THE GEOMETRIC DISTRIBUTION

The geometric distribution

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

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

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?

The trials are Bernoulli trials with a sequence *FFFFS*.

$p = P(3 \text{ or } 6) = \frac{2}{8} = \frac{1}{4}$ $\qquad q = 1 - \frac{1}{4} = \frac{3}{4}$

$$
\begin{aligned}
P(FFFFS) &= \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} \times \frac{1}{4} \\
&= \frac{3^4}{4^5} \\
&= \frac{81}{1024} \\
&= 0.079\,101\ldots
\end{aligned}
$$

The probability of 5 rolls is about 0.079 or 7.9%.

The pattern in the previous example is termed a **Geometric Distribution.**

Note: unlike the other distributions this distribution is **not finite.**

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

The probability function is $P(X = x) = q^x p \;\; X$ has the values $0, 1, 2, 3, \ldots X \epsilon J^+$

The probability function is a geometric sequence.

Recall the infinite sum of a Geometric Series with a common ratio r is given by $\frac{a}{1-r}$ provided $-1 \leq r \leq 1$

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

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

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

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

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

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

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

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

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

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

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

$$S_n = \frac{a(1 - r^6)}{1 - r} = \frac{\frac{1}{4}\left[1 - \left(\frac{3}{4}\right)^6\right]}{1 - \frac{3}{4}} = 1 - \left(\frac{3}{4}\right)^6 \approx 0.8220$$

$$E(X) = \frac{1 - p}{p} = \frac{1 - \frac{1}{4}}{\frac{1}{4}} = \frac{3}{4} \times \frac{4}{1} = 3$$

$E(X) = 3$. There will be an average of 3 bottles labelled correctly before the first incorrectly labelled bottle occurs.

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 $P(2 \le Y \le 7)$

$p = 60\% = 0.6$ and $q = 1 - 0.6 = 0.4$

$P(Y = y) = (0.4)^{y-1}(0.6)$

$$\begin{aligned} P(Y = 8) &= (0.4)^{8-1}(0.6) \\ &= (0.4)^{7}(0.6) \\ &\approx 0.000\,98 \end{aligned}$$

$$\begin{aligned} P(2 \le Y \le 7) &= P(Y = 2) + P(Y = 3) + P(Y = 4) + \cdots + P(Y = 7) \\ &= (0.4)(0.6) + (0.4)^2(0.6) + (0.4)^3(0.6) + \cdots + (0.4)^6(0.6) \end{aligned}$$

This is a GP with $a = (0.4)(0.6)$, $r = 0.4$ and $n = 6$

$$\begin{aligned} S_n &= \frac{a(1-r^6)}{1-r} \\ &= \frac{(0.4)(0.6)\left[1-(0.4)^6\right]}{1-0.4} \\ &\approx 0.3984 \end{aligned}$$

Exercise 5.02 Page 208 Nelson Senior Maths

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

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

This is failure, failure, ..., failure, success.

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

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

2	B	because $(0.2)^4(0.8) = 0.001\,28$
3	D	0.0723, $P(X = 4) = (0.7)^4(0.3) = 0.07203$
4	B	$P(X > 3) = 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3)$ $= 1 - [0.8 + 0.2 \times 0.8 + (0.2)^2 \times 0.8 + (0.2)^3 \times 0.8]$ $= 0.0016$
5	A	$E(X) = \dfrac{1-p}{p} = \dfrac{1-0.8}{0.8} = 0.25$

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.

a	Yes, as failure, failure, success with the same probabilities.
b	No, as the probabilities do not stay the same.
c	Yes, as a sequence of q, q, ..., q, p with the same probabilities.
d	Yes, as a sequence of B, B, ..., B, G with the same probabilities.
e	Yes, as a sequence of N, N, ..., N, Y with the same probabilities.

