

The Binomial Distribution

37.2

Introduction

A situation in which an experiment (or trial) is repeated a fixed number of times can be modelled, under certain assumptions, by the binomial distribution. Within each trial we focus attention on a particular outcome. If the outcome occurs we label this as a success. The binomial distribution allows us to calculate the probability of observing a certain number of successes in a given number of trials.

You should note that the term 'success' (and by implication 'failure') are simply labels and as such might be misleading. For example counting the number of defective items produced by a machine might be thought of as counting successes if you are looking for defective items! Trials with two possible outcomes are often used as the building blocks of random experiments and can be useful to engineers. Two examples are:

1. A particular mobile phone link is known to transmit 6% of 'bits' of information in error. As an engineer you might need to know the probability that two bits out of the next ten transmitted are in error.
2. A machine is known to produce, on average, 2% defective components. As an engineer you might need to know the probability that 3 items are defective in the next 20 produced.

The binomial distribution will help you to answer such questions.



Prerequisites

Before starting this Section you should . . .

- understand the concepts of probability



Learning Outcomes

On completion you should be able to . . .

- recognise and use the formula for binomial probabilities
- state the assumptions on which the binomial model is based

1. The binomial model

We have introduced random variables from a general perspective and have seen that there are two basic types: discrete and continuous. We examine four particular examples of distributions for random variables which occur often in practice and have been given special names. They are the **binomial** distribution, the **Poisson** distribution, the **Hypergeometric** distribution and the **Normal** distribution. The first three are distributions for discrete random variables and the fourth is for a continuous random variable. In this Section we focus attention on the binomial distribution.

The binomial distribution can be used in situations in which a given experiment (often referred to, in this context, as a **trial**) is repeated a number of times. For the binomial model to be applied the following four criteria must be satisfied:

- the trial is carried out a fixed number of times n
- the outcomes of each trial can be classified into two 'types' conventionally named 'success' or 'failure'
- the probability p of success remains constant for each trial
- the individual trials are independent of each other.

For example, if we consider throwing a coin 7 times what is the probability that exactly 4 Heads occur? This problem can be modelled by the binomial distribution since the four basic criteria are assumed satisfied as we see.

- here the trial is 'throwing a coin' which is carried out 7 times
- the occurrence of Heads on any given trial (i.e. throw) may be called a 'success' and Tails called a 'failure'
- the probability of success is $p = \frac{1}{2}$ and remains constant for each trial
- each throw of the coin is independent of the others.

The reader will be able to complete the solution to this example once we have constructed the general binomial model.

The following two scenarios are typical of those met by engineers. The reader should check that the criteria stated above are met by each scenario.

1. An electronic product has a total of 30 integrated circuits built into it. The product is capable of operating successfully only if at least 27 of the circuits operate properly. What is the probability that the product operates successfully if the probability of any integrated circuit failing to operate is 0.01?
2. Digital communication is achieved by transmitting information in "bits". Errors do occur in data transmissions. Suppose that the number of bits in error is represented by the random variable X and that the probability of a communication error in a bit is 0.001. If at most 2 errors are present in a 1000 bit transmission, the transmission can be successfully decoded. If a 1000 bit message is transmitted, find the probability that it can be successfully decoded.

Before developing the *general* binomial distribution we consider the following examples which, as you will soon recognise, have the basic characteristics of a binomial distribution.



Example 7

In a box of floppy discs it is known that 95% will work. A sample of three of the discs is selected at random.

Find the probability that (a) none (b) 1, (c) 2, (d) all 3 of the sample will work.

Solution

Let the event {the disc works} be W and the event {the disc fails} be F . The probability that a disc will work is denoted by $P(W)$ and the probability that a disc will fail is denoted by $P(F)$. Then $P(W) = 0.95$ and $P(F) = 1 - P(W) = 1 - 0.95 = 0.05$.

- (a) The probability that none of the discs works equals the probability that all 3 discs fail. This is given by:

$$\begin{aligned} P(\text{none work}) &= P(FFF) = P(F) \times P(F) \times P(F) \quad \text{as the events are independent} \\ &= 0.05 \times 0.05 \times 0.05 = 0.05^3 = 0.000125 \end{aligned}$$

- (b) If only one disc works then you could select the three discs in the following orders

(FFW) or (FWF) or (WFF) hence

$$\begin{aligned} P(\text{one works}) &= P(FFW) + P(FWF) + P(WFF) \\ &= P(F) \times P(F) \times P(W) + P(F) \times P(W) \times P(F) + P(W) \times P(F) \times P(F) \\ &= (0.05 \times 0.05 \times 0.95) + (0.05 \times 0.95 \times 0.05) + (0.95 \times 0.05 \times 0.05) \\ &= 3 \times (0.05)^2 \times 0.95 = 0.007125 \end{aligned}$$

