)^3$

d $\binom{8}{7}(0.25)^7(0.75)^1$ **e** $\binom{10}{0}(0.09)^0(0.91)^{10}$

6 Example 9 For each of the following probability distributions, identify the parameters p, q, n and x.

a $\binom{10}{2}(0.5)^2(0.5)^8$ **b** $\binom{20}{0}(0.85)^{20}$ **c** $\binom{15}{12}\left(\frac{3}{5}\right)^{12}\left(\frac{2}{5}\right)^3$

d $\binom{9}{8}(0.11)^8(0.89)$ **e** $\binom{7}{4}(0.25)^4(0.75)^3$

7 A binomial variable, X, has the probability function $P(X = x) = \binom{7}{x}(0.8)^x(0.2)^{7-x}$.
Find:

a the number of trials

b the probability of success in any trial

c the probability distribution as a table (with probabilities correct to four decimal places).

8 A binomial variable, Z, has the probability function $P(Z = z) = \binom{7}{z}(0.15)^z(0.85)^{7-z}$.
Find:

a the number of trials

b the probability of success in any trial

c the probability distribution as a table (with probabilities correct to four decimal places)

9 A fair six-sided die is rolled 7 times and the number of 4s is recorded. Find the probability of obtaining exactly three 4s.

Reasoning and communication

10 Compare the probability distributions found in questions **7** and **8** and comment on the differences you observe.

5.04 USING THE BINOMIAL DISTRIBUTION

Calculating binomial probabilities ‘by hand’ is very tedious. You can use your CAS calculator instead, unless you are required to give an exact answer.

Example 10

CAS Calculate the required probabilities in each of the following situations.

a A binomial experiment has 10 trials and $p = 0.35$. Find the probability that there are 3 successes obtained from the 10 trials.

b Find the binomial probability of 8 successes from 14 trials, where the probability of success is 0.75.

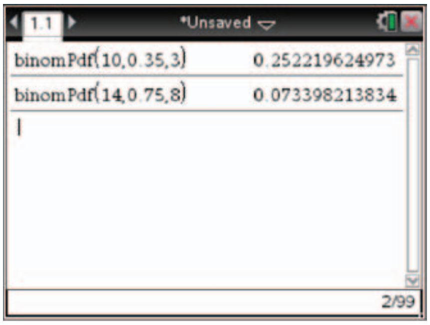
Solution

a List the parameters. $n = 10, p = 0.35, x = 3$

b List the parameters. $n = 14, p = 0.75$ and $x = 8$

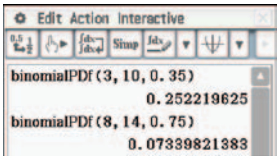
TI-Nspire CAS

Use a calculator page.
Use menu, 5: Probability, 5: Distributions and D: Binomial Pdf. Fill in the parameters.



ClassPad

Use the Main menu.
Tap Action, Distribution/Inv. Dist, Discrete and choose binomialPDf.
Complete the parameters in the order x, n, p.
You can do the same thing starting with Interactive. Fill in the table in the same order.



Write the answers.

a About 0.2522.
b About 0.0734.

9780170250290

Sometimes you need to find the probability of a number of values of a binomial distribution and add them together.

Example 11

A binomial experiment has 9 trials. The probability of success in any trial is 0.7. If X = number of successes, calculate the probability that:

a X is at most 2

b X is more than 2

c X is more than 6.

Solution

a List the parameters.	$n = 9, p = 0.7$
'At most 2' means 2 or fewer.	$P(X \le 2) = P(X = 0) + P(X = 1) + P(X = 2)$
Find the values.	$= 0.3^9 + 9 \times 0.7 \times 0.3^8 + 36 \times 0.7^2 \times 0.3^7$
Use your calculator.	$= 0.004\ 290\ 8\ldots$
Write the answer.	The probability that X is at most 2 is about 0.004 291.
b Use complements.	$P(X > 2) = 1 - P(X \le 2)$
Substitute in the value.	$= 1 - 0.004\ 290\ 8\ldots$ $= 0.995\ 709\ 1\ldots$
Write the answer.	The probability that X is more than 2 is about 0.9957.
c Write the probability.	$P(X > 6) = P(X = 7) + P(X = 8) + P(X = 9)$
Find the values.	$= 36 \times 0.7^7 \times 0.3^2 + 9 \times 0.7^8 \times 0.3 + 0.7^9$
Use your calculator.	$= 0.462\ 831\ldots$
Write the answer.	The probability that X is more than 6 is about 0.4628.

You can also use your CAS calculator to find combined binomial probabilities using the cumulative binomial distribution.

Example 12

CAS A binomial experiment has 20 trials. The probability of success in any trial is 0.48. The random variable X is the number of successes. Calculate the probability that X is between 5 and 9 inclusive.

Solution

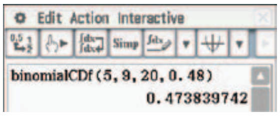
Identify the parameters. $n = 20, p = 0.48, 5 \le x \le 9$

TI-Nspire CAS

Use a calculator page.
Use menu, 5: Probability, 5: Distributions and E: 20Binomial Cdf. Fill in the parameters.

ClassPad

Use the Main menu.
Tap Action, Distribution/Inv. Dist, Discrete and choose binomialCdf. Complete the parameters in the order *lower, upper, n, p*.
You can do the same thing if you start with Interactive and fill in the table.



Write the answer. The probability that X is between 5 and 9 inclusive is about 0.4738.

You can always type the commands into the CAS calculators instead of negotiating the menus if you prefer to do it that way.

EXERCISE 5.04 Using the binomial distribution

Concepts and techniques

1 Example 10 In a particular binomial experiment, a success occurs in 5 out of every 6 trials. The probability that there will be exactly 5 successes in the next 6 trials is closest to:
A 0.0231 **B** 0.1667 **C** 0.3349 **D** 0.4019 **E** 0.8333

WS Using the binomial probability distribution

2 Example 11 X is a random variable with a binomial probability distribution, with $n = 7$ and $p = \frac{2}{5}$. The probability that X is at least 6 is:
A 0.0139 **B** 0.0188 **C** 0.0221 **D** 0.9959 **E** 0.9992

9780170250290

3 In a certain binomial experiment, the probability of success is 0.4. If there are five trials, what is the probability that at most one successful outcome results?

