

DP IB Maths: AI SL



Your notes

4.4 Probability Distributions

Contents

- * 4.4.1 Discrete Probability Distributions
- * 4.4.2 Expected Values



Your notes

4.4.1 Discrete Probability Distributions

Discrete Probability Distributions

What is a discrete random variable?

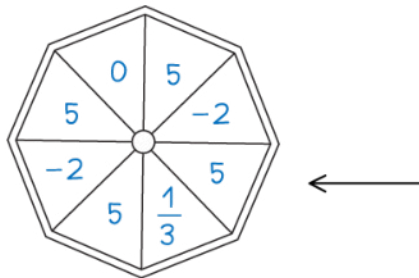
- A **random variable** is a variable whose value depends on the outcome of a **random event**
 - The value of the random variable is not known until the event is carried out (this is what is meant by 'random' in this case)
- **Random variables** are denoted using **upper case letters** (X , Y , etc)
- **Particular outcomes** of the event are denoted using **lower case letters** (x , y , etc)
- $P(X = x)$ means "the probability of the random variable X taking the value x "
- A **discrete** random variable (often abbreviated to DRV) can only take **certain values** within a set
 - Discrete random variables **usually count** something
 - Discrete random variables usually can only take a finite number of values but it is possible that it can take an infinite number of values (see the examples below)
- **Examples** of discrete random variables include:
 - The number of times a coin lands on heads when flipped 20 times
 - this has a finite number of outcomes: $\{0, 1, 2, \dots, 20\}$
 - The number of emails a manager receives within an hour
 - this has an infinite number of outcomes: $\{1, 2, 3, \dots\}$
 - The number of times a dice is rolled until it lands on a 6
 - this has an infinite number of outcomes: $\{1, 2, 3, \dots\}$
 - The number that a dice lands on when rolled once
 - this has a finite number of outcomes: $\{1, 2, 3, 4, 5, 6\}$

What is a probability distribution of a discrete random variable?

- A **discrete probability distribution** fully describes **all the values** that a discrete random variable can take along with their **associated probabilities**
 - This can be given in a **table**
 - Or it can be given as a **function** (called a discrete probability distribution function or "pdf")
 - They can be represented by **vertical line graphs** (the possible values for along the horizontal axis and the probability on the vertical axis)
- The **sum of the probabilities** of **all the values** of a discrete random variable is **1**
 - This is usually written $\sum P(X = x) = 1$
- A **discrete uniform distribution** is one where the random variable takes a finite number of values each with an **equal probability**
 - If there are n values then the probability of each one is $\frac{1}{n}$

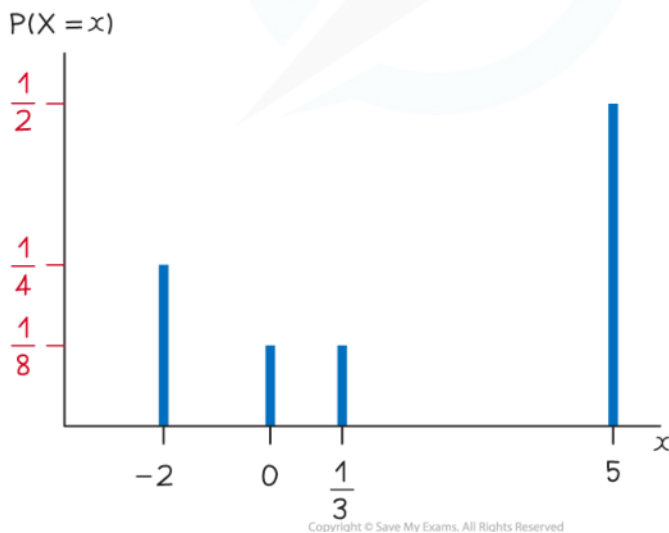


Your notes



LET x BE THE NUMBER THAT THE SPINNER LANDS ON

x	-2	0	$\frac{1}{3}$	5
$P(X=x)$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$

$$P(X=x) = \begin{cases} \frac{1}{8} & x = 0, \frac{1}{3} \\ \frac{1}{4} & x = -2 \\ \frac{1}{2} & x = 5 \\ 0 & \text{OTHERWISE} \end{cases}$$


How do I calculate probabilities using a discrete probability distribution?