- (c) If 2 discs work you could select them in order

(FWW) or (WFW) or (WWF) hence

$$\begin{aligned} P(\text{two work}) &= P(FWW) + P(WFW) + P(WWF) \\ &= P(F) \times P(W) \times P(W) + P(W) \times P(F) \times P(W) + P(W) \times P(W) \times P(F) \\ &= (0.05 \times 0.95 \times 0.95) + (0.95 \times 0.05 \times 0.95) + (0.95 \times 0.95 \times 0.05) \\ &= 3 \times (0.05) \times (0.95)^2 = 0.135375 \end{aligned}$$

- (d) The probability that all 3 discs work is given by $P(WWW) = 0.95^3 = 0.857375$.

Notice that since the 4 outcomes we have dealt with are *all possible outcomes* of selecting 3 discs, the probabilities should add up to 1. It is an easy check to verify that they do.

One of the most important assumptions above is that of **independence**. The probability of selecting a working disc remains unchanged no matter whether the previous selected disc worked or not.



Example 8

A worn machine is known to produce 10% defective components. If the random variable X is the number of defective components produced in a run of 3 components, find the probabilities that X takes the values 0 to 3.

Solution

Assuming that the production of components is independent and that the probability $p = 0.1$ of producing a defective component remains constant, the following table summarizes the production run. We let G represent a *good* component and let D represent a *defective* component.

Note that since we are only dealing with two possible outcomes, we can say that the probability q of the machine producing a good component is $1 - 0.1 = 0.9$. More generally, we know that $q + p = 1$ if we are dealing with a binomial distribution.

Outcome	Value of X	Probability of Occurrence
GGG	0	$(0.9)(0.9)(0.9) = (0.9)^3$
GGD	1	$(0.9)(0.9)(0.1) = (0.9)^2(0.1)$
GDG	1	$(0.9)(0.1)(0.9) = (0.9)^2(0.1)$
DGG	1	$(0.1)(0.9)(0.9) = (0.9)^2(0.1)$
DDG	2	$(0.1)(0.1)(0.9) = (0.9)(0.1)^2$
DGD	2	$(0.1)(0.9)(0.1) = (0.9)(0.1)^2$
GDD	2	$(0.9)(0.1)(0.1) = (0.9)(0.1)^2$
DDD	3	$(0.1)(0.1)(0.1) = (0.1)^3$

From this table it is easy to see that

$$P(X = 0) = (0.9)^3$$

$$P(X = 1) = 3 \times (0.9)^2(0.1)$$

$$P(X = 2) = 3 \times (0.9)(0.1)^2$$

$$P(X = 3) = (0.1)^3$$

Clearly, a pattern is developing. In fact you may have already realized that the probabilities we have found are just the terms of the expansion of the expression $(0.9 + 0.1)^3$ since

$$(0.9 + 0.1)^3 = (0.9)^3 + 3 \times (0.9)^2(0.1) + 3 \times (0.9)(0.1)^2 + (0.1)^3$$

We now develop the binomial distribution from a more general perspective. If you find the theory getting a bit heavy simply refer back to this example to help clarify the situation.

First we shall find it convenient to denote the probability of failure on a trial, which is $1 - p$, by q , that is:

$$q = 1 - p.$$

What we shall do is to calculate probabilities of the number of 'successes' occurring in n trials, beginning with $n = 1$.

$n = 1$ With only one trial we can observe either 1 success (with probability p) or 0 successes (with probability q).

$n = 2$ Here there are 3 possibilities: We can observe 2, 1 or 0 successes. Let S denote a success and F denote a failure. So a failure followed by a success would be denoted by FS whilst two failures followed by one success would be denoted by FFS and so on.

Then

$$P(\text{2 successes in 2 trials}) = P(SS) = P(S)P(S) = p^2$$

(where we have used the assumption of independence between trials and hence multiplied probabilities). Now, using the usual rules of basic probability, we have:

$$P(\text{1 success in 2 trials}) = P[(SF) \cup (FS)] = P(SF) + P(FS) = pq + qp = 2pq$$

$$P(\text{0 successes in 2 trials}) = P(FF) = P(F)P(F) = q^2$$

The three probabilities we have found – q^2 , $2qp$, p^2 – are in fact the terms which arise in the binomial expansion of $(q + p)^2 = q^2 + 2qp + p^2$. We also note that since $q = 1 - p$ the probabilities sum to 1 (as we should expect):

$$q^2 + 2qp + p^2 = (q + p)^2 = ((1 - p) + p)^2 = 1$$



List the outcomes for the binomial model for the case $n = 3$, calculate their probabilities and display the results in a table.

