Discrete Random Variables

Prerequisites

You should be familiar with the use of probability trees to solve elementary problems in probability. You should also be familiar with calculating the mean and variance (standard deviation) from a frequency table.

Discrete and continuous variables

A *variable* is an expression X that can take a value. For example, the length of tables in a classroom is a variable; the particular lengths that the tables have are the values of this variable. Variables can take either *discrete* or *continuous* values. Informally, the term *continuous* means that between any two values that the variable can take there is another value. For example, one table might be 1.0 metres long, and another 1.2 metres long, but it is possible for a third table to be 1.1 metres long, or any length lying between 1.0 and 1.2 metres. The term *discrete* means the values that a variable can take are separated and distinct from one another. For example, the number of books on a shelf can be 2 or 3, but not 2.5. The values 2 and 3 are cut off, separated from each other. The variable either takes one value or the other, but not any value in between.

Example (1)

Classify the following variables as either discrete or continuous.

(*a*) The number of cars in a car park.

(*b*) The number of fish in the lake.

(*c*) The weight of rainfall over Scandinavia in one month.

Solution

(*a*) Discrete

(*b*) Discrete

(*c*) Continuous

(*d*) Continuous

In this chapter we shall be concerned with variables that take discrete values. They are called *discrete variables*.

Random variables

Let X be a variable. Then X is an expression that takes values. We will denote the expression, "X takes the value 1" by $X = 1$. A variable may take more than one value. When the variable is discrete then we denote the 1st, 2nd, 3rd,... ith values by $x_1, x_2, x_3, ..., x_i$. The expression, "X takes the 1st value" is denoted $X = x_1$ and the expression X takes the ith value is denoted $X = x_i$.

Example (2)

Translate the following into symbols

(a) The number of cars X in the parking lot is 52.

(b) The number of fish Y in the lake is y_i.

Solution

(a) $X = 52$

(b) $Y = y_i$

Example (3)

In a bag there are four balls. Each ball has a number written on it. The number on the two of the balls is 1. The number on one ball is 2. The number on the last ball is 3. In an experiment two balls are drawn from the bag at random in two successive trials without replacement. Let X be the variable representing the sum of the numbers on the two balls chosen.

(a) Draw a probability tree for the experiment.

(b) Find the probability that

(i) $X < 2$

(ii) $X = 2$

(iii) $X = 3$

(iv) $X = 4$

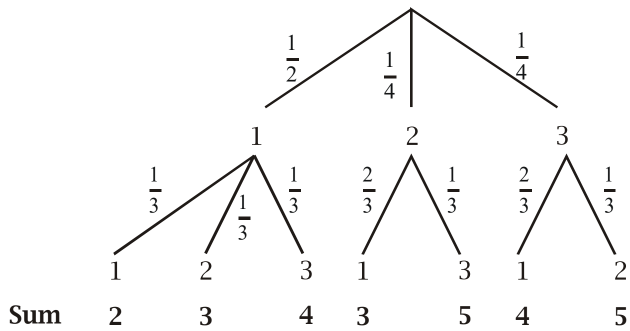
(v) $X = 5$

(vi) $X > 5$

(c) Find the sum of all the probabilities found in (b).

Solution

(a)



(b) Here we denote the expression "the probability that $X > 0$" by $P(X < 2)$

(i) $P(X<2)=0$

(ii) $P(X=2)=\frac{1}{6}$

(iii) $P(X=3)=\frac{1}{3}$

(iv) $P(X=4)=\frac{1}{3}$

(v) $P(X=5)=\frac{1}{6}$

(vi) $P(X>5)=0$

(c) The sum of all these probabilities is

$$\frac{1}{6}+\frac{1}{3}+\frac{1}{3}+\frac{1}{6}=1$$

The answer to part (c) of the above example agrees with the law of total probability – that the sum of the probabilities of all possible (mutually exclusive) events is equal to 1.

A variable X whose values are determined by chance is called a *random variable*. In the above example (3) the possible values of the variable X are 2, 3, 4 and 5. The actual event that occurs is determined by chance. Each event has associated with it a probability. The probability that X takes the value 2 is $\frac{1}{6}$. We write this as $P(X=2)=\frac{1}{6}$. The four possible events $(X=2,\ X=3,\ X=4 \text{ and } X=5)$ are mutually exclusive. Since each of these four mutually exclusive events is determined by chance the sum of their probabilities must obey the law of total probability, in other words, be equal to 1. In example (3) the random variable X also takes discrete values. Therefore, it is an example of a *discrete random variable*

Example (3) continued

Using the information from the preceding part of the question, complete the following table showing the *probability distribution* of the discrete random variable *X*.

