Probability Distributions

Lesson objectives

- I can recognise and identify a discrete random variable
- I can generate a probability distribution by calculating the probabilities for all values of a random variable
- I can represent a proability distribution using a table and a probability histogram
- I can make connections between the frequency histogram and the probability distribution
- I can calculate and interpret the expected value for a probability distribution
- I can make connections between the expected value and the weighted mean of the values of the discrete random variables

1.1

Lesson objectives | Teachers' notes | Lesson notes

MHR Page 151 #s 1 - 7 & 9

Definitions

Probability Distribution

- The probabilities for all possible outcomes of an experiment or sample space
- Often shown as a graph of probability versus the value of a random variable

Random Variable

- A quantity that can have a range of values
- Designated by a capital letter X, with individual values designated by a lower-case x

Discrete Random Variable

- A variable that can have only certain values within a given range, such as the sum of two dice

Continuous Random Variable

- A variable that can have an infinite number of possible values in a given range, often measurements, such as volume or time

Probability Histogram

- A graph of a probability distribution in which equal intervals are marked on the horizontal axis and the probabilities associated with these intervals are indicated by the areas of the bars

Weighted Mean

- The mean of a set of numbers that are given weightings based on their frequency
- Multiply each number by its weight (or frequency) and divide by the sum of the weights

Expectation (Expected Value)

- Written E(X)
- E(X) of a probability distribution is the predicted average of all possible outcomes
- E(X) is equal to the sum of the products of each outcome, x, with its probability, P(x)
- $$E(X) = \sum_{i=1}^{n} x_i \cdot P(x_i)$$

Let's take a look at the Investigate on Page 144.
We are going to use a spreadsheet to help us.

Example 1

Constructing a Probability Histogram

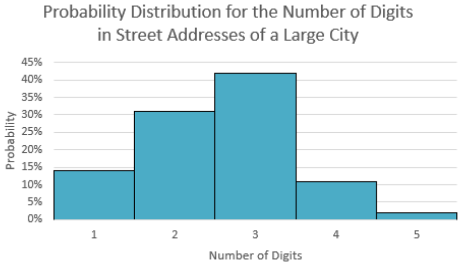
The table gives the probability distribution of the number of digits in street addresses of a large city.

a) Identify the random variable.
b) Construct a probability histogram.
c) Explain the meaning of the individual bars in the histogram.
d) Describe the distribution.
e) Calculate the sum of the probabilities. Comment on the result.

Number of Digits, *x*	Probability, *P*(*x*)
1	14%
2	31%
3	42%
4	11%
5	2%

a) The random variable is the number of digits used in a street address.

b)



c) The area of each bar represents the probability of each outcome.

d) The probability increases to 3 digit numbers and then decreases

e) The sum of the probabilities is 100%. It means that all of the outcomes are accounted for.

Your Turn

The table gives the percent breakdown of the number of rooms in apartments in a particular complex.

Number of Rooms, x	Percent, $P(x)$
2	15
3	30
4	42
5	10
6	3

a) Identify the random variable.

b) Construct a probability histogram.

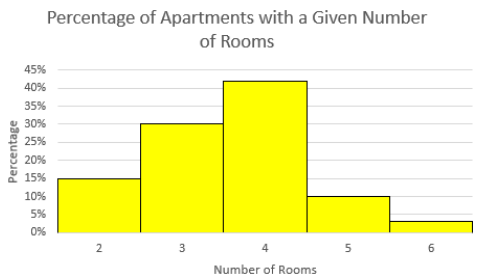
c) Explain the meaning of the individual bars in the histogram.

d) Describe the distribution.

e) Calculate the sum of the probabilities. Does this confirm the results in the example?

a) The random variable is the number of rooms

b)



c) The area of each bar represents the probability of each outcome

d) The number of rooms increases to three, and then decreases

e) The probabilities add to 100%, meaning that all the outcomes are accounted for

Example 2

Expectation of a Probability Distribution

a) Make a tree diagram and show the probability distribution for the number of girls in a family of three children.

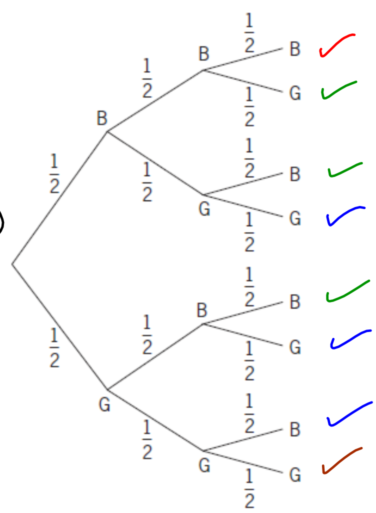
b) Make a probability histogram for this distribution.

c) Calculate the weighted mean number of girls in a "typical" family of three children.

d) Calculate the expectation for the number of girls in a family of three children. Compare it to the weighted mean.

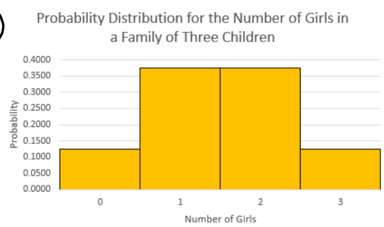
e) Interpret the results in parts c) and d).

a)



# of Girls, x	Freq	Prob, $P(x)$
0	1	1/8
1	3	3/8
2	3	3/8
3	1	1/8

b)



multiply

c) Weighted mean $= \dfrac{\sum_{x=0}^{3} x \cdot f}{\sum f}$

$= \dfrac{(0\times1)+(1\times3)+(2\times3)+(3\times1)}{8}$

$= \dfrac{12}{8} = 1.5$

d) Expected value $= \sum_{x=0}^{3} x \cdot P(x)$

$= \dfrac{12}{8} = 1.5$

They both give an answer of 1.5

e) A family of 3 children would have 1.5 girls

Your Turn

A spinner has two equal sectors, coloured red and blue.