Your solution

Answer

{three successes, two successes, one success, no successes}

Three successes occur only as SSS with probability p^3 .

Two successes can occur as SSF with probability (p^2q) , as SFS with probability (pqp) or as FSS with probability (qp^2) .

These are mutually exclusive events so the combined probability is the sum $3p^2q$.

Similarly, we can calculate the other probabilities and obtain the following table of results.

Number of successes	3	2	1	0
Probability	p^3	$3p^2q$	$3pq^2$	q^3

Note that the probabilities you have obtained:

$$q^3, 3q^2p, 3qp^2, p^3$$

are the terms which arise in the binomial expansion of $(q + p)^3 = q^3 + 3q^2p + 3qp^2 + p^3$



Repeat the previous Task for the binomial model for the case with $n = 4$.

Your solution

Answer

Number of successes	4	3	2	1	0
Probability	p^4	$4p^3q$	$6p^2q^2$	$4pq^3$	q^4

Again we explore the connection between the probabilities and the terms in the binomial expansion of $(q + p)^4$. Consider this expansion

$$(q + p)^4 = q^4 + 4q^3p + 6q^2p^2 + 4qp^3 + p^4$$

Then, for example, the term $4p^3q$, is the probability of 3 successes in the four trials. These successes can occur anywhere in the four trials and there must be one failure hence the p^3 and q components which are multiplied together. The remaining part of this term, 4, is the number of ways of selecting three objects from 4.

Similarly there are ${}^4C_2 = \frac{4!}{2!2!} = 6$ ways of selecting two objects from 4 so that the coefficient 6 combines with p^2 and q^2 to give the probability of two successes (and hence two failures) in four trials.

The approach described here can be extended for any number n of trials.



Key Point 4

The Binomial Probabilities

Let X be a discrete random variable, being the number of successes occurring in n independent trials of an experiment. If X is to be described by the binomial model, the probability of exactly r successes in n trials is given by

$$P(X = r) = {}^n C_r p^r q^{n-r}.$$

Here there are r successes (each with probability p), $n - r$ failures (each with probability q) and

${}^n C_r = \frac{n!}{r!(n-r)!}$ is the number of ways of placing the r successes among the n trials.

Notation

If a random variable X follows a binomial distribution in which an experiment is repeated n times each with probability p of success then we write $X \sim B(n, p)$.



Example 9

A worn machine is known to produce 10% defective components. If the random variable X is the number of defective components produced in a run of 4 components, find the probabilities that X takes the values 0 to 4.

Solution

From Example 8, we know that the probabilities required are the terms of the expansion of the expression:

$$(0.9 + 0.1)^4 \quad \text{so} \quad X \sim B(4, 0.1)$$

Hence the required probabilities are (using the general formula with $n = 4$ and $p = 0.1$)

$$P(X = 0) = (0.9)^4 = 0.6561$$

$$P(X = 1) = 4(0.9)^3(0.1) = 0.2916$$

$$P(X = 2) = \frac{4 \times 3}{1 \times 2}(0.9)^2(0.1)^2 = 0.0486$$

$$P(X = 3) = \frac{4 \times 3 \times 2}{1 \times 2 \times 3}(0.9)(0.1)^3 = 0.0036$$

$$P(X = 4) = (0.1)^4 = 0.0001$$

Also, since we are using the expansion of $(0.9 + 0.1)^4$, the probabilities should sum to 1, This is a useful check on your arithmetic when you are using a binomial distribution.



Example 10

In a box of switches it is known 10% of the switches are faulty. A technician is wiring 30 circuits, each of which needs one switch. What is the probability that (a) all 30 work, (b) at most 2 of the circuits do not work?

Solution

The answers involve binomial distributions because there are only two states for each circuit - it either works or it doesn't work.

A trial is the operation of testing each circuit.

A success is that it works. We are given $P(\text{success}) = p = 0.9$

Also we have the number of trials $n = 30$

Applying the binomial distribution $P(X = r) = {}^n C_r p^r (1 - p)^{n-r}$.

