Discrete Probability Distributions

37.1

Introduction

It is often possible to model real systems by using the same or similar random experiments and their associated random variables. Numerical random variables may be classified in two broad but distinct categories called discrete random variables and continuous random variables. Often, discrete

1. Discrete probability distributions

We shall look at discrete distributions in this Workbook and continuous distributions in HELM 38. In order to get a good understanding of discrete distributions it is advisable to familiarise yourself with two related topics: permutations and combinations. Essentially we shall be using this area of

Example 1

Write out the possible permutations of the letters A, B, C and D.

Solution

The possible permutations are

$$\begin{array}{cccccc} ABCD & ABDC & ADBC & ADCB & ACBD & ACDB \\ BADC & BACD & BCDA & BCAD & BDAC & BDCA \\ CABD & CADB & CDBA & CDAB & CBAD & CBDA \\ DABC & DACB & DCAB & DCBA & DBAC & DBCA \end{array}$$

There are $4! = 24$ permutations of the four letters A, B, C and D.

In general we can order n distinct objects in $n!$ ways.
Suppose we have r different types of object. It follows that if we have n_1 objects of one kind, n_2 of another kind and so on then the n_1 objects can be ordered in $n_1!$ ways, the n_2 objects in $n_2!$ ways and so on. If $n_1 + n_2 + \cdots + n_r = n$ and if p is the number of permutations possible from n objects we may write

$$p \times (n_1! \times n_2! \times \cdots \times n_r!) = n!$$

and so p is given by the formula

$$p = \frac{n!}{n_1! \times n_2! \times \cdots \times n_r!}$$

Very often we will find it useful to be able to calculate the number of permutations of n objects taken r at a time. Assuming that we do not allow repetitions, we may choose the first object in n ways, the second in $n - 1$ ways, the third in $n - 2$ ways and so on so that the r^{th} object may be chosen in $n - r + 1$ ways.

Example 2

Find the number of permutations of the four letters A, B, C and D taken three at a time.

Solution

We may choose the first letter in 4 ways, either A, B, C or D. Suppose, for the purposes of illustration we choose A. We may choose the second letter in 3 ways, either B, C or D. Suppose, for the purposes of illustration we choose B. We may choose the third letter in 2 ways, either C or D. Suppose, for the purposes of illustration we choose C. The total number of choices made is $4 \times 3 \times 2 = 24$.

In general the numbers of permutations of n objects taken r at a time is

$$n(n-1)(n-2)\ldots(n-r+1) \qquad \text{which is the same as} \quad \frac{n!}{(n-r)!}$$

This is usually denoted by ^nP_r so that

$$^nP_r = \frac{n!}{(n-r)!}$$

If we allow repetitions the number of permutations becomes n^r (can you see why?).

Example 3

Find the number of permutations of the four letters A, B, C and D taken two at a time.

Solution

We may choose the first letter in 4 ways and the second letter in 3 ways giving us

$$4 \times 3 = \frac{4 \times 3 \times 2 \times 1}{1 \times 2} = \frac{4!}{2!} = 12 \quad \text{permutations}$$

Combinations

A **combination** of objects takes **no account of order** whereas a permutation does. The formula $^nP_r = \dfrac{n!}{(n-r)!}$ gives us the number of ordered sets of r objects chosen from n. Suppose the number of sets of r objects (taken from n objects) in which order is not taken into account is C. It follows that

$$C \times r! = \frac{n!}{(n-r)!} \quad \text{and so } C \text{ is given by the formula} \qquad C = \frac{n!}{r!(n-r)!}$$

We normally denote the right-hand side of this expression by ^nC_r so that

$$^nC_r = \frac{n!}{r!(n-r)!} \qquad \text{A common alternative notation for } {}^nC_r \text{ is } \binom{n}{r}.$$

Example 4

How many car registrations are there beginning with $NP05$ followed by three letters? Note that, conventionally, I, O and Q may not be chosen.

Solution

We have to choose 3 letters from 23 allowing repetition. Hence the number of registrations beginning with $NP05$ must be $23^3 = 12167$.

(a) How many di□erent signals consisting of five symbols can be sent using the dot and dash of Morse code?

(b) How many can be sent if five symbols *or less* can be sent?

Your solution

Answer

(a) Clearly, the order of the symbols is important. We can choose each symbol in two ways, either a dot or a dash. The number of distinct signals is

$2 \times 2 \times 2 \times 2 \times = 2^5 = 32$

(b) If five *or less* symbols may be used, the total number of signals may be calculated as follows:

Using one symbol: 2 ways
Using two symbols: $2 \times 2 = 4$ ways
Using three symbols: $2 \times 2 \times 2 = 8$ ways
Using four symbols: $2 \times 2 \times 2 \times 2 = 16$ ways

2. Random variables

A random variable X is a quantity whose value cannot be predicted with certainty. We assume that for every real number a the probability $\mathrm{P}(X = a)$ in a trial is well-defined. In practice, engineers are often concerned with two broad types of variables and their probability distributions: discrete random variables and their distributions, and continuous random variables and their distributions. Discrete distributions arise from experiments involving counting, for example, road deaths, car production and aircraft sales, while continuous distributions arise from experiments involving measurement, for example, voltage, corrosion and oil pressure.

