M1 Maths

P6-3 Discrete Random Variables

- discrete random variables
- discrete probability distributions
- expected value and variation

Summary Learn Solve Revise Answers

Summary

A random variable is a numerical variable (like the number of spots that come up when you roll two dice) that takes on different values at random in an experiment. It can be discrete or continuous.

A discrete probability distribution (or probability function) consists of the set of all the possible outcomes of an experiment each with its probability. It can be expressed as a table or as a graph.

The expected value of a discrete random variable is the average value of the variable over a very large number of trials. It is calculated by multiplying each possible value of the variable by the probability that that value will occur and then adding the products.

To determine if a gambling game is worth playing, we can work out the expected value of the winnings and compare it to the cost of playing.

The value of a random variable can vary either side of the expected value. The variance and standard deviation are measures of this variation.

Learn

Random Variables

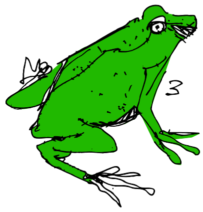
A random experiment is one where the outcome is a matter of chance and cannot be predicted with certainty. Tossing a coin is a random experiment – you can't tell if it will come down heads or tails; rolling a die is another.

A random variable is a variable whose values are numbers that correspond to the possible outcomes of a random experiment. If the experiment is rolling a die, the random variable might be the number that comes up. It can take on the values 1, 2, 3, 4, 5 or 6. In tossing a coin, the possible outcomes are H and T. This does not constitute a random variable because the results aren't numbers. But if we say a head is 1 and a tail 0, then these numbers do constitute random variable with possible values 0 and 1.

If we toss 10 coins and count the heads, then the number of heads is a random variable with possible values 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10.

These example were discrete random variables, because they can only take on certain discrete values, with values in between those not being possible. You can't get 3.7 when you roll a normal die and you couldn't get 7.19 heads if you toss 10 coins.

However, suppose the random variable was the time taken for a toad to fall off the table in a toad race. This variable can take on a continuous range of values including 7.393 seconds and 824.18054 seconds. The time taken to fall off the table is a continuous random variable.



This module focuses on discrete random variables. A later module, P6-5, deals with continuous random variables.

Practice

Q1 For each of the following, say whether it is a discrete random variable, a continuous random variable or not a random variable at all.

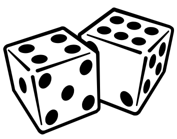
(a) the colour of the marble drawn from a bag

(b) the number of blue marbles drawn from a bag

(c) the mass of a marble drawn from a bag

(d) whether more red marbles are drawn from a bag than green marbles

(e) the sum of the numbers that come up when two dice are rolled



(f) the average of the numbers that come up on five dice rolls

(g) the height reached by a toy rocket fired from a spring launcher

(h) the floor of a building that has the largest mass of humans on it

(i) the direction a ball rolls (chosen from N, NE, E, SE etc.)

(j) the bearing on which a ball rolls

(k) the quadrant in which a point on the Cartesian plane lies

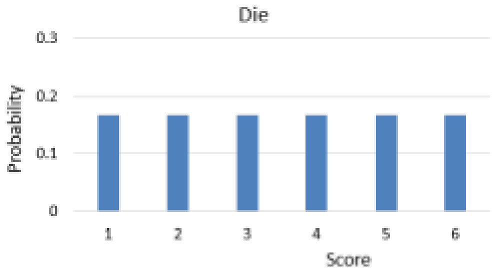
Discrete Probability Distributions

A discrete random variable has a number of possible values. For instance, when rolling a die, the score is a discrete random variable with possible values 1, 2, 3, 4, 5, 6.