(a) Probability that all 30 work is $P(X = 30) = {}^{30} C_{30} (0.9)^{30} (0.1)^0 = 0.04239$

(b) The statement that "at most 2 circuits do not work" implies that 28, 29 or 30 work. That is $X \geq 28$

$$P(X \geq 28) = P(X = 28) + P(X = 29) + P(X = 30)$$

$$P(X = 30) = {}^{30} C_{30} (0.9)^{30} (0.1)^0 = 0.04239$$

$$P(X = 29) = {}^{30} C_{29} (0.9)^{29} (0.1)^1 = 0.14130$$

$$P(X = 28) = {}^{30} C_{28} (0.9)^{28} (0.1)^2 = 0.22766$$

$$\text{Hence } P(X \geq 28) = 0.41135$$



Example 11

A University Engineering Department has introduced a new software package called SOLVIT. To save money, the University's Purchasing Department has negotiated a bargain price for a 4-user licence that allows only four students to use SOLVIT at any one time. It is estimated that this should allow 90% of students to use the package when they need it. The Students' Union has asked for more licences to be bought since engineering students report having to queue excessively to use SOLVIT. As a result the Computer Centre monitors the use of the software. Their findings show that on average 20 students are logged on at peak times and 4 of these want to use SOLVIT. Was the Purchasing Department's estimate correct?

Solution

$$P(\text{student wanted to use SOLVIT}) = \frac{4}{20} = 0.2$$

Let X be the number of students wanting to use SOLVIT at any one time, then

$$P(X = 0) = {}^{20}C_0(0.2)^0(0.8)^{20} = 0.0115$$

$$P(X = 1) = {}^{20}C_1(0.2)^1(0.8)^{19} = 0.0576$$

$$P(X = 2) = {}^{20}C_2(0.2)^2(0.8)^{18} = 0.1369$$

$$P(X = 3) = {}^{20}C_3(0.2)^3(0.8)^{17} = 0.2054$$

$$P(X = 4) = {}^{20}C_4(0.2)^4(0.8)^{16} = 0.2182$$

Therefore

$$\begin{aligned} P(X \leq 4) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) \\ &= 0.01152 + 0.0576 + 0.1369 + 0.2054 + 0.2182 \\ &= 0.61862 \end{aligned}$$

The probability that more than 4 students will want to use SOLVIT is

$$P(X > 4) = 1 - P(X \leq 4) = 0.38138$$

That is, 38% of the time there will be more than 4 students wanting to use the software. The Purchasing Department has grossly overestimated the availability of the software on the basis of a 4-user licence.



Using the binomial model, and assuming that a success occurs with probability $\frac{1}{5}$ in each trial, find the probability that in 6 trials there are

- (a) 0 successes (b) 3 successes (c) 2 failures.

Let X be the number of successes in 6 independent trials.

Your solution

(a) $P(X = 0) =$

Answer

In each case $p = \frac{1}{5}$ and $q = 1 - p = \frac{4}{5}$.

Here $r = 0$ and

$$P(X = 0) = q^6 = \left(\frac{4}{5}\right)^6 = \frac{4096}{15625} \approx 0.262$$

Your solution

(b) $P(X = 3) =$

Answer

$$r = 3 \text{ and } P(X = 3) = {}^6C_3 p^3 q^3 = \frac{6 \times 5 \times 4}{1 \times 2 \times 3} \times \left(\frac{1}{5}\right)^3 \times \left(\frac{4}{5}\right)^3 = \frac{20 \times 64}{5^6} = \frac{12 \times 80}{15625} = 0.0819$$

Your solution

(c) $P(X = 4) =$

Answer

$$\text{Here } r = 4 \text{ and } P(X = 4) = {}^6C_4 p^4 q^2 = \frac{6 \times 5}{1 \times 2} \times \left(\frac{1}{5}\right)^4 \times \left(\frac{4}{5}\right)^2 = \frac{15 \times 4^2}{5^6} = \frac{240}{15625} = 0.01536$$

2. Expectation and variance of the binomial distribution

For a binomial distribution $X \sim B(n, p)$, the mean and variance, as we shall see, have a simple form. While we will not prove the formulae in general terms - the algebra can be rather tedious - we will illustrate the results for cases involving small values of n .

The case $n = 2$

Essentially, we have a random variable X which follows a binomial distribution $X \sim B(2, p)$ so that the values taken by X (and X^2 - needed to calculate the variance) are shown in the following table:

x	x^2	$P(X = x)$	$xP(X = x)$	$x^2P(X = x)$
0	0	q^2	0	0
1	1	$2qp$	$2qp$	$2qp$
2	4	p^2	$2p^2$	$4p^2$

We can now calculate the mean of this distribution:

$$E(X) = \sum xP(X = x) = 0 + 2qp + 2p^2 = 2p(q + p) = 2p \quad \text{since } q + p = 1$$

Similarly, the variance $V(X)$ is given by

$$V(X) = E(X^2) - [E(X)]^2 = 0 + 2qp + 4p^2 - (2p)^2 = 2qp$$



Calculate the mean and variance of a random variable X which follows a binomial distribution $X \sim B(3, p)$.