Discrete random variables and probability distributions

A random variable X and its distribution are said to be discrete if the values of X can be presented as an ordered list say $x_1, x_2, x_3, \ldots$ with probability values $p_1, p_2, p_3, \ldots$. That is $\mathrm{P}(X = x_i) = p_i$. For example, the number of times a particular machine fails during the course of one calendar year is a discrete random variable.

More generally a discrete distribution $f(x)$ may be defined by:

$$f(x) = \begin{cases} p_i & \text{if } x = x_i \quad i = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases}$$

The distribution function $F(x)$ (sometimes called the cumulative distribution function) is obtained by taking sums as defined by

$$F(x) = \sum_{x_i \le x} f(x_i) = \sum_{x_i \le x} p_i$$

We sum the probabilities p_i for which x_i is less than or equal to x. This gives a step function with jumps of size p_i at each value x_i of X. The step function is defined for all values, not just ts, -114ned fox ju]T/F277.97Tf6.652-1.793Td[(i)]T/F2

Example 5

Turbo Generators plc manufacture seven large turbines for a customer. Three of these turbines do not meet the customer's specification. Quality control inspectors choose two turbines at random. Let the discrete random variable X be defined to be the number of turbines inspected which meet the customer's specification.

(a) Find the probabilities that X takes the values $0, 1$ or 2.

(b) Find and graph the cumulative distribution function.

Solution

(a) The possible values of X are clearly $0, 1$ or 2 and may occur as follows:

Sample Space	Value of X
Turbine faulty, Turbine faulty	0
Turbine faulty, Turbine good	1
Turbine good, Turbine faulty	1
Turbine good, Turbine good	2

We can easily calculate the probability that X takes the values $0, 1$ or 2 as follows:

$$\mathrm{P}(X=0)=\frac{3}{7}\times\frac{2}{6}=\frac{1}{7} \qquad \mathrm{P}(X=1)=\frac{4}{7}\times\frac{3}{6}+\frac{3}{7}\times\frac{4}{6}=\frac{4}{7} \qquad \mathrm{P}(X=2)=\frac{4}{7}\times\frac{3}{6}=\frac{2}{7}$$

The values of $F(x)=\sum_{x_i\leq x}\mathrm{P}(X=x_i)$ are clearly

$$F(0)=\frac{1}{7} \qquad F(1)=\frac{5}{7} \qquad \text{and} \qquad F(2)=\frac{7}{7}=1$$

(b) The graph of the step function $F(x)$ is shown below.

Figure 1

3. Mean and variance of a discrete probability distribution

If an experiment is performed N times in which the n possible outcomes $X = x_1, x_2, x_3, \ldots, x_n$ are observed with frequencies $f_1, f_2, f_3, \ldots, f_n$ respectively, we know that the mean of the distribution of outcomes is given by

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + \ldots + f_nx_n}{f_1 + f_2 + \ldots + f_n} = \frac{\sum_{i=1}^{n} f_ix_i}{\sum_{i=1}^{n} f_i} = \frac{1}{N}\sum_{i=1}^{n} f_ix_i = \sum_{i=1}^{n}\left(\frac{f_i}{N}\right)x_i$$

(Note that $\sum_{i=1}^{n} f_i = f_1 + f_2 + \cdots + f_n = N$.)

The quantity $\dfrac{f_i}{N}$ is called the **relative frequency** of the observation x_i

$$E[g(X)] = \sum_{i}^{n} g(x_i) f(x_i)$$

In particular if $g(X) = X^2$, we obtain $$E[X^2] = \sum_{i}^{n} x_i^2 f(x_i)$$

Example 6

A tra□c engineer is interested in the number of vehicles reaching a particular crossroads during periods of relatively low tra□c flow. The engineer finds that the number of vehicles X reaching the crossroads per minute is governed by the probability distribution:

x	0	1	2	3	4
$\mathrm{P}(X = x)$	0.37	0.39	0.19	0.04	0.01

(a) Calculate the expected value, the variance and the standard deviation of the random variable X.

(b) Graph the probability distribution $\mathrm{P}(X = x)$ and the corresponding cumulative probability distribution $F(x) = \sum_{x_i \le x} \mathrm{P}(X = x_i)$.

Solution

(a) The expectation, variance and standard deviation and cumulative probability values are calculated as follows:

x	x^2	$\mathrm{P}(X = x)$	$F(x)$
0	0	0.37	0.37
1	1	0.39	0.76
2	4	0.19	0.95
3	9	0.04	0.99
4	16	0.01	1.00

$$\begin{aligned}\mathrm{E}(X) &= \sum_{x=0}^{4} x\mathrm{P}(X = x)\\ &= 0 \times 0.37 + 1 \times 0.39 + 2 \times 0.19 + 3 \times 0.04 + 4 \times 0.01\\ &= 0.93\end{aligned}$$

$$\begin{aligned}\mathrm{V}(X) &= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2\\ &= \sum_{x=0}^{4} x^2\mathrm{P}(X = x) - \left[\sum_{x=0}^{4} x\mathrm{P}(X = x)\right]^2\\ &= 0 \times 0.37 + 1 \times 0.39 + 4 \times 0.19 + 9 \times 0.04 + 16 \times 0.01 - (0.93)^2\\ &= 0.8051\end{aligned}$$

The standard deviation is given by $\sigma = \sqrt{}$

Solution (contd.)