A 0.077 76 B 0.2592 C 0.3104 D 0.336 96 E 0.663 04

4 X is a random variable with a binomial probability distribution, with $n = 8$ and $p = \frac{1}{3}$. The probability that $X \geq 1$ is closest to:

A 0.0390 B 0.1951 C 0.8049 D 0.9610 E 0.9998

5 In a certain binomial experiment, a success occurs 75% of the time. If there are 10 trials, the probability that there are fewer than 3 successful outcomes is closest to:

A 0.000 415 B 0.000 416 C 0.003 090
D 0.003 506 E 0.004 218

6 A binomial experiment has 7 trials. The probability of success in any trial is 0.4. If X = the number of successes, calculate the probability that:

a $X = 3$ b X is at least 3 c X is more than 5

7 Example 12 CAS In a binomial experiment, $n = 11$, $p = 0.82$ and X = the number of successes. Calculate the probability that:

a $X = 7$ b X is between 3 and 6 (inclusive)
c X is less than or equal to 3 d X is greater than 7.

8 CAS A certain binomial experiment has 15 trials. The probability of success in any trial is 0.27. The random variable X is the number of successes. Calculate the probability that:

a $X = 11$ b X is at least 5
c X is greater than or equal to 9 d X is between 2 and 9 (exclusive)

9 CAS In a binomial experiment, $n = 25$, $p = 0.725$ and X = number of successes. Calculate the probabilities of the following.

a $X = 15$ b X is between 5 and 15 (inclusive)
c X is greater than or equal to 14 d X is less than 10

10 A binomial experiment has 8 trials and $p = 0.4$. Calculate the probability that:

a 2 successes occur b 5 successes occur
c at least 2 successes occur.

11 CAS A binomial experiment has 17 trials and the probability of success in any trial is 0.65. Calculate the probabilities of the following.

a 5 successes occur
b between 4 and 8 successes (inclusively) occur
c at least 14 successes occur

Reasoning and communication

12 A binomial experiment has 6 trials and the probability of success in any trial is 0.5. If X is the number of successes, find the value of x if:

a $P(X = x) = 0.3125$ b $P(X = x) = 0.093\ 75$ c $P(X = x) = 0.234\ 375$

13 A binomial experiment has 5 trials and the probability of failure in any trial is 0.5. If X is the number of successes, find the value of x if:

a $P(X < x) = 0.1875$ b $P(X \geq x) = 0.1875$ c $P(X > x) = 0.5$

5.05 PROPERTIES OF THE BINOMIAL DISTRIBUTION

You can find the values of a binomial distribution 'by hand' or by using your calculator. It is also useful to show binomial distributions as graphs. Even though the variable is discrete, column graphs of binomial distributions are often referred to as histograms.

Example 13

Draw a histogram of the binomial distribution for $n = 14$ and $p = 0.65$. Comment on the shape of the graph.

Solution

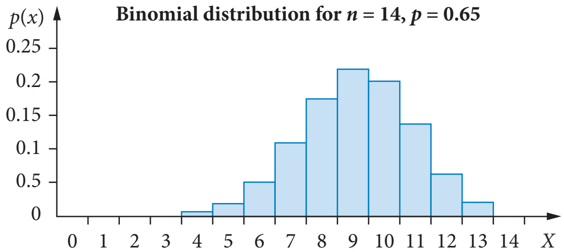
State the parameters.

$n = 14, p = 0.65, 0 \le X \le 14$

Use a your calculator or the formula to calculate the probabilities.
Many of the probabilities are too small to show on the graph.

$P(X = 0) = 0.000\ 000\ 413\ldots$
$P(X = 1) = 0.000\ 010\ldots$
$P(X = 2) = 0.000\ 129\ldots$
$P(X = 3) = 0.000\ 965\ldots$
$P(X = 4) = 0.004\ 92\ldots$
$P(X = 5) = 0.0183\ldots$
and so on.

Draw the histogram.



Comment on the shape.

The graph is skewed left.

You should remember that a **skewed** statistics or probability graph is stretched out more to one side that the other. In Example **13**, the highest point (the mode) is at 9 and the graph stretches from 0 to 9 on the left and from 9 to 14 on the right. It stretches a distance of 9 on the left but only 5 on the right, even though some of the probabilities are too small to see on the graph.

You can use your CAS calculator to investigate the shapes of binomial probability distributions.

9780170250290

INVESTIGATION The effect of changing n and p on a binomial distribution histogram

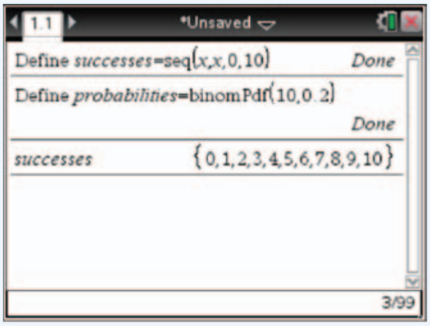
Part A

A binomial experiment has 10 trials and the probability of success in any trial is 0.2. The number of successful outcomes is observed.

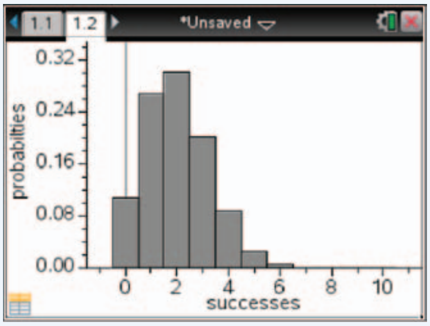
- What are the parameters for this experiment?
- Use a CAS calculator to draw a histogram of the probability distribution for this experiment.

TI-Nspire CAS

Use a Calculator page.
Type 'Define successes=seq(x,x,0,10)'.
Type 'Define probabilities=binompdf(10,0.2)
Type successes to see the list.

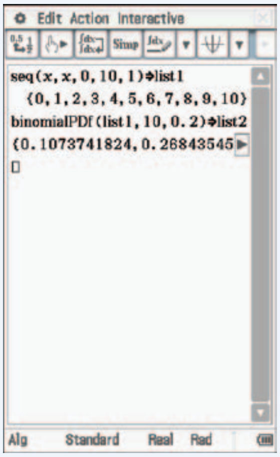


Add a Data & Statistics page.
Use menu, 2: Plot Properties and 5: Add X Variable and choose successes.
Use menu, 1: Plot Type and 3: Histogram and choose a histogram.
Finally use menu, 2: Plot Properties and 9: Add Y Summary List and choose probabilities.