Your solution

Answer

The table of values appropriate to this case is:

x	x^2	$P(X = x)$	$xP(X = x)$	$x^2P(X = x)$
0	0	q^3	0	0
1	1	$3q^2p$	$3q^2p$	$3q^2p$
2	4	$3qp^2$	$6qp^2$	$12qp^2$
3	9	p^3	$3p^3$	$9p^3$

Hence $E(X) = \sum xP(X = x) = 0 + 3q^2p + 6qp^2 + 3p^3 = 3p(q + p)^2 = 3p$ since $q + p = 1$

$$\begin{aligned}
 V(X) &= E(X^2) - [E(X)]^2 \\
 &= 0 + 3q^2p + 12qp^2 + 9p^3 - (3p)^2 \\
 &= 3p(q^2 + 4qp + 3p^2 - 3p) \\
 &= 3p((1 - p)^2 + 4(1 - p)p + 3p^2 - 3p) \\
 &= 3p(1 - 2p + p^2 + 4p - 4p^2 + 3p^2 - 3p) = 3p(1 - p) = 3pq
 \end{aligned}$$

From the results given above, it is reasonable to assert the following result in Key Point 5.

**Key Point 5****Expectation and Variance of the Binomial Distribution**

If a random variable X which can assume the values $0, 1, 2, 3, \dots, n$ follows a binomial distribution $X \sim B(n, p)$ so that

$$P(X = r) = {}^nC_r p^r q^{n-r} = {}^nC_r p^r (1 - p)^{n-r}$$

then the expectation and variance of the distribution are given by the formulae

$$E(X) = np \quad \text{and} \quad V(X) = np(1 - p) = npq$$



A die is thrown repeatedly 36 times in all. Find $E(X)$ and $V(X)$ where X is the number of sixes obtained.

Your solution

Answer

Consider the occurrence of a six, with X being the number of sixes thrown in 36 trials.

The random variable X follows a binomial distribution. (Why? Refer to page 18 for the criteria if necessary). A trial is the operation of throwing a die. A success is the occurrence of a 6 on a particular trial, so $p = \frac{1}{6}$. We have $n = 36$, $p = \frac{1}{6}$ so that

$$E(X) = np = 36 \times \frac{1}{6} = 6 \quad \text{and} \quad V(X) = npq = 36 \times \frac{1}{6} \times \frac{5}{6} = 5.$$

Hence the standard deviation is $\sigma = \sqrt{5} \simeq 2.236$.

$E(X) = 6$ implies that in 36 throws of a fair die we would expect, on average, to see 6 sixes. This makes perfect sense, of course.

Exercises

- The probability that a mountain-bike rider travelling along a certain track will have a tyre burst is 0.05. Find the probability that among 17 riders:
 - exactly one has a burst tyre
 - at most three have a burst tyre
 - two or more have burst tyres.
- A transmission channel transmits zeros and ones in strings of length 8, called 'words'. Possible distortion may change a one to a zero or vice versa; assume this distortion occurs with probability .01 for each digit, independently. An error-correcting code is employed in the construction of the word such that the receiver can deduce the word correctly if at most one digit is in error. What is the probability the word is decoded incorrectly?
 - Assume that a word is a sequence of 10 zeros or ones and, as before, the probability of incorrect transmission of a digit is .01. If the error-correcting code allows correct decoding of the word if no more than two digits are incorrect, compute the probability that the word is decoded correctly.
- An examination consists of 10 multi-choice questions, in each of which a candidate has to deduce which one of five suggested answers is correct. A completely unprepared student guesses each answer completely randomly. What is the probability that this student gets 8 or more questions correct? Draw the appropriate moral!
- The probability that a machine will produce all bolts in a production run within specification is 0.998. A sample of 8 machines is taken at random. Calculate the probability that
 - all 8 machines, (b) 7 or 8 machines, (c) at least 6 machineswill produce all bolts within specification
- The probability that a machine develops a fault within the first 3 years of use is 0.003. If 40 machines are selected at random, calculate the probability that 38 or more will not develop any faults within the first 3 years of use.
- A computer installation has 10 terminals. Independently, the probability that any one terminal will require attention during a week is 0.1. Find the probabilities that
 - 0, (b), 1 (c) 2, (d) 3 or more, terminals will require attention during the next week.
- The quality of electronic chips is checked by examining samples of 5. The frequency distribution of the number of defective chips per sample obtained when 100 samples have been examined is:

No. of defectives	0	1	2	3	4	5
No. of samples	47	34	16	3	0	0

Calculate the proportion of defective chips in the 500 tested. Assuming that a binomial distribution holds, use this value to calculate the expected frequencies corresponding to the observed frequencies in the table.