$X = x_i$	2	3	4	5
$P(X = x_i)$	$\frac{1}{6}$			

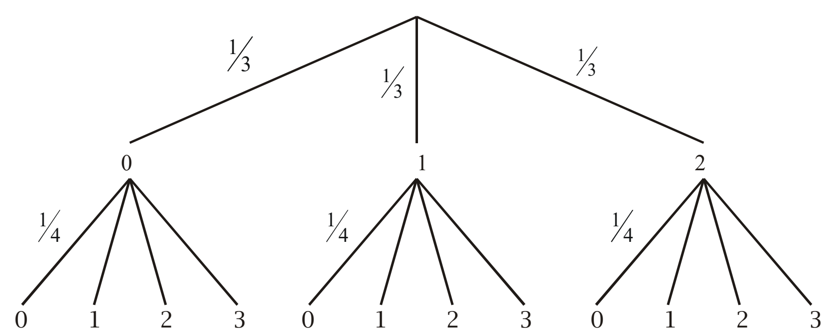
Solution

$X = x_i$	2	3	4	5
$P(X = x_i)$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$

Example (4)

Two bags contain a mixture of balls. In the first bag there are three balls having the numbers 0, 1 and 2; in the second bag there are four balls, numbered 0, 1, 2 and 3. One ball is drawn from the first bag, and another ball from the second. Write down the discrete probability distribution of the product of the two values.

Solution



Product **0** **0** **0** **0** **0** **1** **2** **3** **0** **2** **4** **6**

For the balls in the first bag

$P(X_1 = 0) = P(X_1 = 1) = P(X_1 = 2) = 1/3$

For the balls in the second bag

$P(X_2 = 0) = P(X_2 = 1) = P(X_2 = 2) = P(X_2 = 3) = 1/4$

Let X represent the discrete random variable for the product of the numbers of the two balls. Then $X = X_1 \times X_2$.

This gives the following probabilities

$P(X=0)=\frac{6}{12}=\frac{1}{2}$

$P(X=1)=\frac{1}{12}$

$P(X=2)=\frac{2}{12}=\frac{1}{6}$

$P(X=3)=\frac{1}{12}$

$P(X=4)=\frac{1}{12}$

$P(X=6)=\frac{1}{12}$

In tabular form

$X=x_i$	0	1	2	3	4	6
$P(X=x_i)$	$\frac{1}{2}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$

Formal definition of a discrete random variable

Let X be a variable such that

(*a*) It is discrete, meaning it can only take n exact values $x_1, x_2, \ldots, x_n$. When X takes the value x_i we write $X=x_i$.

(*b*) It is random, meaning that with each value that the variable takes, there is associated a probability p_i. We write this $P(X=x_i)=p_i$, which is read, "The probability that the random variable X takes the value x_i is p_i".

(*c*) Because it is random it obeys the law of total probability. The sum of all the probabilities for all n values is equal to 1.

A function that assigns to each value of a random variable X a probability is called a *probability density function,* which is abbreviated to *pdf.* If X is also discrete the pdf is called a *discrete probability density function.* A table giving the values of the pdf for each value x_i of the variable X is called a *discrete probability distribution.*

Example (5)

Which of the following tables represents a discrete probability function?

(*a*)

$X=x_i$	0	1	2	3	4
$P(X=x_i)$	$\frac{1}{4}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$

(*b*)

$X = x_i$	0	1	2	3	4
$P(X = x_i)$	$\frac{1}{2}$	$\frac{3}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$

Solution

For (*a*) the sum of the entries in the second row is

$$\frac{1}{4}+\frac{1}{16}+\frac{1}{8}+\frac{1}{8}+\frac{1}{16}=\frac{5}{8}\neq 1$$

As this does not sum to 1, it cannot be a pdf (probability density function).

For (*b*) the sum is

$$\frac{1}{2}+\frac{3}{16}+\frac{1}{8}+\frac{1}{8}+\frac{1}{16}=1$$

so this is a pdf.