Each of these values has a certain probability. We can put the values and their probabilities into a table like this:

Score	1	2	3	4	5	6
Probability	$^1/_6$	$^1/_6$	$^1/_6$	$^1/_6$	$^1/_6$	$^1/_6$

As a graph, this data would look like this:



This is the probability distribution for rolling a die. The sum of the probabilities is 1 (it must always be in any probability distribution), but the distribution shows how the probabilities are distributed across the various possible values of the random variable. In this case they are evenly distributed. This kind of even distribution where all the probabilities are the same is often called a *uniform distribution* or a *rectangular distribution*.

Let's look at a different discrete random variable – the number of heads you get when you toss 8 coins.

The possible values are 0, 1, 2, . . . 8. In this case the probabilities are distributed quite differently – not evenly at all.

Number of Hs	0	1	2	3	4	5	6	7	8
Probability	0.004	0.031	0.109	0.219	0.273	0.219	0.109	0.031	0.004

As a graph, this distribution would look like this:

This distribution is not uniform, but is more bell-shaped with a peak in the middle, dropping off either side, but with tails at both ends. It is actually an example of a binomial distribution, a common and important type of discrete probability distribution. We will look at these in more detail in the next module (P6-4 Binomial Distributions).

Probability Functions

A probability distribution is sometimes called a probability function. Look again at the probability distribution for tossing 8 coins.

Number of Hs	0	1	2	3	4	5	6	7	8
Probability	0.004	0.031	0.109	0.219	0.273	0.219	0.109	0.031	0.004

It is a relation with *number of heads* as the independent variable and *probability* as the dependent variable. The same goes for the graph. Because for each number of heads there is exactly one probability (for each value of the independent variable, there is exactly one value for the dependent variable), we can say that the relation is a function.

The same will go for any probability distribution. So *probability function* is just another name for *probability distribution*.

Practice

Q2 Draw a table and graph of the probability distribution for the number of heads when tossing 4 coins.

Q3 Draw a table and graph of the probability function for the total score when rolling 2 dice.

Q4 The probability distribution below is for the way a matchbox lands when dropped. What number should replace the x?

Flat	82%
On edge	x

On end 4%

Q5 Draw a table of the probability function for the number of children a couple would have to have in order to get a girl. List any assumptions you need to make.

Expected Value

The expected value of a random variable is its average (mean) value over a large number of trials. It is obtained by multiplying each possible value by its probability then adding them.

Consider the case of the number of heads when tossing 8 coins that we looked at above. The probability distribution is

Number of Hs	0	1	2	3	4	5	6	7	8
Probability	0.004	0.031	0.109	0.219	0.273	0.219	0.109	0.031	0.004

We can add another row with the products of the value of the variable and the probability.

x	0	1	2	3	4	5	6	7	8
$P(X = x)$	0.004	0.031	0.109	0.219	0.273	0.219	0.109	0.031	0.004
$x \times P(X = x)$	0	0.031	0.218	0.657	1.092	1.095	0.654	0.217	0.032

$\sum xP(X = x) = 3.996$

Note the change of notation here. It is common to call the random variable we are dealing with X and to call its value x. In this case, X is the variable *number of heads* and x is its value, e.g. 3, 1, 5 etc.

Then $P(X = x)$ is the probability that the variable X will take a particular value x. $x\,P(X = x)$ is x multiplied by $P(X = x)$. $\sum xP(X = x)$ is the sum of the $x\,P(X = x)$ values.

As you will see, the expected value is 4. (The difference is because of rounding errors). This is exactly what one would expect: if you toss 8 coins, on average you will get 4 heads.

Practice

Solve the following problems using tables of x, $P(X = x)$ and $x\,P(X = x)$ as shown above.

Q6 Use the table of the probability function for the total score when rolling 2 dice that you made in Q3 to find the expected value.

Q7 Use the table of the probability function for the number of children a couple would have in order to get a girl that you made for Q5 to find the expected number.

Q8 What is the expected number of children for couples who keep having children until they have one of each sex?

Q9 A bag contains 4 red balls and 4 black balls. If you take balls out without looking and without replacement, what is the average number you will have to take out to get at least one of each colour?