Exercises continued

8. In a large school, 80% of the pupils like mathematics. A visitor to the school asks each of 4 pupils, chosen at random, whether they like mathematics.

- (a) Calculate the probabilities of obtaining an answer *yes* from 0, 1, 2, 3, 4 of the pupils
- (b) Find the probability that the visitor obtains the answer *yes* from at least 2 pupils:
 - (i) when the number of pupils questioned remains at 4
 - (ii) when the number of pupils questioned is increased to 8.

9. A machine has two drive belts, one on the left and one on the right. From time to time the drive belts break. When one breaks the machine is stopped and both belts are replaced. Details of n consecutive breakages are recorded. Assume that the left and right belts are equally likely to break first. Let X be the number of times the break is on the left.

- (a) How many possible different sequences of “left” and “right” are there?
- (b) How many of these sequences contain exactly j “lefts”?
- (c) Find an expression, in terms of n and j , for the probability that $X = j$.
- (d) Let $n = 6$. Find the probability distribution of X .

10. A machine is built to make mass-produced items. Each item made by the machine has a probability p of being defective. Given the value of p , the items are independent of each other. Because of the way in which the machines are made, p could take one of several values. In fact $p = X/100$ where X has a discrete uniform distribution on the interval $[0, 5]$. The machine is tested by counting the number of items made before a defective is produced. Find the conditional probability distribution of X given that the first defective item is the thirteenth to be made.

11. Seven batches of articles are manufactured. Each batch contains ten articles. Each article has, independently, a probability of 0.1 of being defective. Find the probability that there is at least one defective article

- (a) in exactly four of the batches,
- (b) in four or more of the batches.

12. A service engineer is can be called out for maintenance on the photocopiers in the offices of four large companies, A, B, C and D. On any given week there is a probability of 0.1 that he will be called to each of these companies. The event of being called to one company is independent of whether or not he is called to any of the others.

- (a) Find the probability that, on a particular day,
 - (i) he is called to all four companies,
 - (ii) he is called to at least three companies,
 - (iii) he is called to all four given that he is called to at least one,
 - (iv) he is called to all four given that he is called to Company A.
- (b) Find the expected value and variance of the number of these companies which call the engineer on a given day.

Exercises continued

13. There are five machines in a factory. Of these machines, three are working properly and two are defective. Machines which are working properly produce articles each of which has independently a probability of 0.1 of being imperfect. For the defective machines this probability is 0.2. A machine is chosen at random and five articles produced by the machine are examined. What is the probability that the machine chosen is defective given that, of the five articles examined, two are imperfect and three are perfect?

14. A company buys mass-produced articles from a supplier. Each article has a probability p of being defective, independently of other articles. If the articles are manufactured correctly then $p = 0.05$. However, a cheaper method of manufacture can be used and this results in $p = 0.1$.

- (a) Find the probability of observing exactly three defectives in a sample of twenty articles
 - (i) given that $p = 0.05$
 - (ii) given that $p = 0.1$.
- (b) The articles are made in large batches. Unfortunately batches made by both methods are stored together and are indistinguishable until tested, although all of the articles in any one batch will be made by the same method. Suppose that a batch delivered to the company has a probability of 0.7 of being made by the correct method. Find the conditional probability that such a batch is correctly manufactured given that, in a sample of twenty articles from the batch, there are exactly three defectives.
- (c) The company can either accept or reject a batch. Rejecting a batch leads to a loss for the company of £150. Accepting a batch which was manufactured by the cheap method will lead to a loss for the company of £400. Accepting a batch which was correctly manufactured leads to a profit of £500. Determine a rule for what the company should do if a sample of twenty articles contains exactly three defectives, in order to maximise the expected value of the profit (where loss is negative profit). Should such a batch be accepted or rejected?
- (d) Repeat the calculation for four defectives in a sample of twenty and hence, or otherwise, determine a rule for how the company should decide whether to accept or reject a batch according to the number of defectives.