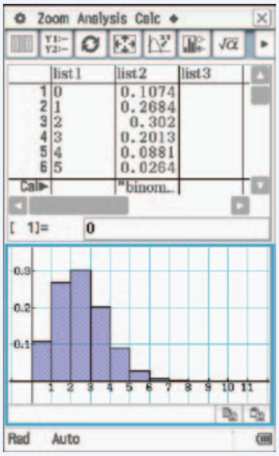


ClassPad

Tap Action, then List, Create and seq.
Enter x, x, 0, 10, 1 to create a list (sequence) of numbers from 0 to 10, going up by 1s. Name this list1.
Tap Action, Distribution/Inv. Dist, Discrete and binomialPDf.
Enter list1 for the number of successes, x, with $n = 10$ and $p = 0.2$. Name this list2.
It will create a list of probabilities for $x = 0$ to $x = 10$.



Now use the Statistics menu. The lists will already be in place.
Use the View Window ( ) to set $-1 \le x \le 11$, scale 1, and $-0.1 \le y \le 0.5$, scale 0.1.
Tap SetGraph, Setting and choose Histogram, List1 for XList and List2 for Freq.
Tap to draw the histogram.
Set Hstart to 0 and Hstep to 1.
You may have to do the View Window again.



A second binomial experiment also has 10 trials but the probability of success in any trial is 0.5. The number of successful outcomes is observed.

- What are the parameters for the second experiment?
- Use a CAS calculator to draw a histogram of the probability distribution for the second experiment.

A third binomial experiment also has 10 trials but the probability of success in any trial is 0.8. The number of successful outcomes is observed.

- What are the parameters for the third experiment?
- Use a CAS calculator to draw a histogram of the probability distribution for the third experiment.
- Compare the histograms for each of the binomial experiments and write a statement that describes the effect of changing p (the probability of success in any trial) on the shape of the histogram. What can you say about the graphs for $p = 0.2$ and $p = 0.8$?

Part B

A binomial experiment is performed in which the probability of success in any trial is 0.6. There are 6 trials and the number of successful outcomes is observed.

- What are the parameters for this experiment?
- Use a CAS calculator to draw a histogram of the probability distribution for this experiment.

This experiment is extended so that there are 20 trials and all other parameters are the same.

- Use a CAS calculator to draw a histogram of the probability distribution for the experiment.

Now the same experiment is extended to 50 trials.

- Use a CAS calculator to draw a histogram of the probability distribution for the experiment.
- Compare the three histograms for this binomial experiment. Write a statement that describes the effect of changing n (the number of trials) on the shape of the histogram.

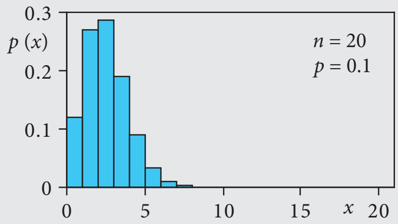
9780170250290

From the previous investigation you should have noticed a number of things about the shape of a probability distribution histogram.

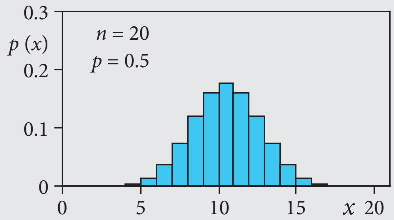
IMPORTANT

Shape of the binomial distribution

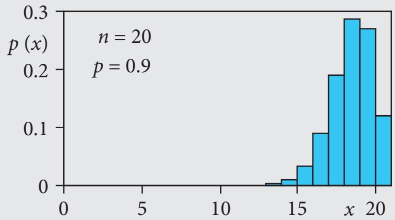
For $p < 0.5$, the distribution is **skewed to the right** (a **positive skew**).



For $p = 0.5$ the distribution is **symmetrical**.



For $p > 0.5$ the distribution is **skewed to the left** (a **negative skew**).



The most likely number of successes of a binomial distribution is the one with the highest probability. You can easily calculate the mean and variance of a binomial distribution using your CAS calculator. Remember that the expected value and variance are given by

$$\mu = E(X) = \sum x \cdot p(x) \qquad \text{and}$$

$$\begin{aligned} Var(X) &= E\left[(X-\mu)^2\right] \\ &= E\left(X^2\right) - \mu^2 \\ &= \sum x^2 \cdot p(x) - \left[\sum x \cdot p(x)\right]^2 \end{aligned}$$

Example 14

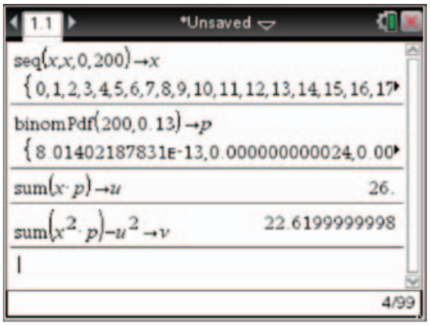
Find the expected value (mean) and variance of a binomial distribution with $n = 200$ and $p = 0.13$.

Solution

Identify the parameters. $n = 200$, $p = 0.13$

TI-Nspire CAS

Use a calculator page.
Store 0–200 in x and store the probabilities in p.
Use sum($x \times p$) to find the expected value, u.
Use sum($x^2 \times p$) – u^2 to find the variance, v.
You could also do this as a spreadsheet, as shown in the Casio.



ClassPad

This is done in a similar way to the earlier investigation on pages 221–222.
Tap Action, then List, Create and seq.
Enter x, x, 0, 200, 1 to create a list (sequence) of numbers from 0 to 200, going up by 1s.
Name this list1.
Tap Action, Distribution/Inv. Dist, Discrete and binomialPDf.
Enter list1 for the number of successes, x, with $n = 200$ and $p = 0.13$. Name this list2.
It will create a list of probabilities from $x = 0$ to $x = 200$.
Go to the Statistics menu.
Tap Calc and One-Variable.
Make XList list1 and Freq List2.
Read the value of the mean $\bar{x}$ and the standard deviation σ_x.

The variance is the square of the standard deviation.

$\text{Variance} \approx 4.756...^2$
≈ 22.62

Write the answer. The mean is 26 and the variance is 22.62.