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 $P(X \le 5)$ c $E(X)$

7 $P(\text{exactly six rolls are required}) = \left(\frac{4}{6}\right)^5\left(\frac{2}{6}\right) = 0.0439$

8 $P(\text{H}) + P(\text{MH}) + P(\text{MMH}) = \frac{1}{3} + \frac{2}{3}\times\frac{1}{3} + \frac{2}{3}\times\frac{2}{3}\times\frac{1}{3} = \frac{9+6+4}{27} = \frac{19}{27} = 0.704$

9 **a** $P(X=5) = 0.8^5 \times 0.2 = 0.0655$

b $P(X \le 5) = 0.2 + 0.8^1 \times 0.2 + 0.8^2 \times 0.2 + 0.8^3 \times 0.2 + 0.8^4 \times 0.2 + 0.8^5 \times 0.2$

$= 0.2(1 + 0.8 + 0.8^2 + 0.8^3 + 0.8^4 + 0.8^5)$

$= 0.738$

c $E(\text{X}) = \frac{1-p}{p} = \frac{1-0.2}{0.2} = 4$

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.

10 $P(\text{underweight}) = \frac{3}{4}$

a $P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{1}{4} = \frac{3}{4}$

b $P(1 \leq X \leq 5) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5$

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

$$= \frac{3}{4} \times \frac{1}{4}\left[1 + \frac{3}{4} + \left(\frac{3}{4}\right)^2 + \left(\frac{3}{4}\right)^3 + \left(\frac{3}{4}\right)^4\right] = 0.572$$

11 $P(\text{hit his drive straight}) = 0.1$

a $P(4 \leq X \leq 14) = (0.9)^4(0.1) + (0.9)^5(0.1) + \ldots + (0.9)^{14}(0.1)$

$= (0.9)^4(0.1)\{1 + 0.9 + 0.9^2 + \ldots + (0.9)^{10}\}$

(Note: Geometric series)

$= (0.9)^4(0.1)\{6.861\ 894\}$

$= 0.45$

b $E(X) = \frac{1-p}{p} = \frac{1-0.1}{0.1} = 9$

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 $P(2 \leq Y \leq 8)$

12 $P(\text{correct}) = 0.1$

Y denotes the number of questions answered for the first correct answer.

a $P(Y = 7) = (0.9)^6(0.1) = 0.0531$

b $P(2 \leq Y \leq 8) = (0.9)(0.1) + (0.9)^2(0.1) + (0.9)^3(0.1) + (0.9)^4(0.1) + (0.9)^5(0.1)$
$+ (0.9)^6(0.1) + (0.9)^7(0.1)$
$= (0.9)(0.1)\{1 + (0.9)^1 + (0.9)^2 + (0.9)^3 + (0.9)^4 + (0.9)^5 + (0.9)^6\}$

$P(2 \leq Y \leq 8) = 0.4695$

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.

13 $P(\text{even}) = 0.5$

a $P(\text{exactly } 2) = 0.5 \times 0.5 = 0.25$

b $P(\text{at least } 2) = P(X \geq 2) = 1 - P(X = 0) - P(X = 1)$

$= 1 - 0.5 - 0.25 = 0.25$

c $P(\text{no more than } 2) = P(X = 0) + P(X = 1) + P(X = 2)$

$= 0.5 + 0.5 \times 0.5 + 0.5 \times 0.5 \times 0.5$

$= 0.875$

14 $P(\text{sinks putt}) = \frac{3}{4}$

$P(\text{qqqqp}) = \left(\frac{1}{4}\right)^4 \times \frac{3}{4} = 0.00293$

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?

15 $P(\text{speed}) = 0.1$

$P(\text{qqqqqp}) = (0.9)^5 \times 0.1 = 0.059$

16 $P(\text{fraud}) = 0.25$

a $P(\text{qqqqq}) = 0.75^5 = 0.237$

b $E(X) = \dfrac{1-p}{p} = \dfrac{1-0.25}{0.25} = 3$

i.e., you would get away with two and the third one would be expected to be detected.

17 $P(\text{job}) = 0.75$

$P(\text{qqqp}) = (0.25)^3 0.75 \approx 0.0117$

EXERCISE 5.02 (PAGE 208)

- Question 1, 3, 4, 6, 8, 10