- First **draw a table** to represent the probability distribution
 - If it is given as a function then find each probability
 - If any probabilities are unknown then use algebra to represent them
- **Form an equation** using $\sum P(X=x) = 1$
 - Add together all the probabilities and make the sum equal to 1
- To find $P(X=k)$

- If k is a possible value of the random variable X then $P(X = k)$ will be given in the table
- If k is not a possible value then $P(X = k) = 0$
- To find $P(X \leq k)$
 - Identify all possible values, x_i , that X can take which satisfy $x_i \leq k$
 - Add together all their corresponding probabilities
 - $P(X \leq k) = \sum_{x_i \leq k} P(X = x_i)$
 - Some mathematicians use the notation $F(x)$ to represent the cumulative distribution
 - $F(x) = P(X \leq x)$
- Using a similar method you can find $P(X < k)$, $P(X > k)$ and $P(X \geq k)$
- As all the probabilities add up to 1 you can form the following equivalent equations:
 - $P(X < k) + P(X = k) + P(X > k) = 1$
 - $P(X > k) = 1 - P(X \leq k)$
 - $P(X \geq k) = 1 - P(X < k)$

How do I know which inequality to use?

- $P(X \leq k)$ would be used for phrases such as:
 - At most, no greater than, etc
- $P(X < k)$ would be used for phrases such as:
 - Fewer than
- $P(X \geq k)$ would be used for phrases such as:
 - At least, no fewer than, etc
- $P(X > k)$ would be used for phrases such as:
 - Greater than, etc



Your notes



Your notes

Worked example

The probability distribution of the discrete random variable X is given by the function

$$P(X=x) = \begin{cases} kx^2 & x = -3, -1, 2, 4 \\ 0 & \text{otherwise.} \end{cases}$$

a) Show that $k = \frac{1}{30}$.

Construct a table

x	-3	-1	2	4
$P(X=x)$	$9k$	k	$4k$	$16k$

Substitute in the values of x
e.g. $P(X=-3) = k(-3)^2 = 9k$

The probabilities add up to 1

$$9k + k + 4k + 16k = 1$$

$$30k = 1$$

$$k = \frac{1}{30}$$

b) Calculate $P(X \leq 3)$.

Substitute k into the probabilities

x	-3	-1	2	4
$P(X=x)$	$\frac{3}{10}$	$\frac{1}{30}$	$\frac{2}{15}$	$\frac{8}{15}$

$$X \leq 3 : X = -3, -1, 2$$

$$P(X \leq 3) = P(X = -3) + P(X = -1) + P(X = 2)$$

$$= \frac{3}{10} + \frac{1}{30} + \frac{2}{15}$$

$$P(X \leq 3) = \frac{7}{15}$$



Your notes

4.4.2 Expected Values

Expected Values $E(X)$

What does $E(X)$ mean and how do I calculate $E(X)$?

- $E(X)$ means the **expected value** or the **mean** of a **random variable X**
 - The expected value does not need to be an obtainable value of X
 - For example: the expected value number of times a coin will land on tails when flipped 5 times is 2.5
- For a **discrete** random variable, it is calculated by:
 - **Multiplying each value** of X with its corresponding **probability**
 - **Adding** all these terms together

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

- This is given in the **formula booklet**
- Look out for **symmetrical** distributions (where the values of X are symmetrical and their probabilities are symmetrical) as the mean of these is the same as the median
 - For example: if X can take the values 1, 5, 9 with probabilities 0.3, 0.4, 0.3 respectively then by symmetry the mean would be 5

How can I decide if a game is fair?

- Let X be the random variable that represents the **gain/loss** of a player in a game
 - X will be **negative** if there is a **loss**
- Normally the expected gain or loss is calculated by **subtracting** the **cost to play** the game from the **expected value** of the **prize**
- If $E(X)$ is **positive** then it means the player can **expect to make a gain**
- If $E(X)$ is **negative** then it means the player can **expect to make a loss**
- The game is called **fair** if the **expected gain is 0**
 - $E(X) = 0$



Your notes

Worked example

Daphne pays \$15 to play a game where she wins a prize of \$1, \$5, \$10 or \$100. The random variable W represents the amount she wins and has the probability distribution shown in the following table:

w	1	5	10	100
$P(W = w)$	0.35	0.5	0.05	0.1

- a) Calculate the expected value of Daphne's prize.

Formula booklet

Expected value of a discrete random variable X	$E(X) = \sum x P(X = x)$
--	--------------------------

$$E(W) = \sum w P(W = w)$$

$$= 1 \times 0.35 + 5 \times 0.5 + 10 \times 0.05 + 100 \times 0.1$$

$$\text{Expected value} = \$13.35$$

- b) Determine whether the game is fair.

A game is fair if expected gain/loss is 0

Prize - cost

$$13.35 - 15 = -1.65$$

Expected loss is \$1.65 so game is not fair