Notice that there appears to be a rounding error in the calculation of the variance on the TI-Nspire in Example **14**. This does illustrate that a CAS calculator is a computer so it has limited storage. This can sometimes produce inaccurate or incorrect answers.

Notice also that the expected value for the binomial distribution with $n = 200$ and $p = 0.13$ is $200 \times 0.13 = 26$.

This follows from the fact that the binomial distribution is a fixed number of Bernoulli trials, each one of which constitutes a Bernoulli distribution. The expected value of a Bernoulli distribution is p, so the expected value of the binomial distribution will just be the sum of the expected value of each of the n trials, np.

Similarly, the variance of the Bernoulli distribution is $p(1-p)$, so the variance of the binomial distribution will be $np(1-p)$.

The standard deviation is just the square root of the variance.

IMPORTANT

The **expected value** (**mean**) of a binomial probability distribution is given by $\mu = E(X) = np$.

The **variance** of a binomial probability distribution is given by $Var(X) = np(1-p)$.

The **standard deviation** of a binomial distribution is given by $\sigma = SD(X) = \sqrt{np(1-p)}$.

You may prefer to use $q = 1 - p$ and write $\mu = np$, $Var(X) = npq$ and $\sigma = \sqrt{npq}$.

Example 15

In the general population, 15% of people are left-handed. A class has 24 students and the number who are left-handed is counted.

a Confirm this is an example of a binomial experiment.

b How many in the class would you expect to be left-handed?

c What is the probability that the number of left-handed students in the class is within one standard deviation of the mean?

Solution

a	Check this situation to see if it has the characteristics of a binomial distribution.	There are 24 independent trials with 2 outcomes, p is fixed and the variable is the number of successes.
	State the result.	This is a binomial distribution.
b	Write the formula for expected value	$E(X) = np$
	Substitute in the values and find the answer.	$= 24 \times 0.15$ $= 3.6$
	State the result.	You would expect 3 or 4 left-handers.
c	Write the formula for σ.	$SD(X) = \sqrt{npq}$
	Substitute in the values.	$= \sqrt{24 \times 0.15 \times 0.85}$
	Evaluate.	≈ 1.75
	Calculate the values of X within 1 standard deviation (σ) of the mean (μ).	$\mu + \sigma \approx 3.6 + 1.75 = 5.35$ $\mu - \sigma \approx 3.6 - 1.75 = 1.85$
	Identify the relevant values of X.	Only the values from 2 to 5 are *within* 1 SD.

State the required probability.	$P(\mu - \sigma \le X \le \mu + \sigma) = P(2 \le X \le 5)$
Use the CDF on your calculator.	$= 0.754\ 637\ldots$
State the result.	There is about a 75% chance that the number of left-handed students in the class is within one standard deviation of the mean.

Even if the values of $\mu - \sigma$ and $\mu + \sigma$ worked out to 1.000 000 001 and 5.999 999 999, you would use $P(2 \le X \le 5)$ in Example **15** because it had to be *within* one standard deviation. It is always very important to read a question carefully to make sure you are answering the question asked.

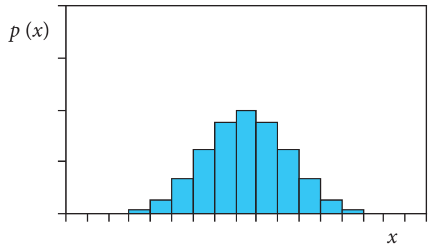
EXERCISE 5.05 Properties of the binomial distribution

WS Binomial probability – Mean and standard deviation

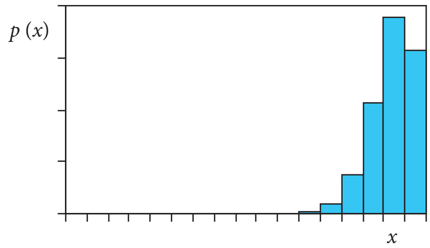
Concepts and techniques

1 Example 13 Which of the following graphs best represents the shape of a binomial probability distribution of the random variable X with 12 trials if the probability of success is 0.3?

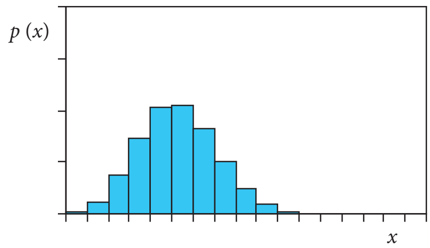
A



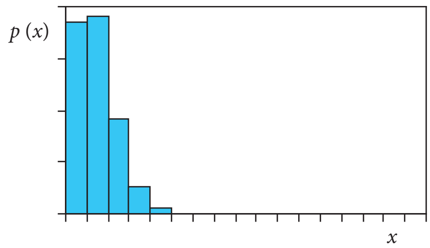
B



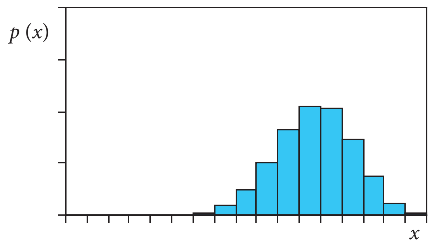
C



D



E



2 Example 15 If a binomial random variable, X, has parameters $n = 32$ and $p = \frac{1}{4}$, then the mean and variance of X are closest to:

A $\mu = 6, \sigma^2 = 8$ B $\mu = 24, \sigma^2 = 12$ C $\mu = 16, \sigma^2 = 8$
D $\mu = 8, \sigma^2 = 6$ E $\mu = 12, \sigma^2 = 24$

3 a Draw graphs of the binomial distributions for $n = 8$ for $p = 0.2$ and $p = 0.8$.
b Comment on the shapes of the graphs.

4 a Draw graphs of the binomial distributions for $n = 10$ for $p = 0.4$ and $p = 0.6$.
b Comment on the shapes of the graphs.

5 In a binomial experiment, the probability of success in any trial is 0.8. There are 10 trials.
a Draw the histogram of the probability distribution for the number of successes.
b Use the histogram to determine the most likely number of successes.

6 Example 14 CAS For each of the following binomial distributions, calculate the mean (expected value) and standard deviation without using the formulas. Use your results to verify the formulas.
a $n = 5$ and $p = 0.5$ b $n = 8$ and $p = 0.3$
c $n = 12$ and $p = 0.8$ d $n = 9$ and $p = 0.6$

7 Use the formulas to find the mean and standard deviation of each of the following binomial distributions.
a $n = 7$ and $p = 0.1$ b $n = 7$ and $p = 0.9$
c $n = 20$ and $p = 0.65$ d $n = 30$ and $p = 0.34$

Reasoning and communication

8 A binomial experiment has a random variable, X. If there are 10 trials and $Var(X) = 0.9$, what is the probability of success?