(b)

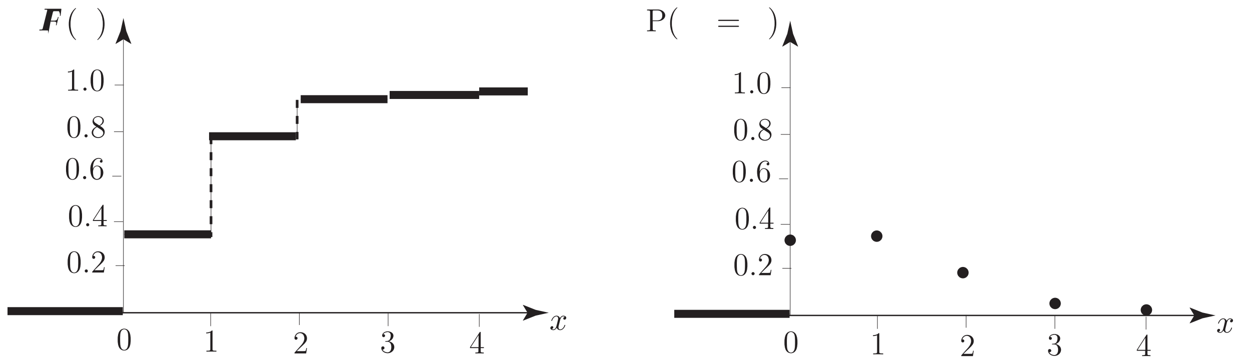

Figure 2

Exercises

1. A machine is operated by two workers. There are sixteen workers available. How many possible teams of two workers are there?

2. A factory has 52 machines. Two of these have been given an experimental modification. In the first week after this modification, problems are reported with thirteen of the machines. What is the probability that both of the modified machines are among the thirteen with problems assuming that all machines are equally likely to give problems,?

3. A factory has 52 machines. Four of these have been given an experimental modification. In the first week after this modification, problems are reported with thirteen of the machines. What is the probability that exactly two of the modified machines are among the thirteen with problems assuming that all machines are equally likely to give problems?

4. A random number generator produces sequences of independent digits, each of which is as likely to be any digit from 0 to 9 as any other. If X denotes any single digit, find $\mathrm{E}(X)$.

5. A hand-held calculator has a clock cycle time of 100 nanoseconds; these are positions numbered $0, 1, \ldots, 99$. Assume a flag is set during a particular cycle at a random position. Thus, if X is the position number at which the flag is set.

$$\mathrm{P}(X = k) = \frac{1}{100} \qquad k = 0, 1, 2, \ldots, 99.$$

Evaluate the average position number $\mathrm{E}(X)$, and σ, the standard deviation.

(Hint: The sum of the first k integers is $k(k + 1)/2$ and the sum of their squares is:

Answers

1. The required number is

$$\binom{16}{2} = \frac{16 \times 15}{2 \times 1} = 120.$$

2. There are

$$\binom{52}{13}$$

possible different selections of 13 machines and all are equally likely. There is only

$$\binom{2}{2} = 1$$

way to pick two machines from those which were modified but there are

$$\binom{50}{11}$$

different choices for the 11 other machines with problems so this is the number of possible selections containing the 2 modified machines.

Hence the required probability is

$$\begin{aligned} \frac{\binom{2}{2}\binom{50}{11}}{\binom{52}{13}} &= \frac{\binom{50}{11}}{\binom{52}{13}} \\ &= \frac{50!/(11!39!)}{52!/(13!39!)} \\ &= \frac{50!13!}{52!11!} \\ &= \frac{13 \times 12}{52 \times 51} \approx 0.0588 \end{aligned}$$

Alternatively, let S be the event "first modified machine is in the group of 13" and C be the event "second modified machine is in the group of 13". Then the required probability is

$$\mathrm{P}(S) \times \mathrm{P}(C \mid S) = \frac{13}{52} \times \frac{12}{51}.$$

Answers

3. There are $\begin{pmatrix} 52 \\ 13 \end{pmatrix}$ different selections of 13, $\begin{pmatrix} 4 \\ 2 \end{pmatrix}$ different choices of two modified machines and $\begin{pmatrix} 48 \\ 11 \end{pmatrix}$ different choices of 11 non-modified machines.