Games

Expected values are commonly used in working out if you will win in the long run in games.

For example, suppose you pay \$1 to roll a die, then you win \$4 if you get a 6 and \$1 if you get a 5. Would you expect to make money, lose money or break even if you played many times?

To answer this, we calculate the expected value of the winnings (often called the expected return) and see if it is more, less or the same as the cost of playing.

The expected value of the winnings is $\$4 \times {}^1/_6 + \$1 \times {}^1/_6 = \$0.83$.

So on average we will receive less than we pay out and so we will lose money in the long run – on average we will lose 17c per game.

Practice

Q10 In a game you pay a dollar for a go. A go consists of rolling a die. If you score a 6, you win \$3, if you score a 5, you win \$2. What is the expected return? Will you make money, lose money or break even in the long term? Is the game worth playing?

Q11 In another game, you pay \$1, then roll two dice. You win \$12 if you get a total of 12, \$6 for 11, \$3 for 10, \$2 for 9 and \$1 for 8. What is the expected return? Is the game worth playing?

Q12 In a game, you toss a coin until both a head and a tail have come up or until you have tossed 4 times (which ever happens first). What is the average number of tosses that you will make? [Note that *average number* and *expected number* are the same thing.]

Q13 In a game, there are 10 cards numbered 1 to 10. You pick one card and receive a number of dollars equal to the square of the number you picked. How much should it cost to play if the game is to be fair? [Fair means that in the long run you will neither win nor lose.]

Q14 In a game you roll a die. If you roll a 6, you win \$5; if you roll a 5, you win \$1; if you roll a 2, 3 or 4, you lose; if you roll a 1, you get another go. If it costs \$$x$ to play, for what values of x will you make money in the long run?

Q15 In a casino game, you put down whatever you like. A die is then rolled. If it comes up 6, you get three times your money back. If it comes up 5, you get twice your money back. Otherwise you get nothing. What is the expected return as a fraction of what you pay? Is the game worth playing?

Q16 In a game you keep picking cards from a pack until you have two the same suit. You pay \$2 per card and win \$5 when you get your two the same suit. How much would you expect to pay, i.e. what is the expected value of what you end up paying before you win? Is the game worth playing?

Q17 In a quiz game, you win some money. Then the quiz master offers you a deal. He will toss a coin. If it comes down heads, you will double your winnings. If it comes down tails, you will halve your winnings. What is your expected return if you take his offer compared to if you don't? Should you take the deal?

Q18 In another quiz game you win some money. The quiz master offers you a deal. He will roll a die. If it comes up 5 or 6, you will double your winnings. If it comes up 1, 2, 3 or 4, you will halve your winnings. What is your expected return if you take his offer compared to if you don't? Should you take the deal?

Variation

As we saw above, the expected number of heads when we toss 8 coins is 4. But we won't always get 4 heads. Sometimes we will get 3 or 5 or even 7 or 0. The number of heads will average 4, but there will be some variation. We can quantify the variation using the variance or the standard deviation. (Variance is the mean square deviation; standard deviation is the square root of the variance.)

Standard deviation and variance were dealt with in Module S4-1 (Quantiles and Spread). Go back and refresh your memory if you think you need to.

Here is the probability distribution again.

Number of Hs	0	1	2	3	4	5	6	7	8
Probability	0.004	0.031	0.109	0.219	0.273	0.219	0.109	0.031	0.004

To find the variance, we find the deviation from the mean for each number of heads, then square it, then multiply it by the probability (this is in effect a frequency), then add them. To get the standard deviation, we then take the square root of the variance.