9 A random binomial variable X has a mean of 24 and a standard deviation of 3. What is the probability of success?

10 A normal deck of playing cards is shuffled and cut. This is done 60 times. The number of aces that are cut is noted.
a Confirm this is an example of a binomial experiment.
b How many aces would you expect to get?
c What is the probability that the number of aces cut is within two standard deviations of the mean?

11 The mean of a binomial distribution is 5 and the standard deviation is 2. What is the actual probability of 5 successes?

12 About how many heads would you expect from 200 tosses of a fair coin?

13 The probability that a lettuce seed germinates is about 0.8. About how many seeds should you plant if you want 500 lettuce seedlings?

14 The probability that a car battery fails within 2 years of purchase is about 60%.
a From 420 batteries, how many would you expect to last longer than 2 years?
b What is the standard deviation of the number of failures in 2 years?

5.06 APPLICATIONS OF THE BINOMIAL DISTRIBUTION

There are many practical applications of the binomial probability distribution. The binomial probability distribution is often used in business and government to help make decisions.

Example 16

A study found that 38% of all swordfish sold in a particular seafood market contained levels of mercury that were above the Food Standards Australia (FSA) recommended maximum levels. Food inspectors took a random sample of twelve pieces of swordfish from the market for testing. Find the probability that:

Alamy/Gianni Muratore

- a five of the pieces have mercury levels above the FSA maximum
- b at most four pieces have mercury levels above the FSA maximum
- c at least three pieces have mercury levels above the FSA maximum.

Solution

a Check this situation to see if it has the characteristics of a binomial distribution.

There are 12 independent trials with 2 outcomes (above level or not), $p = 0.38$ is fixed and the variable is the number of "successes", so it is binomial.

State the required probability.

$$P(X = 5) = \binom{12}{5} p^5 q^7$$

Substitute and calculate or use the PDF.

$$= \binom{12}{5} (0.38)^5 (0.62)^7$$

$$= 0.220\,996\ldots$$

State the answer.

The probability of 5 bad fish is about 22%.

b State the required probability.

$$P(\text{at most } 4) = P(0 \le X \le 4)$$

Use the CDF.

$$= 0.495\,719\ldots$$

State the answer.

The probability of at most 4 bad fish is about 49.6%.

9780170250290

c State the required probability.	$P(\text{at least } 3) = P(3 \le X \le 12)$
Use the CDF.	$= 0.893\,056\ldots$
State the answer.	The probability of at least 3 bad fish is about 89.3%.

FSA say consumers should eat not more than 1 serve of large predatory fish like swordfish or shark (flake) in a week, and no other fish or shellfish at all the same week.

Example 17

At a particular telemarketing company, it is known that the probability that a call to a potential client (cold calling) results in a sale is 0.05. What is the least number of calls that must be made to ensure that the probability of making at least 2 sales is 90%?

Solution

Confirm that this is an example of a binomial experiment.

There are 2 outcomes (sale or not), p is fixed and the variable is the number of "successes". The calls are independent. Determining the number of calls will make it binomial.

State the required probability.

$$P(X \ge 2) > 0.9$$

This is easier with the complement.

$$1 - P(X < 2) > 0.9$$

Simplify.

$$P(X < 2) > 0.1$$

Rewrite the LHS.

$$P(X < 2) = P(X = 0) + P(X = 1)$$

Use the binomial function with n as the number of trials.

$$= \binom{n}{0}(0.05)^0(0.95)^n + \binom{n}{1}(0.05)^1(0.95)^{n-1}$$

Substitute $\binom{n}{0} = 1$ and $\binom{n}{1} = n$.

$$= 0.95^n + n \times 0.05(0.95)^{n-1}$$

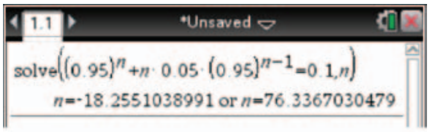
Write what needs to be solved.

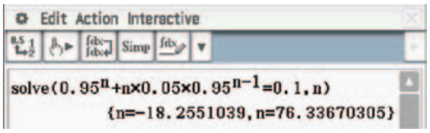
$$0.95^n + n \times 0.05(0.95)^{n-1} < 0.1$$

Write this as an equation.

$$0.95^n + n \times 0.05(0.95)^{n-1} = 0.1$$

This equation cannot be solved algebraically. Use the Solve function of your CAS calculator. The syntax is the same for both calculators shown, and they both give warnings that there may be other solutions.





n cannot be negative.

$$n = 76.336\ldots$$

State the result.

At least 77 calls must be made to ensure the probability of making at least 2 sales is 90%.

You CAS calculator gives a warning because they use numerical methods to find a solution by iteration. Since it is not an algebraic method, there is no guarantee that there are no other solutions. In real situations, you would check that the solution does work.

Example 18

It is known that about 8% of all people in Australia will exceed their recommended dose of caffeine from drinking coffee or energy drinks. A company employs 200 staff at a factory and the manager wants to know how the workforce might be affected, given that they have vending machines for coffee, tea and energy drinks. Some of the effects of too much caffeine are headaches, restlessness and inattention.

Calculate the probability that the number of workers who will get too much caffeine is between two standard deviations of the mean and explain what this result means for the factory.

iStockphoto/NoDerog

Solution

Check this situation to see if it has the characteristics of a binomial distribution.	There are 200 trials with 2 outcomes (overdose or not), p is fixed and the variable is the number of "successes", so it is binomial.
Calculate the mean.	$\mu = np$
Substitute and evaluate.	$= 200 \times 0.08$ $= 16$
Calculate the standard deviation.	$\sigma = \sqrt{npq}$
Substitute and evaluate.	$= \sqrt{200 \times 0.08 \times 0.92}$ $= 3.836\,665\ldots$
Calculate the X values that are two standard deviations of either side of the mean.	$\mu + 2\sigma = 16 + 2 \times 3.836\ldots = 23.673\ldots$ $\mu - 2\sigma = 16 - 2 \times 3.836\ldots = 8.326\ldots$
State the required probability.	$P(\text{within } 2\sigma) = P(8 \le X \le 23)$
Use your CAS calculator.	$= 0.961\,419\ldots$
State the result.	There is about a 96% chance that from 8 to 23 staff will overdose with caffeine. Get rid of the vending machines and educate staff about the dangers!!