Thus the required probability is

$$\frac{\begin{pmatrix} 4 \\ 2 \end{pmatrix}\begin{pmatrix} 48 \\ 11 \end{pmatrix}}{\begin{pmatrix} 52 \\ 13 \end{pmatrix}} = \frac{(4!/2!2!)(48!/11!37!)}{(52!/13!39!)}$$

$$= \frac{4!48!13!39!}{52!2!2!11!37!}$$

$$= \frac{4 \times 3 \times 13 \times 12 \times 39 \times 38}{52 \times 51 \times 50 \times 49 \times 2} \approx 0.2135$$

Alternatively, let $I(i)$ be the event "modified machine i is in the group of 13" and $O(i)$ be the negation of this, for $i = 1, 2, 3, 4.$ The number of choices of two modified machines is

$$\begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

so the required probability is

$$\begin{pmatrix} 4 \\ 2 \end{pmatrix} \mathrm{P}\{I(1)\} \times \mathrm{P}\{I(2) \mid I(1)\} \times \mathrm{P}\{O(3) \mid I(1), I(2)\} \times \mathrm{P}\{O(4) \mid I(1)I(2)O(3)\}$$

$$= \begin{pmatrix} 4 \\ 2 \end{pmatrix} \frac{13}{52} \times \frac{12}{51} \times \frac{39}{50} \times \frac{38}{49}$$

$$= \frac{4 \times 3 \times 13 \times 12 \times 39 \times 38}{52 \times 51 \times 50 \times 49 \times 2}$$

4.

x	0	1	2	3	4	5	6	7	8	9
$\mathrm{P}(X = x)$	$^1/_{10}$	$^1/_{10}$	$^1/_{10}$	$^1/_{10}$	$^1/_{10}$	$^1/_{10}$	$^1/_{10}$	$^1/_{10}$	$^1/_{10}$	$^1/_{10}$

$$\mathrm{E}(X) = \frac{1}{10}\{0 + 1 + 2 + 3 + \ldots + 9\} = 4.5$$

Answers

5. Same as Q.4 but with 100 positions

$$E(X) = \frac{1}{100}\{0 + 1 + 2 + 3 + ... + 99\} = \frac{1}{100}\left[\frac{99(99 + 1))}{2}\right] = 49.5$$

σ^2 = mean of squares − square of means

$$\therefore \quad \sigma^2 = \frac{1}{100}[1^2 + 2^2 + ... + 99^2] - (49.5)^2$$

$$= \frac{1}{100} \text{[illegible]}$$

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 d26(ite2i4-)-327(erro)28(r.)27ll6oqi4- yf8uceb5d313(t(a)1(d)-380(a)1h)1(d)-q(20)-e1(bi)-1(n422

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**

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 $\mathrm{P}(W)$ and the probability that a disc will fail is denoted by $\mathrm{P}(F)$. Then $\mathrm{P}(W) = 0.95$ and $\mathrm{P}(F) = 1 - \mathrm{P}(W) = 1 - 0.95 = 0.05$.

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

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

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

$$(FFW) \text{ or } (FWF) \text{ or } (WFF) \qquad \text{hence}$$

$$\begin{aligned} \mathrm{P}(\text{one works}) &= \mathrm{P}(FFW) + \mathrm{P}(FWF) + \mathrm{P}(WFF) \\ &= \mathrm{P}(F) \times \mathrm{P}(F) \times \mathrm{P}(W) + \mathrm{P}(F) \times \mathrm{P}(W) \times \mathrm{P}(F) + \mathrm{P}(W) \times \mathrm{P}(F) \times \mathrm{P}(F) \\ &= (0.05 \times 0.05 \times 0.95) + (0.05 \times 0.95 \times 0.05) + (0.95 \times 0.05950. \end{aligned}$$

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.

Note that the probabilities you have obtained:

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

Key Point 4

The Binomial Probabilities

Let X

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(success) = p = 0.9
Also we have the number of trials $n = 30$

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

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

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 $p = \dfrac{1}{5}$ and $q = 1 - p = \dfrac{4}{5}$.

Here $r = 0$ and

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:

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

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

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

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

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

$\mathrm{E}(X) = 6$ 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

1. 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:

   (a) exactly one has a burst tyre

   (b) at most three have a burst tyre

   (c) two or more have burst tyres.

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

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'.

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.

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

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

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

j	$P(Y = 13 \mid X = j)$	$P(X = j \mid Y = 13)$
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) $p_4 = \binom{7}{4}\left(1 - 0.9^{10}\right)^4\left(0.9^{10}\right)^3 = 35 \times 0.6513^4 \times 0.3487^3 = 0.2670.$

(b)

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

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

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

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) $P(Y = 4) = 0.1^4 = 0.0001.$

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

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

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

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

$$\frac{P(Y = 4)}{P(A)}$$ =61(01)]E03B8888800[99239696023905Q45955(27331589[()]F5119557.4440[()]F611252677[

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

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

14.

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

(ii)

$$\begin{aligned}
p_2 &= \binom{20}{2} 0.1^2\times 0.9^{18} = \frac{3}{18}\times 9\times p_3 = 0.28518 \\
p_1 &= \binom{20}{1} 0.1\times 0.9^{19} = \frac{2}{19}\times 9\times p_2 = 0.27017 \\
p_0 &= \binom{20}{0} 0.9^{20} = 0.12158.
\end{aligned}$$

The total probability is 0.867.