In this last example table (*b*) defines a probability density function of a discrete random variable by explicitly listing the probabilities that it takes for each of its values. A pdf (probability density function) may also be given by a rule.

Example (6)

A discrete random variable X has probability density function is given by

$P(X = x) = \frac{1}{k}x^2$ for $x = 1, 2, 3, 4$

$P(X = x) = 0$ otherwise

(*a*) Show that $k = 30$.

(*b*) Write out the discrete probability distribution of X as a table.

Solution

(*a*) Since this is a pdf it obeys the law of total probability. Therefore

$$\frac{1}{k}\left(1^2+2^2+3^2+4^2\right)=1$$
$$1+4+9+16=k$$
$$k=30$$

(*b*)

$X = x_i$	1	2	3	4
$P(X = x_i)$	$\frac{1}{30}$	$\frac{2}{15}$	$\frac{9}{30}$	$\frac{8}{15}$

Example (7)

The discrete random variable X has the distribution given by the table below. Find the value of d.

x	0	1	2	3	4	5
$P(X = x)$	0.3	0.2	d	0.1	0.1	0.1

Solution

The sum of all the probabilities must equal 1

$$0.3 + 0.2 + d + 0.1 + 0.1 + 0.1 = 1$$
$$d = 1 - (0.3 + 0.2 + 0.1 + 0.1 + 0.1)$$
$$d = 0.2$$

Remark

Using the symbol $\sum$ to denote a sum, the expression "the sum of all the probabilities equals 1" is written $\sum P(X = x) = 1$. This is the technical way of writing sums. You should learn to use it if you are intending to study mathematics or statistics to a higher level than, say, represented by this chapter. It is not really possible to understand the optional proofs subsequently given below without grasp of the meaning of this symbol.

The uniform distribution of a discrete random variable

Let X be a discrete random variable with n values $x_1, x_2, \ldots, x_n$. Let the probability density function of X assign the probability $P(X = x_i) = \frac{1}{n}$ for each of these n values. That is, the probability is the same for each value that X can take. Then the discrete random variable X is said to follow a *uniform distribution*.

Example (8)

A fair cubical die is to be thrown once. Let X be the discrete random variable that represents the score of the die.

(a) What is the probability that $X = 6$?

(b) Write out the probability distribution of X as a table of values and assigned probabilities.

(c) Write a definition of the probability density function in the form $P(X = x_i) = \frac{1}{n}$ that gives this probability distribution.

(d) What is the background assumption being made throughout when answering the first three parts of this question?

Solution

(*a*) $P(X=6)=\frac{1}{6}$

(*b*)

$X=x_i$	1	2	3	4	5	6
$P(X=x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

(*c*) $P(X=x)=\frac{1}{6}$ for $x=1, 2, 3, 4, 5, 6$

$P(X=x)=0$ otherwise

(*d*) We are assuming that all outcomes are equally likely.

In example (7) the probability distribution is uniform. The assumption that all outcomes are equally likely, contained in the question by the description of the die as *fair*, makes it possible for us to construct this as the *theoretical* distribution for one throw of a cubical die.

Comparing theory with experiment

Probabilities can be determined by experiment, or from theory. In example (7) we obtained the theoretical distribution for one throw of a fair cubical die as

$X=x_i$	1	2	3	4	5	6
$P(X=x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Example (9)

The cubical die of example (7) is to be thrown 120 times.

What are the expected frequencies that X shall take each of the values $x=1, 2, 3, 4, 5, 6$?

Solution

$X=x_i$	1	2	3	4	5	6
$P(X=x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
Expected frequency	20	20	20	20	20	20

As the example shows to find the expected frequency we multiply the probability $P(X = x_i)$ by the total number of trials, *n*.

Experimental probability

When probabilities are determined by experiment we begin by making observations of the values that the variable takes. This is called a *sample*. If *n* observations are made then the *sample size* is *n*. To derive the experimental probability p_i that *X* takes the value x_i, $P(X = x_i) = p_i$, we observe the number of times *X* takes that value. This number is called the *frequency* that $X = x_i$ and is denoted f_i. Then the experimental probability $P(X = x_i) = p_i$ is the ratio $p_i = \frac{f_i}{n}$. A set of frequencies for each value that the variable can take is called a *frequency distribution*.