9780170250290

EXERCISE 5.06 Applications of the binomial distribution

WS Applications of binomial probability

Concepts and techniques

In this exercise, you are expected to use your CAS calculator.

1 Example 16 A certain plant variety produces 25% red flowers and the flower colour red forms a binomial distribution. One of these plants has 3 flowers.
  a What are the values of n, p and q?
  b What is the probability that all flowers are red?
  c What is the probability that no flowers are red?
  d What is the probability that at least 2 flowers are red?

2 The probability that the Melbourne Storm win is $\frac{2}{3}$ whenever they play. Find the probabilities of winning the following numbers of games from the first 4 of the season.
  a Exactly 2 games.
  b At least 1 game.
  c More than half of the games.

3 A family has 6 children. Find the probabilities of the following family compositions.
  a 3 boys and 3 girls
  b Fewer boys than girls.

4 What proportion of families with exactly 4 children would be expected to have at least 2 girls?

5 Metal welds that are done with robotic devices have a 15% chance of containing a flaw. A pipeline contains 7 welds. Find the probability that:
  a all the welds have flaws
  b at least 1 weld has a flaw.

iStockphoto/Dreef

6 It is a bank's experience that 5% of the cheques received in an automatic teller will 'bounce'. What is the probability that, for 8 cheques deposited
  a none will bounce?
  b exactly 2 will bounce?
  c at least 1 will bounce?

7 If, over a given period in Brisbane, rain falls at random on 4 days out of every 10, find the probability that

a the first 2 days of a given week will be wet and the remainder of the week will be fine
b rain will fall on the first 2 days of the week
c at least 2 days in the week will be wet.

8 A multiple-choice test has 7 questions, each with 4 alternatives, only 1 of which is correct. What is the probability of guessing correct answers to

a all of the 7 questions?
b none of the 7 questions?
c 3 of the 7 questions?
d at least 3 of the 7 questions?

9 A gardener plants 8 seeds. The probability that a seed will germinate is 0.85. What is the probability that at least 6 of the seeds will germinate?

10 A particular family has 3 children. It is known that the genetic makeup of the parents is such that there is a 20% chance that a child born will have curly hair.

a What are the values of n, p and q for this binomial situation?
b What is the probability that all of the children have curly hair?
c What is the probability that none of the children have curly hair?
d What is the probability that at least 1 of the children has curly hair?

Reasoning and communication

11 Example 17 The probability of Jacqui hitting a target is $\frac{1}{4}$.

a If she fires 7 times, what is the probability of her hitting the target at least twice?
b How many times must she fire so that the probability of hitting the target at least once is greater than $\frac{2}{3}$?

12 A survey revealed that 63% of all people in a particular city eat fast food at least once a week. A random sample of 6 people in the city was selected. What is the probability that:

a half of those selected ate fast food at least once a week?
b at least half of those selected ate fast food at least once a week?

13 Example 18 It is known that 68% of all school students in a certain Australian state attend government schools. A new community is planned for establishment in an area near the state's capital city. It is estimated that the community will eventually have 1500 school-aged children. Calculate the probability that the number of students who will attend a government school is between two standard deviations of the mean and explain what this result means.

14 Hospital records confirm that, of patients suffering from a particular complaint, 80% recover during their hospital stay. What is the probability that 6 randomly selected patients suffering from the complaint will all recover?

15 A judge is scheduled to hear 10 appeals for traffic violations. Each appeal has a probability of 0.4 of being approved, independently of other appeals. Find the probability that less than half of the appeals are granted.

9780170250290

16 Boom Boom is the school's best tennis player. From experience, it is known that the probability that Boom Boom serves an ace is 0.15. What is the probability that Boom Boom wins a game with just 4 serves?

17 It is known that 10% of the population carry a particular gene that predisposes them to bladder cancer.

a A study group of 10 people is randomly selected for testing. Find the probability that:

i two people in the group possess the gene

ii at least one person in the group possesses the gene.

b What is the smallest number of people who should be selected for a study group to ensure that there is more than a 70% chance of having at least one person who possesses the gene?

18 After a certain drug in tablet form is stored above 30°C, it is found that an average of one-fifth of the tablets become ineffective. A person who has been on a holiday in the tropics in summer kept 8 of these tablets in a suitcase where the temperature was consistently over 30°C. What is the probability that:

a exactly 3 of the tablets are now ineffective?

b at least 1 of the tablets is now ineffective?

19 Find the probability that, of the first 6 people met in the street on a given day, at least 4 will have their birthday on a Sunday this year. Give your answer as a fraction.

20 It is known that 30% of the components manufactured in a particular assembly line will have some kind of defect.

a A random sample of 6 components is selected for testing. Find the probability that:

i none has a defect ii at least one has a defect of some kind.

b What is the smallest number of components that should be selected to ensure that there is more that an 80% chance of having at least one defective component?

21 A manufacturer finds that, in the long run, 15% of the manufactured articles are defective. If a sample of 10 articles is randomly selected, find the probability that:

a the sample contains 2 defective articles

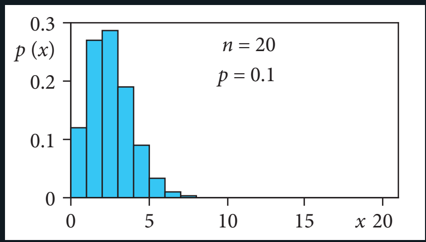
b the first 3 articles selected are defective

c the sample contains at least 3 defective articles.

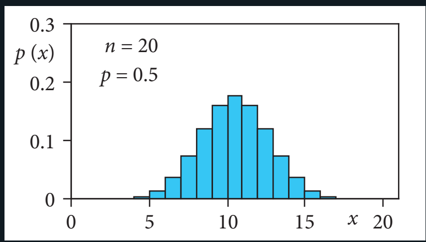
22 Wendy owns 10 holiday units in a remote location and does not want to risk leaving DVD players permanently in the units, so she hires them out to occupants as they require them. She finds that about 80% of people who occupy the units want to hire a DVD player, so she buys 7. Find the probability that, at a time when all the units are occupied, the demand for DVD players will exceed supply.