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

$$\begin{aligned}
p_0 &= \mathrm{P}(0 \text{ out of } 16) = 0.9^{16} = 0.185302 \\
p_1 &= \mathrm{P}(1 \text{ out of } 16) = \frac{16}{9}\times p_0 = 0.3294258 \\
p_2 &= \mathrm{P}(2 \text{ out of } 16) = \frac{15}{2}\times\frac{1}{9}\times p_1 = 0.2745215
\end{aligned}$$

(b)

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

The Poisson Distribution

37.3

Introduction

In this Section we introduce a probability model which can be used when the outcome of an experiment is a random variable taking on positive integer values and where the only information available is a measurement of its average value. This has widespread applications, for example in analysing traffic flow, in fault prediction on electric cables and in the prediction of randomly occurring accidents. We shall look at the Poisson distribution in two distinct ways. Firstly, as a distribution in its own right. This will enable us to apply statistical methods to a set of problems which cannot be solved using the binomial distribution. Secondly, as an approximation to the binomial distribution $X \quad B(n, p)$ in the case where n is large and p is small. You will find that this approximation can often save the need to do much tedious arithmetic.

Prerequisites

Before starting this Section you should . . .

Applying condition (1) allows us to approximate terms such as $(n-1), (n-2), \ldots$ to n (mathematically, we are allowing $n \to \infty$) and the right-hand side of our expansion becomes

$$1 + np + \frac{n^2}{2!}p^2 + \cdots + \frac{n^r}{r!}p^r + \ldots$$

Note that the term $p^n \to 0$ under these conditions and hence has been omitted.
We now have the series

$$1 + np + \frac{(np)^2}{2!} + \cdots + \frac{(np)^r}{r!} + \ldots$$

which, using condition (3) may be written as

$$1 + \lambda + \frac{(\lambda)^2}{2!} + \cdots + \frac{(\lambda)^r}{r!} + \ldots$$

You may recognise this as the expansion of e^{λ}.

If we are to be able to claim that the terms of this expansion represent probabilities, we must be sure that the sum of the terms is 1. We divide by e^{λ} to satisfy this condition. This gives the result

$$\frac{e^{\lambda}}{e^{\lambda}} = 1 = \frac{1}{e^{\lambda}}(1 + \lambda + \frac{(\lambda)^2}{2!} + \cdots + \frac{(\lambda)^r}{r!} + \ldots)$$

$$= e^{-\lambda} + e^{-\lambda}\lambda + e^{-\lambda}\frac{\lambda^2}{2!} + e^{-\lambda}\frac{\lambda^3}{3!} + \cdots + e^{-\lambda}\frac{\lambda^r}{r!} + \cdots +$$

The terms of this expansion are very good approximations to the corresponding binomial expansion under the conditions

1. n is large
2. p is small
3. $np = \lambda$ (λ constant)

The Poisson approximation to the binomial distribution is summarized below.

Key Point 6

Poisson Approximation to the Binomial Distribution

Assuming that n is large, p is small and that np is constant, the terms

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

of a binomial distribution may be closely approximated by the terms

$$P(X = r) = e^{-\lambda}\frac{\lambda^r}{r!}$$

of the Poisson distribution for corresponding values of r.

Example 12

We introduced the binomial distribution by considering the following scenario. 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.

Suppose now that a similar machine which is known to produce 1% defective components is used for a production run of 40 components. We wish to calculate the probability that two defective items are produced. Essentially we are assuming that $X \quad B(40, 0.01)$ and are asking for $\mathrm{P}(X = 2)$. We use both the binomial distribution and its Poisson approximation for comparison.

Solution

Using the binomial distribution we have the solution

$$\mathrm{P}(X = 2) = {}^{40}C_2(0.99)^{40-2}(0.01)^2 = \frac{40 \times 39}{1 \times 2} \times 0.99^{38} \times 0.01^2 = 0.0532$$

Note that the arithmetic involved is unwieldy. Using the Poisson approximation we have the solution

$$\mathrm{P}(X = 2) = \mathrm{e}^{-0.4}\frac{0.4^2}{2!} = 0.0536$$

Note that the arithmetic involved is simpler and the approximation is reasonable.

Practical considerations

In practice, we can use the Poisson distribution to very closely approximate the binomial distribution provided that the product np is constant with

$n \geq 100$ and $p \leq 0.05$

Note that this is not a hard-and-fast rule and we simply say that

'the larger n is the better and the smaller p is the better provided that np is a sensible size.'

The approximation remains good provided that $np < 5$ for values of n as low as 20.

Mass-produced needles are packed in boxes of 1000. It is believed that 1 needle in 2000 on average is substandard. What is the probability that a box contains 2 or more defectives? The correct model is the binomial distribution with $n = 1000$, $p = \dfrac{1}{2000}$ (and $q = \dfrac{1999}{2000}$).

(a) Using the binomial distribution calculate $\mathrm{P}(X = 0)$, $\mathrm{P}(X = 1)$ and hence $\mathrm{P}(X \geq 2)$:

Your solution

Answer

$$\mathrm{P}(X = 0) = \left(\frac{1999}{2000}\right)^{1000} = 0.60645$$

$$\mathrm{P}(X = 1) = 1000\left(\frac{1999}{2000}\right)^{999} \times \left(\frac{1}{2000}\right) = \frac{1}{2}\left(\frac{1999}{2000}\right)^{999} = 0.30338$$

$\therefore \quad \mathrm{P}(X = 0) + \mathrm{P}(X = 1) = 0.60645 + 0.30338 = 0.90983 \quad 0.9098 \quad$ (4 d.p.)