Answers

1. Binomial distribution $P(X = r) = {}^nC_r p^r (1-p)^{n-r}$ where p is the probability of single 'success' which is 'tyre burst'.

$$(a) P(X = 1) = {}^{17}C_1 (0.05)^1 (0.95)^{16} = 0.3741$$

(b)

$$\begin{aligned} P(X \leq 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= (0.95)^{17} + 17(0.05)(0.95)^{16} + \frac{17 \times 16}{2 \times 1} (0.05)^2 (0.95)^{15} \\ &\quad + \frac{17 \times 16 \times 15}{3 \times 2 \times 1} (0.05)^3 (0.95)^{14} = 0.9912 \end{aligned}$$

$$(c) P(X \geq 2) = 1 - P[(X = 0) \cup (X = 1)] = 1 - (0.95)^{17} - 17(0.05)(0.95)^{16} = 0.2077$$

2.

- (a) $P(\text{distortion}) = 0.01$ for each digit. This is a binomial situation in which the probability of 'success' is $0.01 = p$ and there are $n = 8$ trials.

A word is decoded incorrectly if there are two or more digits in error

$$\begin{aligned} P(X \geq 2) &= 1 - P[(X = 0) \cup (X = 1)] \\ &= 1 - {}^8C_0 (0.99)^8 - {}^8C_1 (0.01)(0.99)^7 = 0.00269 \end{aligned}$$

- (b) Same as (a) with $n = 10$. Correct decoding if $X \leq 2$

$$\begin{aligned} P(X \leq 2) &= P[(X = 0) \cup (X = 1) \cup (X = 2)] \\ &= (0.99)^{10} + 10(0.01)(0.99)^9 + 45(0.01)^2 (0.99)^8 = 0.99989 \end{aligned}$$

3. Let X be a random variable 'number of answers guessed correctly' then for each question (i.e. trial) the probability of a 'success' = $\frac{1}{5}$. It is clear that X follows a binomial distribution with $n = 10$ and $p = 0.2$.

$$P(\text{randomly choosing correct answer}) = \frac{1}{5} \quad n = 10$$

$$\begin{aligned} P(8 \text{ or more correct}) &= P[(X = 8) \cup (X = 9) \cup (X = 10)] \\ &= {}^{10}C_8 (0.2)^8 (0.8)^2 + {}^{10}C_9 (0.2)^9 (0.8) + {}^{10}C_{10} (0.2)^{10} = 0.000078 \end{aligned}$$

4. (a) 0.9841 (b) 0.9999 (c) 1.0000

$$5. P(X \geq 38) = P(X = 38) + P(X = 39) + P(X = 40) = 0.00626 + 0.1067 + 0.88676 = 0.99975$$

6. (a) 0.3487 (b) 0.3874 (c) 0.1937 (d) 0.0702

$$7. 0.15 \text{ (total defectives = } 0 + 34 + 32 + 9 + 0 \text{ out of 500 tested); } 44, 39, 14, 2, 0, 0$$

8. (a) 0.0016, 0.0256, 0.1536, 0.4096, 0.4096; (b)(i) 0.9728 (b)(ii) 0.9988

Answers

9.

(a) There are 2^n possible sequences.

(b) The number containing exactly j “lefts” is $\binom{n}{j}$.

(c) $P(X = j) = \binom{n}{j} 2^{-n}$.

(d) With $n = 6$ the distribution of X is

j	0	1	2	3	4	5	6
$P(X = j)$	0.015625	0.09375	0.234375	0.3125	0.234375	0.09375	0.015625

10. Let Y be the number of the first defective item.

$$P(X = j | Y = 13) = \frac{P(X = j) \times P(Y = 13 | X = j)}{\sum_{i=0}^5 P(X = i) \times P(Y = 13 | X = i)} = \frac{P(Y = 13 | X = j)}{\sum_{i=0}^5 P(Y = 13 | X = i)}$$

since $P(X = j) = 1/6$ for $j = 0, \dots, 5$.

$$P(Y = 13 | X = j) = \left(1 - \frac{X}{100}\right)^{12} \left(\frac{X}{100}\right)$$

j	$P(Y = 13 X = j)$	$P(X = j Y = 13)$
0	0.00000	0.0000
1	0.00886	0.0707
2	0.01569	0.1251
3	0.02082	0.1660
4	0.02451	0.1954
5	0.02702	0.2154
6	0.02856	0.2277
Total	0.12546	1

Answers

11.

The probability of at least one defective in a batch is $1 - 0.9^{10} = 0.6513$.