5 CHAPTER SUMMARY
BINOMIAL DISTRIBUTIONS

- A **Bernoulli distribution** is a distribution with two possible outcomes, called **success** and **failure**, with a fixed probability of success. The probability of success is normally written as p and the probability of failure as q, so that $p + q = 1$.
- The **Bernoulli random variable** X is given the values $P(X = 0) = q$, $P(X = 1) = p$.
- The expected value, variance and standard deviation of a Bernoulli distribution are given by $E(X) = \mu = p$, $Var(X) = p(1 - p)$ and $SD(X) = \sigma = \sqrt{p(1-p)}$
- **Bernoulli trials** are independent trials that have only two possible outcomes, called **success** and **failure**, where all the trials have the same probability of success. The probability of success is normally written as p and the probability of failure as q.
- The **geometric distribution** is the probability distribution of the number of failures of Bernoulli trials before the first success.
- The **geometric random variable** X has the values 0, 1, 2, 3, …: $X \in J^+$, the probability function is $P(X = x) = q^x p$ and the expected value is $E(X) = \frac{q}{1-q} = \frac{1-p}{p}$.
- The **binomial distribution** is the probability distribution of the number of successes arising from a fixed number of Bernoulli trials. A situation that produces a binomial distribution may be called a **binomial experiment**.
- A binomial distribution has the following properties:
  - each trial has only two outcomes – success and failure with probabilities p and q
  - there are a fixed number, n, of independent trials with a fixed probability of success
  - and the **binomial random variable** is the number of successes from n trials.
- If X is a binomial random variable, then the probability of x successes from n trials is given by $P(X = x) = \binom{n}{x} p^x q^{n-x}$, or $P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$.
- For $p < 0.5$, the binomial distribution is **skewed to the right** (a **positive skew**).

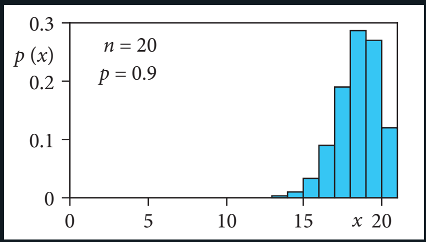


- For $p = 0.5$, the distribution is **symmetrical**.



9780170250290

- For $p > 0.5$, the distribution is **skewed to the left** (a **negative skew**).



- The **expected value** (**mean**) of a binomial probability distribution is $\mu = E(X) = np$.
- The **variance** of a binomial probability distribution is given $Var(X) = np(1 - p)$.
- The **standard deviation** of a binomial distribution is given by $\sigma = SD(X) = \sqrt{np(1-p)}$.

5 CHAPTER REVIEW
BINOMIAL DISTRIBUTIONS

Multiple choice

1 Example 1 Which of the following can be considered as a Bernoulli distribution?

A Tossing a coin 5 times and counting the number of heads
B Rolling a die and seeing if it comes up with a 6 on top
C Cutting cards and checking the suit, with replacement
D All of the above
E None of the above

2 Example 3 A Bernoulli distribution has a probability of success of 0.2. What is the standard deviation?

A 0.04 B 0.16 C 0.4 D 0.8 E 1.6

3 Example 4 Which of the following variables will have a geometric probability distribution?

A the number of times a coin must be flipped until 2 successive heads have been flipped
B the number of red cars sold each week by a car dealership
C the number on which the ball rests when a roulette wheel is spun
D the number of cards drawn from a well-shuffled deck of playing cards before an ace occurs.
E the number of cars that are able to pass through a set of lights at an intersection while the light is green

4 Example 5 X has a geometric distribution, where X is number of failures before the first success. If the probability of success is 0.4, then $P(X = 3)$ is equal to:

A 0.0384 B 0.0576 C 0.0684 D 0.0834 E 0.0864

5 Example 7 For which of the following situations would a binomial probability model be most reasonable?

A the number of hearts that occur when 7 cards are drawn with replacement from a well-shuffled deck of 52 cards
B the number of tosses of a fair coin required before a head occurs
C the number of blue marbles that are observed when 8 marbles are drawn without replacement from a bag containing 4 blue, 6 white and 3 green marbles
D the number of spades dealt in a hand of five cards from a well-shuffled deck of 52 cards
E all of the above

6 Example 9 A card is drawn from a well-shuffled deck of 52 playing cards and its suit is noted before it is returned to the deck. This process is repeated 12 times. If X is the number of hearts observed, then which of the following best describes the probability distribution for X?

A a uniform distribution with $n = 12$ and $p = 0.25$
B a binomial distribution with $n = 52$ and $p = 0.5$
C a geometric distribution with $n = 12$ and $p = 0.75$
D a hypergeometric distribution with $k = 4$ and $n = 12$
E a binomial distribution with $n = 12$ and $p = 0.25$

9780170250290

7 Example 11 In a certain binomial experiment the chances of a successful outcome are 5 in every 12 trials. If there are 7 trials, what is the probability that at most one successful outcome results?
A 0.0149 B 0.1379 C 0.1793 D 0.3842 E 0.4167

8 Example 12 A binomial experiment has 20 trials. If the probability of success in any trial is 90%, what is the probability of obtaining 18 or more successful outcomes:
A 0.190 B 0.270 C 0.285 D 0.677 E 0.900

9 Example 14 A binomial probability distribution has 20 trials and the probability of success is 0.7. The expected value is:
A 7 B 10 C 12.5 D 14 E 17

Short answer

10 Example 1 State which of the following describe a Bernoulli distribution.
- a Drawing a card from a deck with replacement and recording its suit.
- b Rolling a six-sided die 15 times and recording the number of times a prime number comes up.
- c Spinning a spinner with equal-sized segments coloured red, green, pink, blue, yellow, orange, black and white and recording the whether the spinner stops on orange.
- d Tossing a coin 15 times and recording the number of tails.
- e Drawing a card from a deck without replacement and recording if a king is drawn.
- f Drawing a marble from a bag containing 6 red and 3 blue marbles with replacement and recording the number of blue marbles drawn.
- g Observing whether a person signing a contract is left-handed.

11 Examples 2, 8 Six balls and then two more called supplementaries drawn from a tumbler with balls numbered 1 to 45 without replacement are used for a lotto game. What is the probability of successfully guessing one of the numbers drawn?

12 Example 3 When you are playing Ludo, what is the exact variance of the random variable X arising from rolling a normal die if you want to get a six to start?