Hence P(2 or more defectives) $\quad 1 - 0.9098 = 0.0902$.

(b) Now choose a suitable value for λ in order to use a Poisson model to approximate the probabilities:

Your solution

$\lambda =$

Answer

$$\lambda = np = 1000 \times \frac{1}{2000} = \tfrac{1}{2}$$

Now recalcula1

In the above Task we have obtained the same answer to 4 d.p., as the exact binomial calculation, essentially because p was so small. We shall not always be so lucky!

Example 13

In the manufacture of glassware, bubbles can occur in the glass which reduces the status of the glassware to that of a 'second'. If, on average, one in every 1000 items produced has a bubble, calculate the probability that exactly six items in a batch of three thousand are seconds.

Solution

Suppose that X = number of items with bubbles, then $X \quad B(3000, 0.001)$

Since $n = 3000 > 100$ and $p = 0.001 < 0.005$ we can use the Poisson distribution with $\lambda = np = 3000 \times 0.001 = 3$. The calculation is:

$$\mathrm{P}(X = 6) = \mathrm{e}^{-3}\frac{3^6}{6!} \approx 0.0498 \times 1.0125 \approx 0.05$$

The result means that we have about a 5% chance of finding exactly six seconds in a batch of three thousand items of glassware.

Example 14

A manufacturer produces light-bulbs that are packed into boxes of 100. If quality control studies indicate that 0.5% of the light-bulbs produced are defective, what percentage of the boxes will contain:

(a) no defective? (b) 2 or more defectives?

Solution

As n is large and p, the P(defective bulb), is small, use the Poisson approximation to the binomial probability distribution. If X = number of defective bulbs in a box, then

$$X \quad \mathrm{P}(\mu) \text{ where } \mu = n \times p = 100 \times 0.005 = 0.5$$

(a) $\mathrm{P}(X = 0) = \dfrac{e^{-0.5}(0.5)^0}{0!} = \dfrac{e^{-0.5}(1)}{1} = 0.6065 \approx 61\%$

(b) $\mathrm{P}(X = 2 \text{ or more}) = \mathrm{P}(X = 2) + \mathrm{P}(X = 3) + \mathrm{P}(X = 4) + \ldots$ but it is easier to consider:

$$\mathrm{P}(X \geq 2) = 1 - [\mathrm{P}(X = 0) + \mathrm{P}(X = 1)]$$

$$\mathrm{P}(X = 1) = \frac{e^{-0.5}(0.5)^1}{1!} = \frac{e^{-0.5}(0.5)}{1} = 0.3033$$

i.e. $\mathrm{P}(X \geq 2) = 1 - [0.6065 + 0.3033] = 0.0902 \approx 9\%$

2. The Poisson distribution

The Poisson distribution is a probability model which can be used to find the probability of a single event occurring a given number of times in an interval of (usually) time. The occurrence of these events must be determined by chance alone which implies that information about the occurrence of any one event cannot be used to predict the occurrence of any other event. It is worth noting that only the *occurrence* of an event can be counted; the *non-occurrence* of an event cannot be counted. This contrasts with Bernoulli trials where we know the number of trials, the number of events occurring and therefore the number of events not occurring.

The Poisson distribution has widespread applications in areas such as analysing tra□c flow, fault prediction in electric cables, defects occurring in manufactured objects such as castings, email messages arriving at a computer and in the prediction of randomly occurring events or accidents. One well known series of accidental events concerns Prussian cavalry who were killed by horse kicks. Although not discussed here (death by horse kick is hardly an engineering application of statistics!) you will find accounts in many statistical texts. One example of the use of a Poisson distribution where the events are not necessarily time related is in the prediction of fault occurrence along a long weld - faults may occur anywhere along the length of the weld. A similar argument applies when scanning castings for faults - we are looking for faults occurring in a volume of material, not over an interval if time.

The following definition gives a theoretical underpinning to the Poisson distribution.

Definition of a Poisson process

Suppose that events occur at random throughout an interval. Suppose further that the interval can be divided into subintervals which are so small that:

Key Point 7

The Poisson Probabilities

If X is the random variable

'number of occurrences in a given interval'

for which the average rate of occurrence is λ then, according to the **Poisson** model, the probability of r occurrences in that interval is given by

$$\mathrm{P}(X = r) = e^{-\lambda}\,\frac{\lambda^r}{r!} \qquad r = 0, 1, 2, 3, \ldots$$

Using the Poisson distribution $\mathrm{P}(X = r) = \mathrm{e}^{-\lambda}\frac{\lambda^r}{r!}$ write down the formulae for $\mathrm{P}(X = 0)$, $\mathrm{P}(X = 1)$, $\mathrm{P}(X = 2)$ and $\mathrm{P}(X = 6)$, noting that $0! = 1$.