Example (10)

A cubical die was thrown 120 times and the results recorded in the following frequency distribution.

Score	1	2	3	4	5	6
Frequency, *f*	19	23	17	17	21	23

Let *X* be the discrete random variable that represents the score of the die. Find the experimental probability corresponding to each value of *X*. Give your answers to 3 significant figures. How does this experimental distribution differ from the theoretical distribution?

Solution

For each entry in the table we divide the frequency by the total sample size. For example

$P(X = 1) = \frac{18}{120} = 0.15$

X	1	2	3	4	5	6
f_i	19	23	17	17	21	23
$P(X = x_i)$	0.158	0.192	0.142	0.142	0.175	0.192
Expected frequencies	20	20	20	20	20	20

The experimental frequencies are not exactly the same as the expected frequencies. Consequently, the experimental probabilities also differ from the expected probabilities. The experimental probability distribution is not uniform.

Remarks

(1) If we sum the probabilities in this table we get

$$\begin{aligned}\text{Sum} &= 0.158 + 0.192 + 0.142 + 0.142 + 0.175 + 0.192 \\ &= 1.001 \\ &\neq 1\end{aligned}$$

This is because of rounding errors in taking the probabilities to 3 significant figures.

(2) The fact that the experimental probability distribution is not uniform should not surprise us. Owing to chance factors we expect some *variation* (divergences) from the expected values when in practice we determine experimental frequencies. The difference between the theoretical and the experimental result does *not prove* that the *true* probability distribution for the score of the die is not uniform. If the experimental frequencies differed substantially from the expected frequencies, then we would begin to suspect that the die was not fair.

These examples illustrate the need to distinguish between theoretical and experimental values for a probability distribution.

Mean and variance

You should already be familiar with the process of finding the mean and variance from a frequency table.

Example (10) continued

Calculate the mean, variance and standard deviation for the grouped frequency table below showing 120 throws of a cubical die. Give your answers to 3 significant figures.

Score	1	2	3	4	5	6
Frequency, f	19	23	17	17	21	23

Solution

The total sample size is $n = 120$

For the mean

$$\bar{x} = \frac{\text{Sum}(\text{value } \times \text{ frequency})}{n} \qquad \left[\bar{x} = \frac{\sum xf}{\sum f}\right]$$

$$= \frac{(1\times19)+(2\times23)+(3\times17)+(4\times17)+(5\times21)+(6\times23)}{120}$$

$$= \frac{427}{120}$$

$$= 3.56 \text{ (3 s.f.)}$$

For the variance

$$S^2 = \frac{\text{Sum}(\text{Square of values } \times \text{ frequency})}{n} - \text{square of the mean}^2 \qquad \left[S^2 = \frac{\sum x^2 f}{\sum f} - (\bar{x})^2\right]$$

$$= \frac{(1^2\times19)+(2^2\times23)+(3^2\times17)+(4^2\times17)+(5^2\times21)+(6^2\times23)}{120} - \left(\frac{427}{120}\right)^2$$

$$= \frac{1889}{120} - \left(\frac{421}{120}\right)^2$$

$$= 3.08 \text{ (3 s.f.)}$$

The standard deviation is the square root of the variance

$$S = \sqrt{3.08...} = 1.75 \text{ (3 s.f.)}$$

In this example we have used the symbols $\bar{x}$, S^2 and S to denote the *experimental* values of the mean, variance and standard deviation of a frequency distribution.

Example (11)

A man intends to throw a fair cubical die 120 times.

(*a*) What is the theoretical average (mean) of all the scores he obtains?

(*b*) What is the theoretical variance of all the scores he obtains?

Solution

This is calculated from the theoretical distribution for the cubical die, which we showed earlier to be

$X = x_i$	1	2	3	4	5	6
$P(X = x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
Predicted frequency	20	20	20	20	20	20

We can use the predicted frequencies for the calculations.