The following is the working.

x	0	1	2	3	4	5	6	7	8
$P(X = x)$	0.004	0.031	0.109	0.219	0.273	0.219	0.109	0.031	0.004
$x - \bar{x}$	−4	−3	−2	−1	0	1	2	3	4
$(x - \bar{x})^2$	16	9	4	1	0	1	4	9	16
$P(X = x) \times (x - \bar{x})^2$	0.064	0.279	0.436	0.219	0	0.219	0.436	0.279	0.064

Variance = $\sum P(X = x)(x - \bar{x})^2 = 1.996$ Standard deviation = $\sqrt{1.996} = 1.41$

The exact variance and standard deviation are 2 and $\sqrt{2}$. Again, the difference results from rounding errors.

Variance and standard deviation can be found on a calculator with less work. Put the values for x and $P(X = x)$ into lists and select σs from the one-variable statistics.

Practice

Q19 The following table is the probability distribution for the number of heads when you toss 6 coins. Find the expected value, the variance and the standard deviation. Then check your answers by redoing the calculations with your calculator.

Number of Hs	0	1	2	3	4	5	6
Probability	0.016	0.094	0.234	0.313	0.234	0.094	0.016

Q20 The following table is the probability distribution for the number of heads when you toss a bent coin six times. Find the expected value, the variance and the standard deviation. Then check your answers by redoing the calculations with your calculator.

Number of Hs	0	1	2	3	4	5	6
Probability	0.004	0.037	0.138	0.276	0.311	0.187	0.047

Solve

Q51 In a game you pay $3 for a go. You keep tossing a coin until you get a head. If you get the head on the first toss, you win $1; if you get it on the second, you win $2; on the third, you win $3; on the fourth, you win $4; and so on. What is the expected return? Is the game worth playing?

Q52 This is a challenging and somewhat baffling question. Don't look at the answer until you have had a good think about the question.

In a quiz show, if the contestant answers a set of questions correctly, she gets to choose from 2 envelopes, one yellow, one white. Each contains a cheque. One cheque is for twice as much money as the other. The contestant chooses the white envelope and looks inside. There is a cheque for $40. Then the quizmaster says 'You can change to the yellow one if you like, but if you don't like it, you can't change back. The contestant thinks for a while and decides that, if she changes, she has a 50% chance of getting $80 and a 50% of getting $20. This is an expected return of $50. This is better than the $40 she already has, so she decides to swap. The quizmaster then asks her if she would have swapped if there had been $200 in there. She repeats the calculation and realises that if she swaps, she will have an expected return of $250 (50% of $400 plus 50% of $100). In fact she realises that whatever amount was in there, if she swapped, she would average 25% more. 'So you are always better off swapping to the yellow one?' the quizmaster asks. 'Yes' says the contestant. 'Then why didn't you take the yellow one in the first place?' he asks. The contestant looks at the quiz master with a mixture of suspicion and puzzlement. Should she have taken the yellow one in the first place?

Revise

Revision Set 1

Q61 For each of the following, say whether it is a discrete random variable, a continuous random variable or not a random variable at all.

(a) the total score from rolling 5 dice

(b) which colour die ends up closest to the edge of the table when rolled

(c) the average distance of the dice from the edge of the table

Q62 Find the probability distribution for the total score obtained when rolling two 4-sided dice with sides numbered 1 to 4. Present it as a table and as a graph.

Q63 Complete the following probability function:

Score	1	2	3	4	5	6
Probability	$^1/_{12}$	$^1/_6$	$^1/_6$		$^1/_4$	$^1/_{24}$

Q64 Find the expected value of the score when rolling two 4-sided dice.

Q65 A couple decide to stop having children when they have a boy and a girl or when they have 5 children. What is the expected value for the number of children they will have?

Q66 You pay \$1 to play a game. Then you toss two bent coins, each with a 40% chance of coming down heads. If you get two heads, you win \$4; if you get two tails, you win \$1. What is the expected value of your winnings. If you played a lot of times, would you win or lose in the long run?

Q67 In a game you roll two normal dice. If you get a double, you win \$5; if you get a 6, you win \$4; if you get a double 6, you get \$9. How much should it cost to play if it is to be fair?