Your solution

$\mathrm{P}(X = 0) =$

$\mathrm{P}(X = 1) =$

$\mathrm{P}(X = 2) =$

$\mathrm{P}(X = 6) =$

Answer

$$\mathrm{P}(X = 0) = e^{-\lambda} \times \frac{\lambda^0}{0!} = e^{-\lambda} \times \frac{1}{1} \equiv e^{-\lambda} \qquad \mathrm{P}(X = 1) = e^{-\lambda} \times \frac{\lambda}{1!} = \lambda e^{-\lambda}$$

$$\mathrm{P}(X = 2) = e^{-\lambda} \times \frac{\lambda^2}{2!} = \frac{\lambda^2}{2}e^{-\lambda} \qquad \mathrm{P}(X = 6) = e^{-\lambda} \times \frac{\lambda^6}{6!} = \frac{\lambda^6}{720}e^{-\lambda}$$

Calculate $P(X = 0)$ to $P(X = 5)$ when $\lambda = 2$, accurate to 4 d.p.

Your solution

Answer

r	0	1	2	3	4	5
$P(X = r)$	0.1353	0.2707	0.2707	0.1804	0.0902	0.0361

Notice how the values for $P(X = r)$ in the above answer increase, stay the same and then decrease relatively rapidly (due to the significant increase in $r!$ with increasing r). Here two of the probabilities

Example 15

Calculate the value for $P(X = 6)$ to extend the Table in the previous Task using the recurrence relation and the value for $P(X = 5)$.

Example 17

The mean number of bacteria per millilitre of a liquid is known to be 6. Find the probability that in 1 ml of the liquid, there will be:

(a) 0, (b) 1, (c) 2, (d) 3, (e) less than 4, (f) 6 bacteria.

Exercises

1. Large sheets of metal have faults in random positions but on average have 1 fault per 10 m^2.

   What is the probability that a sheet 5 m × 8 m will have at most one fault?

2. If 250 litres of water are known to be polluted with 10^6 bacteria what is the probability that a sample of 1 cc of the water contains no bacteria?

3. Suppose vehicles arrive at a signalised road intersection at an average rate of 360 per hour and the cycle of the tra□c lights is set at 40 seconds. In what percentage of cycles will the number of vehicles arriving be (a) exactly 5, (b) less than 5? If, after the lights change to green, there is time to clear only 5 vehicles before the signal changes to red again, what is the probability that waiting vehicles are not cleared in one cycle?

4. Previous results indicate that 1 in 1000 transistors are defective on average.

   (a) Find the probability that there are 4 defective transistors in a batch of 2000.

   (b) What is the largest number, N, of transistors that can be put in a box462(t)462i003a655210(i

Exercises continued

10. A factory uses tools of a particular type. From time to time failures in these tools occur and they need to be replaced. The number of such failures in a day has a Poisson distribution with mean 1.25. At the beginning of a particular day there are five replacement tools in stock. A new delivery of replacements will arrive after four days. If all five spares are used before the new delivery arrives then further replacements cannot be made until the delivery arrives. Find

    (a) the probability that three replacements are required over the next four days.

    (b) the expected number of replacements actually made over the next four days.

Answers

1. Poisson Process. In a sheet size 40 m^2 **we expect** 4 faults

   $\therefore \quad \lambda = 4 \qquad P(X = r) = \lambda^r e^{-\lambda}/r!$

   $P(X \le 1) = P(X = 0) + P(X = 1) = e^{-4} + 4e^{-4} = 0.0916$

2. In 1 cc **we expect** 4 bacteria(= $10^6/250000$) $\quad \therefore \quad \lambda = 4$

   $P(X = 0) = e^{-4} = 0.0183$

3. In 40 seconds **we expect** 4 vehicles $\quad \therefore \quad \lambda = 4$

   (a) P (exactly 5) = $\lambda^5 e^{-\lambda}/5! = 0.15629$ i.e. in 15.6% of cycles

   (b) P (less than 5) = $e^{-\lambda}\left(1 + \lambda^4 + \dfrac{\lambda^2}{2!} + \dfrac{\lambda^3}{3!} + \dfrac{\lambda^4}{4!}\right)$

   $= e^{-4}\left(1 + 4 + 8 + \right.$ ^{32}

Answers

5. P(defective) = 0.02. Poisson approximation to binomial $\lambda = np = 100(0.02) = 2$

   P(4 or fewer defectives in sample of 100)

$$= \mathrm{P}(X = 0) + \mathrm{P}(X = 1) + \mathrm{P}(X = 2) + \mathrm{P}(X = 3) + \mathrm{P}(X = 4)$$

$$= e^{-2} + 2e^{-2} + \frac{2^2}{2}e^{-2} + \frac{2^3}{3!}e^{-2} + \frac{2^4}{4!}e^{-2} = 0.947347$$

Inspection costs

Cost c	75	75×50
$\mathrm{P}(X = c)$	0.947347	0.0526

E(Cost) = 75(0.947347) + 75 × 50(0.0526) = 268.5 p

6. (a) 0.51342 (b) 0.00056, (c) 0.14430, (d) 0.14430

7. (a) 0.12038, (b) 0.10698, (c) 0.15552, (d) 0.84448

8. (a) 0.00016, (b) 0.00767, (c) 0.14379, (d) 0.44933, (e) 0.95258, (f) 0.19121

9. Let X be the total number of failures.

   (a) $\mathrm{E}(X) = 10 \times 0.4 = 4.$

   (b) $\mathrm{P}(X < 2) = \mathrm{P}(X = 0) + \mathrm{P}(X = 1) = e^{-4} + 4e^{-4} = 5e^{-4} = 0.0916.$