Mean

$$\text{Theoretical mean} = \frac{\text{Sum}(\text{value } \times \text{ frequency})}{n} \qquad \left[\bar{x} = \frac{\sum xf}{\sum f}\right]$$

$$= \frac{(1\times 20)+(2\times 20)+(3\times 20)+(4\times 20)+(5\times 20)+(6\times 20)}{120}$$

$$= \frac{420}{120}$$

$$= 3.5$$

Variance

$$\text{Theoretical variance} = \frac{\text{Sum}(\text{Square of values } \times \text{ frequency})}{n} - \text{square of the expected mean}$$

$$= \frac{(1^2\times 20)+(2^2\times 20)+(3^2\times 20)+(4^2\times 20)+(5^2\times 20)+(6^2\times 20)}{120} - \left(\frac{7}{2}\right)^2$$

$$= \frac{1820}{120} - \left(\frac{7}{2}\right)^2$$

$$= \frac{35}{12} \quad (2.92 \text{ to } 3 \text{ s.f.})$$

We can compare these values with the experimental ones found earlier

	mean	variance
Theoretical	3.51	3.33
Experimental	3.5	2.92

The experimental mean was very close to the theoretical mean, and there was slightly less variance than the theoretical variance.

Example (11) continued

Was the die a fair die?

Solution

At this stage we can only give a *subjective* answer to this question. Looking at the data the differences between the theoretical and experimental values *seem* insignificant. We would conclude that this is a fair die and that all the scores are equally likely to come up.

The solution to example (11) introduces a topic for subsequent consideration. At present our concern is primarily with evaluating the differences between theoretical and experimental distributions. We should have symbols for the terms "theoretical mean" and "theoretical variance". Furthermore, since they represent different concepts from "experimental mean", denoted $\overline{x}$, and "experimental variance", denoted S^2, we should use different symbols. We use $E(X)$ for theoretical mean and $\text{var}(X)$ for theoretical variance. The theoretical mean $E(X)$ is also called the *expectation*.

It is not necessary to generate theoretical frequencies in order to find $E(X)$ and $\text{var}(X)$ for a theoretical probability distribution. A probability distribution takes the form

$X = x_i$	x_1	x_2	....	x_n
$P(X = x_i)$	$P(X = x_1)$	$P(X = x_2)$		$P(X = x_n)$

The expectation for a theoretical discrete probability distribution is found by

$E(X) = \text{Sum}(\text{value} \times \text{ expected probability})$

In symbols this is

$E(X) = \sum x_i P(X = x_i)$.

Example (12)

The discrete random variable X has the following distribution.

i	0	1	2	3	4	5	6	7
$P(X = i)$	0.05	0.05	0.15	0.25	0.25	0.1	0.1	0.05

Find $E(X)$

Solution.

$$\begin{aligned} E(X) &= \text{Sum}(\text{value} \times \text{ expected probability}) \\ &= (1 \times 0.05) + (2 \times 0.15) + (3 \times 0.25) + (4 \times 0.25) + (5 \times 0.1) + (6 \times 0.1) + (7 \times 0.05) \\ &= 0.05 + 0.30 + 0.75 + 1.00 + 0.5 + 0.6 + 0.35 \\ &= 3.55 \end{aligned}$$

In this process we treat the probabilities as if they were frequencies, and we calculate the mean in exactly the same way we would calculate the mean for a frequency table, noting that the sum of the frequencies (here probabilities) is 1. The difference is that the frequencies used in this calculation derive from theory rather than observation. To find the formula for the expected variance we employ the same method. Namely, we evaluate the variance of the theoretical distribution by treating it as if it were a frequency table where the frequencies are given by the probabilities. We use the formula

$$\text{var}(X) = \frac{\text{Sum}(\text{Square of expected values } \times \text{ expected probability})}{n} - \text{square of the expected mean}$$

The first part of this formula is the expression

$$E(X^2) = \frac{\text{Sum}(\text{Square of expected values } \times \text{ expected probability})}{n} \qquad \left[E(X^2) = \sum (x_i)^2 P(X = x_i)\right]$$

It is usual to evaluate this expression first and to calculate the theoretical variance in a second step, when

$$\text{var}(X) = E(X^2) - \left[E(X)\right]^2$$

Example (13)

Calculate the theoretical expectation and variance for the discrete random variable X whose theoretical probability distribution is:

$X = x_i$	1	2	3	4
$P(X = x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{5}{12}$

$$E(X) = \text{Sum}(\text{value } \times \text{ expected probability}) \qquad \left[E(X) = \sum x_i P(X = x_i)\right]$$

$$= \left(1 \times \frac{1}{6}\right) + \left(2 \times \frac{1}{6}\right) + \left(3 \times \frac{1}{4}\right) + \left(4 \times \frac{5}{12}\right)$$

$$= \frac{2 + 4 + 9 + 20}{12}$$

$$= \frac{35}{12}$$

$$E(X^2) = \frac{\text{Sum}(\text{Square of expected values } \times \text{ expected probability})}{n} \qquad \left[E(X^2) = \sum (x_i)^2 P(X = x_i)\right]$$

$$= \left(1 \times \frac{1}{6}\right) + \left(4 \times \frac{1}{6}\right) + \left(9 \times \frac{1}{4}\right) + \left(16 \times \frac{5}{12}\right)$$

$$= \frac{2 + 8 + 27 + 80}{12} = \frac{117}{12} = \frac{39}{4}$$

$$\text{var}(X) = E(X^2) - \left[E(X)\right]^2$$

$$= \frac{39}{4} - \left(\frac{35}{12}\right)^2$$

$$= \frac{1404 - 1225}{144}$$

$$= \frac{179}{144}$$

$$= 1.243 \quad (3 \text{ d.p.})$$

Expectation of any function of a discrete random variable

In the expression

$E(X^2) = \sum (x_i)^2 P(X = x_i)$

X^2 is a function of the discrete random variable X. If we write $g(X) = X^2$ then we get a more general form

$E(g(X)) = \sum g(x) P(X = x_i)$

Example (14)
A discrete random variable X has the following theoretical probability distribution.

$X = x_i$	1	2	3	4
$P(X = x_i)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

Find $E(g(X))$ where $g(x)$ is the function $g(x) = x^3$.

Solution

$$E(g(X)) = \sum g(x) P(X = x_i)$$
$$= 1^3 \times \left(\frac{1}{2}\right) + 2^3 \times \left(\frac{1}{4}\right) + 3^3 \times \left(\frac{1}{8}\right) + 4^3 \times \left(\frac{1}{8}\right) = \frac{1}{2} + 2 + \frac{27}{8} + 8 = 13\frac{7}{8}$$

Expectation and variance of a discrete probability distribution under a scaling and translation

Let X be a random variable. The values that X can take may be transformed by being multiplied by a constant a or by addition of a constant b.

Example (15)
A random variable X has the following probability distribution

x	16	17	18	19	20	21	22	23	24	25	26
$P(X = x)$	0.1	0.1	0.05	0.1	0.25	0.15	0.15	0.05	0	0	0.05

(*a*) Find the expected mean and variance of *X*.

(*b*) Every value of *X* is multiplied by 2 and to this 4 is added. So it is subject to a scaling of 2 followed by a translation of 4 to give another random variable *Y*.

$Y = 2X + 4$

Construct the probability distribution for *Y*.

(*c*) Find the expected mean and variance of *Y*.

Solution

(*a*) $E(X) = \text{Sum}\left(\text{value } \times \text{ expected probability}\right) \qquad \left[E(X) = \sum x_i P\left(X = x_i\right)\right]$

$$= (16 \times 0.1) + (17 \times 0.1) + (18 \times 0.05) + (19 \times 0.1) + (20 \times 0.25) + (21 \times 0.15) + (22 \times 0.15) + (23 \times 0.05) + (26 \times 0.05)$$

$$= 20$$

$$E\left(X^2\right) = \frac{\text{Sum}\left(\text{Square of values } \times \text{ probability}\right)}{n} \qquad \left[E(X^2) = \sum \left(x_i\right)^2 P\left(X = x_i\right)\right]$$

$$= \left(16^2 \times 0.1\right) + \left(17^2 \times 0.1\right) + \left(18^2 \times 0.05\right) + \left(19^2 \times 0.1\right) + \left(20^2 \times 0.25\right) + \left(21^2 \times 0.15\right) + \left(22^2 \times 0.15\right) + \left(23^2 \times 0.05\right) + \left(26^2 \times 0.05\right)$$

$$= 405.8$$

$$\text{var}(X) = E(X^2) - \left[E\left(X\right)\right]^2$$

$$= 405.8 - 400$$

$$= 5.8$$

(*b*)

x	16	17	18	19	20	21	22	23	24	25	26
$y = 2x + 4$	36	38	40	42	44	46	48	50	52	54	56
$P(X = y)$	0.1	0.1	0.05	0.1	0.25	0.15	0.15	0.05	0	0	0.05