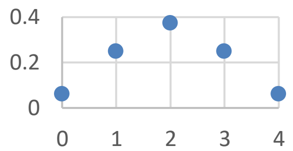
Q68 Calculate by hand the variance and standard deviation of the score when rolling a die. Check by using your calculator.

Answers

Q1 (a) not a random variable
(b) discrete random variable
(c) continuous random variable
(d) not a random variable
(e) discrete random variable
(f) discrete random variable
(g) continuous random variable
(h) discrete random variable
(i) discrete random variable
(j) continuous random variable
(k) discrete random variable

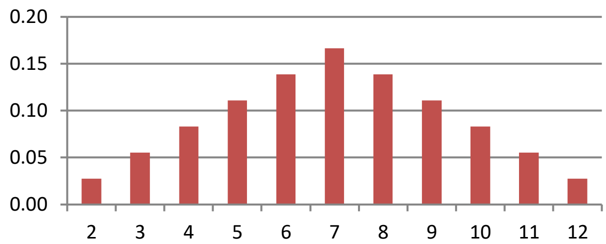
Q2

Number	0	1	2	3	4
Probability	$^1/_{16}$	$^4/_{16}$	$^6/_{16}$	$^4/_{16}$	$^1/_{16}$



Q3

Score	2	3	4	5	6	7	8	9	10	11	12
Probability	$^1/_{36}$	$^2/_{36}$	$^3/_{36}$	$^4/_{36}$	$^5/_{36}$	$^6/_{36}$	$^5/_{36}$	$^4/_{36}$	$^3/_{36}$	$^2/_{36}$	$^1/_{36}$



Q4 14%

Q5

Number	1	2	3	4	5	6	7	8	9	10
Probability	$^1/_2$	$^1/_4$	$^1/_8$	$^1/_{16}$	$^1/_{32}$	$^1/_{64}$	$^1/_{128}$	$^1/_{256}$	$^1/_{512}$	$^1/_{1024}$

Assumptions: There is a ½ probability that any child will be a girl
There are no multiple births
They give up after having 10 boys

Q6 7 Q7 2 Q8 3 Q9 $2^3/_5$

Q10 83.33c. You will lose in the long run. It is not worth playing.

Q11 $1.28. Definitely worth playing.

Q12 2¾

Q13 $38.50

Q14 $2.93. Consider there to be just 5 possible outcomes from a roll: 2, 3, 4, 5 or 6.

Q15 $^5/_6$. No.

Q16 $4.11. Worth playing.

Q17 1.25 times as much. You should take the deal.

Q18 The same. It makes no difference whether you take the deal or not.

Q19 3, 1.5, 1.22

Q20 3.6, 1.44, 1.2

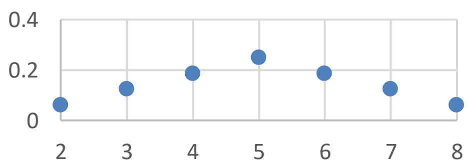
Q51 $2. Not worth playing.

Q52 It makes no difference if she swaps. The false assumption is that any amount of money is as likely as any other amount. So, if the check is for $2 million dollars, then the other envelope is as likely to contain $4 million as $1 million. This would only be true if all amounts were equally likely in an envelope. For this to be the case $200 trillion dollars would have to be as likely as $200. Obviously, it wouldn't be: smaller amounts will have to be more likely than larger amounts.

Q61 (a) discrete random variable
(b) not a random variable
(c) continuous random variable

Q62

Score	2	3	4	5	6	7	8
Probability	$^1/_{16}$	$^2/_{16}$	$^3/_{16}$	$^4/_{16}$	$^3/_{16}$	$^2/_{16}$	$^1/_{16}$



Q63 $^7/_{24}$ Q64 5 Q65 $2^7/_8$

Q66 $1. Neither. Q67 $2.06 Q68 $^{35}/_{12}$, $\sqrt{^{35}/_{12}}$