10. Let the number required over 4 days be X. Then $\mathrm{E}(X) = 4 \times 1.25 = 5$ and X Poisson(5).

   (a) $\mathrm{P}(X = 3) = \dfrac{e^{-5}5^3}{3!} = 0.1404.$

   (b) Let R be the number of replacements made.

$$\mathrm{E}(R) = 0 \times \mathrm{P}(X = 0) + \cdots + 4 \times \mathrm{P}(X = 4) + 5 \times \mathrm{P}(X \geq 5),$$

   and

$$\mathrm{P}(X \geq 5) = 1 - [\mathrm{P}(X = 0) + \cdots + \mathrm{P}(X = 4)]$$

so

$$\begin{aligned}\mathrm{E}(R) &= 5 - 5 \times \mathrm{P}(X = 0) - \cdots - 1 \times \mathrm{P}(X = 4)\\ &= 5 - e^{-5}\ 5 \times \frac{5^0}{0!} + 4 \times \frac{5^1}{1!} + \cdots + 1 \times \frac{5^4}{}\end{aligned}$$

The Hypergeometric Distribution

The calculations involved when using the hypergeometric distribution are usually more complex than their binomial counterparts.

If we sample without replacement we may proceed in general as follows:

- we may select n items from a population of N items in ${}^{N}C_{n}$ ways;
- we may select r defective items from M defective items in ${}^{M}C_{r}$ ways;
- we may select $n - r$ non-defective items from $N - M$ non-defective items in

It is possible to derive formulae for the mean and variance of the hypergeometric distribution. However, the calculations are more di□cult than their binomial counterparts, so we will simple state the results.

Key Point 11

Expectation and Variance of the Hypergeometric Distribution

The expectation (mean) and variance of the hypergeometric random variable

In the manufacture of car tyres, a particular production process is know to yield 10 tyres with defective walls in every batch of 100 tyres produced. From a production batch of 100 tyres, a sample of 4 is selected for testing to destruction. Find:

(a) the probability that the sample contains 1 defective tyre

(b) the expectation of the number of defectives in samples of size 4

(c) the variance of the number of defectives in samples of size 4.

Your solution

Answer

Sampling is clearly without replacement and we use the hypergeometric distribution with $N = 100$, $M = 10$, $n = 4$, $r = 1$ and $p = 0.1$. Hence:

(a) $\mathrm{P}(X = r) = \dfrac{{}^{M}C_r \times {}^{N-M}C_{n-r}}{{}^{N}C_n}$ gives

$$\mathrm{P}(X = 1) = \frac{{}^{10}C_1 \times {}^{100-10}C_{4-1}}{{}^{100}C_4} = \frac{10 \times 117480}{3921225} \approx 0.3$$

(b) The expectation is $\mathrm{E}(X) = np = 4 \times 0.1 = 0.4$

(c) The variance is $\mathrm{V}(X) = np(1-p)\dfrac{N-M}{N-1} = 0.4 \times 0.9 \times \dfrac{90}{99} \approx 0.33$

Answer

(a) Let X = the numbers of defectives in a sample. Then

$$\mathrm{P}(X = d) = \frac{{}^{45}C_{10-d} \times {}^{5}C_{d}}{{}^{50}C_{10}}$$

Hence

$$\mathrm{P}(X = 0) = \frac{{}^{45}C_{10} \times {}^{5}C_{0}}{{}^{50}C_{10}} = 0.311 \qquad \mathrm{P}(X = 1) = \frac{{}^{45}C_{9} \times {}^{5}C_{1}}{{}^{50}C_{10}} = 0.431$$

$$\mathrm{P}(X = 2) = \frac{{}^{45}C_{8} \times {}^{5}C_{2}}{{}^{50}C_{10}} = 0.210 \qquad \mathrm{P}(X = 3) = \frac{{}^{45}C_{7} \times {}^{5}C_{3}}{{}^{50}C_{10}} = 0.044$$

$$\mathrm{P}(X = 4) = \frac{{}^{45}C_{6} \times {}^{5}C_{4}}{{}^{50}C_{10}} = 0.004 \qquad \mathrm{P}(X = 5) = \frac{{}^{45}C_{5} \times {}^{5}C_{5}}{}$$

C

Answer

(a) Let the number of below-standard components in the sample be X. The probability of acceptance is

$$\begin{aligned} \mathrm{P}(X=0)+\mathrm{P}(X=1) &= \frac{\binom{14}{5}\binom{6}{0}}{\binom{20}{5}} + \frac{\binom{14}{4}\binom{6}{1}}{\binom{20}{5}} \\ &= \frac{14}{5} \times \frac{13}{4} \times \frac{12}{3} \times \frac{11}{2} \times 10 \end{aligned}$$