Let the probability of at least one defective in exactly j batches be p_j .

$$(a) \quad p_4 = \binom{7}{4} (1 - 0.9^{10})^4 (0.9^{10})^3 = 35 \times 0.6513^4 \times 0.3487^3 = 0.2670.$$

(b)

$$p_5 = \binom{7}{5} (1 - 0.9^{10})^5 (0.9^{10})^2 = 21 \times 0.6513^5 \times 0.3487^2 = 0.2993.$$

$$p_6 = \binom{7}{6} (1 - 0.9^{10})^6 (0.9^{10})^1 = 7 \times 0.6513^6 \times 0.3487^1 = 0.1863.$$

$$p_7 = \binom{7}{7} (1 - 0.9^{10})^7 (0.9^{10})^0 = 0.6513^7 = 0.0497.$$

The probability of at least one defective in four or more of the batches is

$$p_4 + p_5 + p_6 + p_7 = 0.8023.$$

12.

(a) Let Y be the number of companies to which the engineer is called and let A denote the event that the engineer is called to company A.

$$(i) \quad P(Y = 4) = 0.1^4 = 0.0001.$$

$$(ii) \quad P(Y \geq 3) = \binom{4}{3} \times 0.1^3 \times 0.9^1 + 0.1^4 = 0.0037.$$

$$(iii) \quad P(Y = 4 | Y \geq 1) = \frac{P(Y = 4 \cap Y \geq 1)}{P(Y \geq 1)}$$

$$= \frac{P(Y = 4)}{P(Y \geq 1)} = \frac{0.0001}{1 - 0.9^4} = \frac{0.0001}{0.3439} = \frac{1}{3439} = 0.0003.$$

$$(iv) \quad P(Y = 4 | A) = \frac{P(Y = 4 \cap A)}{P(A)}$$

$$\frac{P(Y = 4)}{P(A)} = \frac{0.0001}{0.1} = 0.0010.$$

(b) The mean is $E(Y) = 4 \times 0.1 = 0.4$. The variance is $V(Y) = 4 \times 0.1 \times 0.9 = 0.36$.

Answers

13. Let D denote the event that the chosen machine is defective and \bar{D} denote the event "not D ".

Let Y be the number of imperfect articles in the sample of five.

Then

$$\begin{aligned} P(D | Y = 2) &= \frac{P(D) \times P(Y = 2 | D)}{P(D) \times P(Y = 2 | D) + P(\bar{D}) \times P(Y = 2 | \bar{D})} \\ &= \frac{\frac{2}{5} \times \binom{5}{2} \times 0.2^2 \times 0.8^3}{\frac{2}{5} \times \binom{5}{2} \times 0.2^2 \times 0.8^3 + \frac{3}{5} \times \binom{5}{2} \times 0.1^2 \times 0.9^3} \\ &= \frac{2 \times 0.2^2 \times 0.8^3}{2 \times 0.2^2 \times 0.8^3 + 3 \times 0.1^2 \times 0.9^3} \\ &= \frac{0.04096}{0.04096 + 0.02187} = 0.6519. \end{aligned}$$

14.

$$(a) (i) \quad p_3 = \binom{20}{3} 0.1^3 \times 0.9^{17} = \frac{20 \times 19 \times 18}{1 \times 2 \times 3} \times 0.1^3 \times 0.9^7 = 0.190.$$

(ii)

$$p_2 = \binom{20}{2} 0.1^2 \times 0.9^{18} = \frac{3}{18} \times 9 \times p_3 = 0.28518$$

$$p_1 = \binom{20}{1} 0.1 \times 0.9^{19} = \frac{2}{19} \times 9 \times p_2 = 0.27017$$

$$p_0 = \binom{20}{0} 0.9^{20} = 0.12158.$$

The total probability is 0.867.

(iii) The required probability is the probability of at most 2 out of 16.

$$p'_0 = P(0 \text{ out of } 16) = 0.9^{16} = 0.185302$$

$$p'_1 = P(1 \text{ out of } 16) = \frac{16}{9} \times p'_0 = 0.3294258$$

$$p'_2 = P(2 \text{ out of } 16) = \frac{15}{2} \times \frac{1}{9} \times p'_1 = 0.2745215$$

(b)

$$\frac{0.2 \binom{4}{1} \times 0.3^1 \times 0.7^3}{0.2 \binom{4}{1} \times 0.3^1 \times 0.7^3 + 0.9 \binom{4}{1} \times 0.1^1 \times 0.9^3} = \frac{0.02058}{0.02058 + 0.05832} = 0.2608.$$