(*c*) $E(Y) = \text{Sum}\left(\text{value } \times \text{ expected probability}\right)$

$$= (36 \times 0.1) + (38 \times 0.1) + (40 \times 0.05) + (42 \times 0.1) + (44 \times 0.25) + (46 \times 0.15) + (48 \times 0.15) + (50 \times 0.05) + (56 \times 0.05)$$

$$= 44$$

$$E\left(Y^2\right) = \frac{\text{Sum}\left(\text{Square of values } \times \text{ probability}\right)}{n}$$

$$= \left(36^2 \times 0.1\right) + \left(38^2 \times 0.1\right) + \left(40^2 \times 0.05\right) + \left(42^2 \times 0.1\right) + \left(44^2 \times 0.25\right) + \left(46^2 \times 0.15\right) + \left(48^2 \times 0.15\right) + \left(50^2 \times 0.05\right) + \left(56^2 \times 0.05\right)$$

$$= 1959.2$$

$$\text{var}(Y) = E(Y^2) - \left[E\left(Y\right)\right]^2$$

$$= 1959.2 - 44^2$$

$$= 23.2$$

It is rather tiresome to find $E(Y)$ and $\text{var}(Y)$ in this way and it would be good to have a short cut to the result. We observe above that

$$Y = 2X + 4$$
$$E(X) = 20$$
$$E(Y) = 44 = 2 \times E(X) + 4$$
$$\text{var}(X) = 5.8$$
$$\text{var}(Y) = 23.2 = 4 \times 5.8$$

This suggests a relationship between $E(Y)$ and $\text{var}(Y)$ and $E(X)$ and $\text{var}(X)$ given by transformation $Y = 2X + 4$. This is as follows

Expectation and variance under a scaling and transformation

Let X be a random variable with expectation $E(X)$. Let a and b be constants. Then

$$E(aX + b) = aE(X) + b$$
$$\text{var}(aX + b) = a^2 \text{var}(X)$$

The expected value of X – that is, the expected mean is translated by the addition of b and multiplied by the scale factor a. The variance is multiplied by the square of the of the scale factor a and unaffected by the addition of b.

Example (16)

The random variable X has mean 20 and variance 6. Find the mean and variance of the random variables

(a) $Y = 2X + 4$

(b) $Z = \frac{1}{3}X - \frac{2}{3}$

Solution

(a)
$$\begin{aligned} E(Y) &= E(2X + 4) \\ &= 2 \times E(X) + 4 \\ &= 2 \times 20 + 4 \\ &= 44 \end{aligned}$$

$$\begin{aligned} \text{var}(Y) &= \text{var}(2X + 4) \\ &= (2^2)\text{var}(X) \\ &= 4 \times 6 \\ &= 24 \end{aligned}$$

(b)
$$\begin{aligned} E(Z) &= E\left(\frac{1}{3} \times X - \frac{2}{3}\right) \\ &= \frac{1}{3} \times E(X) - \frac{2}{3} = \frac{1}{3} \times 20 - \frac{2}{3} = \frac{20}{3} - \frac{2}{3} = 6 \end{aligned}$$

$$\operatorname{var}(Z) = \operatorname{var}\left(\frac{1}{3}X - 9\right)$$
$$= \left(\frac{1}{3}\right)^2 \operatorname{var}(X) = \frac{1}{9} \times 6 = \frac{2}{3}$$

Example (17)

The discrete random variable X has probability distribution given by

$P(X = x) = \frac{1}{k}(2x+1)$ for $x = 1, 2, 3, 4, 5$

$P(X = x) = 0$ otherwise

(*a*) Show that $k = 35$.

(*b*) Find the mean and variance of X.

(*c*) Given that X_1, X_2 are two independent observations of X, evaluate

$P(X_1 + X_2 = 5)$

(*d*) The random variable Y is defined by

$Y = 3X - 3$

Find the mean and variance of Y.