13 Example 4 A fair six-sided die is rolled until the first time that a number less than 3 is observed. What is the probability that exactly five rolls are required?

14 Example 5 A putt-putt golf enthusiast keeps careful records and knows that, in the long run, she completes a hole in three putts or less 30% of the time. She regards this as a 'win'. If she takes more than three putts, she considers this to be a loss. If X represents the number of holes she loses before the first win on a particular round of putt-putt, find:
- a $P(X = 3)$
- b $P(X \leq 5)$
- c $E(X)$


Thinkstock/Karia Yekareva

15 Example 6 An experienced rugby union player knows that he converts 68% of all kicks for goal. If Y represents the number of kicks before his first successful conversion, find:

a $P(Y = 3)$ b $P(2 \le Y \le 9)$

16 Example 7 Which of the following are examples of binomial experiments?

a Spinning a spinner numbered 1 to 12 five times and counting the number of prime numbers that occur.

b Rolling a six-sided die 20 times and recording the number of rolls required before an odd number occurs.

c Drawing a card from a deck with replacement 10 times and recording the number of spades drawn.

d Tossing a coin and counting the number of tosses required before two tails occur.

e Rolling a six-sided die 20 times and counting the number of even numbers that come up.

f Drawing a marble from bag containing 9 green, 3 red and 2 blue marbles without replacement 6 times and recording the number of green marbles drawn.

g Tossing a coin 12 times and recording the number of heads that occur.

17 Example 8 Evaluate the following, correct to 4 decimal places if necessary.

a $\binom{5}{2}(0.4)^2(0.6)^3$ b $\binom{10}{7}(0.43)^7(0.57)^3$ c $\binom{8}{2}\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^6$

18 Example 9 A binomial variable, X, has the probability function:

$$P(X = x) = \binom{8}{x}(0.45)^x(0.55)^{8-x}.$$

Find:

a the number of trials

b the probability of success in any trial

c the probability distribution as a table (with probabilities correct to four decimal places).

19 Example 9 A binomial experiment has 8 trials. The probability of success in any trial is 0.35. If X = number of successes, calculate the probability that:

a $X = 7$ b X is at least 2 c X is more than 4.

20 Example 10 A card is drawn from a well-shuffled deck, checked to see whether it is a picture card (K, Q or J) and then returned to the deck. What is the probability of obtaining 2 picture cards from 5 such draws?

21 Example 11 A certain binomial experiment has 20 trials. The probability of success in any trial is 0.68. The random variable X is the number of successes. Use a graphics calculator to find the probability that:

a $X = 12$ b X is at least 7 c X is at most 9.

22 Example 11 A binomial experiment has 9 trials and $p = 0.6$. Calculate the probability that:

a 3 successes occur b 6 successes occur c at least 2 successes occur.

23 Example 11 In a binomial experiment, the probability of success is 0.7. What is the probability of getting at least 2 successes if there are three trials?

9780170250290

24 Examples 11, 12 CAS Find the following binomial probabilities.

a $P(x = 5)$ when $n = 15$ and $p = 0.3$

b $P(x = 7)$ when $n = 19$ and $p = 0.45$

c $P(x < 6)$ when $n = 18$ and $p = 0.75$

d $P(x \geq 4)$ when $n = 12$ and $p = 0.6$

e $P(3 \leq x \leq 9)$ when $n = 20$ and $p = 0.3$

25 Example 12 A normal coin is tossed 12 times. What is the probability of obtaining more than 3 but less than 7 heads?

26 Example 13 Draw a histogram of the binomial probability distribution for $n = 8$ and $p = 0.65$. Comment on the shape of the distribution.

27 Example 13 In a binomial experiment, the probability of success in any trial is 0.4. There are 8 trials.

a Draw the histogram of the probability distribution for the number of successes.

b Use the histogram to determine the most likely number of successes.

28 Example 14 Find the mean and standard deviation of each of the following binomial distributions.

a $n = 8$ and $p = 0.2$

b $n = 10$ and $p = 0.7$

c $n = 15$ and $p = 0.55$

29 Example 15 The probability that children will be absent from school in a flu epidemic is 15%. What is the probability that from a class of 20 students:

a 7 will be away?

b at least 7 will be away?

c at least 5 will be away?

30 Example 16 It is known that 28% of eggs placed in cartons are slightly underweight. A carton holding a dozen eggs is selected for examination. Find the probability that:

a the first 2 eggs selected are slightly underweight and the remainder are not underweight

b 2 of the eggs in the carton are slightly underweight

c at least 2 of the eggs in the carton are slightly underweight.

Shutterstock.com/Chimpinski

Application

31 A target shooter knows from experience that she hits a target 4 times out of each 5 shots. If she is allowed 10 shots in a competition, what is the probability that she gets:

a 10 hits? b 9 hits? c at least 7 hits?

Shutterstock.com/vvicha

32 A new vaccine is claimed to be 85% effective in immunising young children against a particular childhood disease. In a sample of 15 children who are vaccinated, what is the probability that:

a all become immune to the disease?

b at least 2 have no immunity to the disease?

c fewer than 4 have no immunity?

33 Five per cent of thermostats produced by a particular process are defective. The thermostats are packed 15 to a box for shipping purposes. When a box of thermostats is received by a spare parts dealer, what is the probability that:

a all are operative? b 2 are defective?

c at least 2 are defective? d no more than 2 are defective?

34 There is a 30% chance that a driver will have an accident in their first year of driving. Eighteen Year 12 students get their driver's licenses in June. What is the probability that less than a quarter will have an accident before June the next year?

35 In the general population, it is estimated that 18% of women suffer from iron deficiency. A study group of 25 women is selected and tested for iron deficiency.

a How many of the study group would you expect to be iron deficient?

b Calculate the probability that the number of women with iron deficiency in the study group is within two standard deviations of the mean.

c Explain what the result in part **b** means.

9780170250290

36 The mean of a binomial distribution is 12 and the standard deviation is 3. What is the probability of 9 successes?

37 It is known that 18% of the residents in a large city object to the fluoridisation of the city's water supply.

a A random sample of 10 residents is selected. Find the probability that:

i none has an objection to water fluoridisation

ii at least two object to water fluoridisation.

b What is the smallest number of residents that should be selected to ensure that there is more that a 90% chance of having at least one who objects to water fluoridisation?

Practice quiz