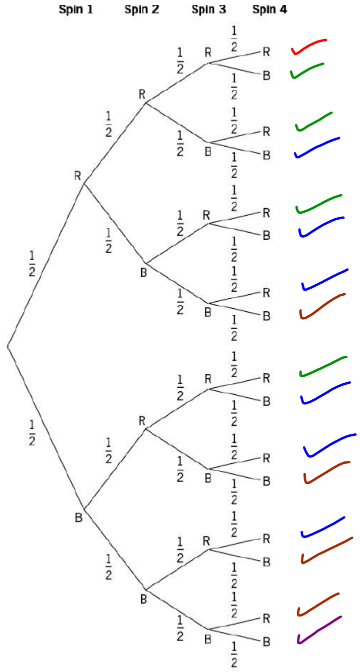
a) Make a tree diagram to show the probability distribution for the number of times the spinner lands on blue when it is spun four times.

b) Make a probability histogram for this distribution.

c) Calculate the expected number of times the spinner lands on blue.

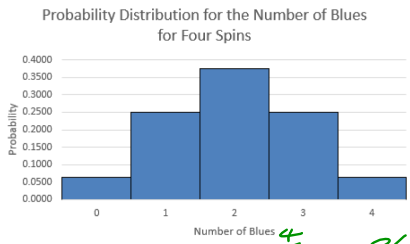
d) Interpret the results in part c).

a)



# of Spins, x	Freq	Prob, P(x)
0	1	1/16
1	4	4/16 = 1/4
2	6	6/16 = 3/8
3	4	4/16 = 1/4
4	1	1/16

b)



c) Expected number of blues $= \sum_{x=0}^{4} x \cdot P(x)$

$= (0 \times \frac{1}{16}) + (1 \times \frac{1}{4}) + (2 \times \frac{3}{8}) + (3 \times \frac{1}{4}) + (4 \times \frac{1}{16})$

$= 0 + \frac{1}{4} + \frac{3}{4} + \frac{3}{4} + \frac{1}{4}$

$= \frac{8}{4} = 2$

d) We would expect to get 2 blues when the spinner is spun four times.

Example 3

Expected Value

A hospital is having a fundraising lottery to raise money for cancer research. A ticket costs $10, and 2 000 000 tickets are available. There are four levels of prizes: one $5 000 000 grand prize, three $100 000 second prizes, ten $1000 prizes, and 2000 free tickets for next year's lottery.

a) What is the expected value of each ticket?

b) Explain its meaning.

a) $E(x) = \frac{\text{Total prizes}}{\text{\# of tickets}} - \text{Cost of ticket}$

$= \frac{1(5,000,000) + 3(100,000) + 10(1,000) + 2,000(10)}{2,000,000} - 10$

$= \frac{5,330,000}{2,000,000} - 10$

$= 2.665 - 10$

$= -7.335$

The expected value of each ticket is -$7.335

b) This means that each ticket, on average, is worth a loss of $7.335

Your Turn

A lottery has a \$10 000 000 grand prize, a \$500 000 second prize, and ten \$50 000 third prizes. A ticket costs \$5, and 4 000 000 tickets were sold.

a) What is the expected value of each ticket?

b) Using the results of this question and of Example 3, are lottery tickets a good investment?

c) How could the lottery be adjusted to make buying a ticket more attractive?

a) $\text{Expected value} = \dfrac{\text{Total prizes}}{\#\text{ of tickets}} - \text{Cost of ticket}$

$= \dfrac{1(10{,}000{,}000) + 1(500{,}000) + 10(50{,}000)}{4{,}000{,}000} - 5$

$= \dfrac{11{,}000{,}000}{4{,}000{,}000} - 5$

$= 2.75 - 5$

$= -2.25$

The expected value of a ticket is -\$2.25

b) This means, that on average, each ticket is worth a loss of \$2.25

Key Concepts

- A probability distribution shows the probabilities of all possible outcomes in an experiment.
- The sum of all probabilities in any distribution is 1.
- A probability histogram graphs the relative frequency of the random variable. The area of each bar represents the probability of the variable.
- Expectation, or expected value, is the weighted average value of the random variable.

$$E(X) = x_1 \cdot P(x_1) + x_2 \cdot P(x_2) + \cdots + x_n \cdot P(x_n)$$
$$= \sum_{i=1}^{n} x_i \cdot P(x_i)$$

The expectation can be a non-integer value.

R1. The expected number of children in a Canadian family is 1.8. Should this be rounded to 2 or left as is? Explain.

Whilst 1.8 children is an impossible number of children to have, an expected value is a predicted value so it should **NOT** be rounded.

R2. Give two examples of a discrete probability distribution. Explain what makes them discrete.

Two examples of a discrete probability distribution are the number of text messages you receive in a day and the number of students in this class. They are examples of discrete data because they have to be whole numbers.

R3. Describe the steps in setting up a probability distribution for the sum of two 12-sided dice.

Create a table showing all the possibilities for the sum of two dice. (You could use a tree diagram, but in this instance it will get VERY messy.)

Determine the frequency of each total and then its probability.

Finally, create a histogram to represent the probability distribution.