Solution

By the law of total probability

(*a*) $\text{Sum}\left[\frac{1}{k}(2x+1)\right] = 1$ for $x = 1, 2, 3, 4, 5$

$\frac{1}{k}(3+5+7+9+11) = 1$

$k = 35$

(*b*)

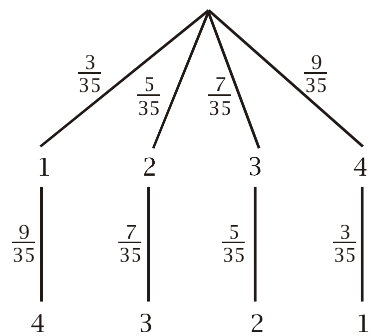
x	1	2	3	4	5
$P(X = x)$	$\frac{3}{35}$	$\frac{5}{35}$	$\frac{7}{35}$	$\frac{9}{35}$	$\frac{11}{35}$

$$E(X) = \sum xp$$
$$= \frac{1}{35}(1\times 3) + (2\times 5) + (3\times 7) + (4\times 9) + (5\times 11)$$
$$= \frac{25}{7}$$

$$E\left(X^2\right)=\sum x^2 p$$
$$=\frac{1}{35}\left(1^2\times 3\right)+\left(2^2\times 5\right)+\left(3^2\times 7\right)+\left(4^2\times 9\right)+\left(5^2\times 11\right)$$
$$=\frac{505}{35}$$
$$=\frac{101}{7}$$

$$\text{var}(X)=E(X^2)-\left[E\left(X\right)\right]^2$$
$$=\frac{101}{7}-\left(\frac{25}{7}\right)^2$$
$$=\frac{82}{49}$$

(*c*) This is stating that there are two trials of the variable *X*.



$$P\left(X_1+X_2=5\right)=2\times\left(\frac{3}{35}\times\frac{9}{35}\right)+2\times\left(\frac{5}{35}\times\frac{7}{35}\right)$$
$$=\frac{124}{1225}$$

(d) $Y=3X-3$

$$E\left(Y\right)=E\left(3X-3\right)$$
$$=3E\left(X\right)-3$$
$$=3\times\frac{25}{7}-3=\frac{54}{7}$$

$$\text{var}\left(3X-3\right)=9\,\text{var}\left(X\right)$$
$$=9\times\frac{82}{49}=\frac{738}{49}$$

Proofs

This section is optional.

<u>Expectation</u>

We will now prove the formula

$E(aX+b) = aE(X)+b$

The definition of the expected mean is

$E(X) = \sum_{\text{all } x} x\, P(X=x)$

Then, given that a and b are constants

$$\begin{aligned}
E(aX+b) &= \sum_{\text{all } x} (ax+b) P(X=x) \\
&= \sum_{\text{all } x} a\big(xP(X=x)\big) + b\,P(X=x) \\
&= \sum_{\text{all } x} a\big(xP(X=x)\big) + \sum_{\text{all } x} b\,P(X=x) \\
&= a\sum_{\text{all } x} xP(X=x) + b\sum_{\text{all } x} P(X=x) \\
&= a\,E(x)+b \qquad \text{since } \sum P(X=x)=1 \text{ and } E(X)=\sum_{\text{all } x} x\ P(X=x)
\end{aligned}$$

<u>Variance</u>

To prove the formula for variance, note that for a given value x the probability

$P(X=ax+b) = P(X=x)$

$$\begin{aligned}
\operatorname{var}(aX+b) &= \sum_{\text{all } x} (ax+b)^2\, P(X=ax+b) - \left[E(aX+b)\right]^2 \\
&= \sum_{\text{all } x} (ax+b)^2\, P(X=x) - \left[E(aX+b)\right]^2 \\
&= \sum_{\text{all } x} \left(a^2x^2+2abx+b^2\right)P(X=x) - \left\{a^2\left[E(x)\right]^2 + 2abE(x)+b^2\right\} \\
&= \sum_{\text{all } x} a^2x^2P(X=x) + \sum_{\text{all } x} 2abx\,P(X=x) + \sum_{\text{all } x} b^2P(X=x) - a^2\left[E(X)\right]^2 - 2abE(X) - b^2 \\
&= \sum_{\text{all } x} a^2\sum x^2\,P(X=x) + 2ab\sum_{\text{all } x} x\,P(X=x) + b^2\sum_{\text{all } x} P(X=x) - a^2\left[E(X)\right]^2 - 2abE(X) - b^2 \\
&= a^2E\left(X^2\right) + 2abE(X) + b^2 - a^2\left[E(X)\right]^2 - 2abE(X) - b^2 \\
&= a^2\left(E\left(X^2\right) - \left[E(X)\right]^2\right) \\
&= a^2\operatorname{var}(X)
\end{aligned}$$