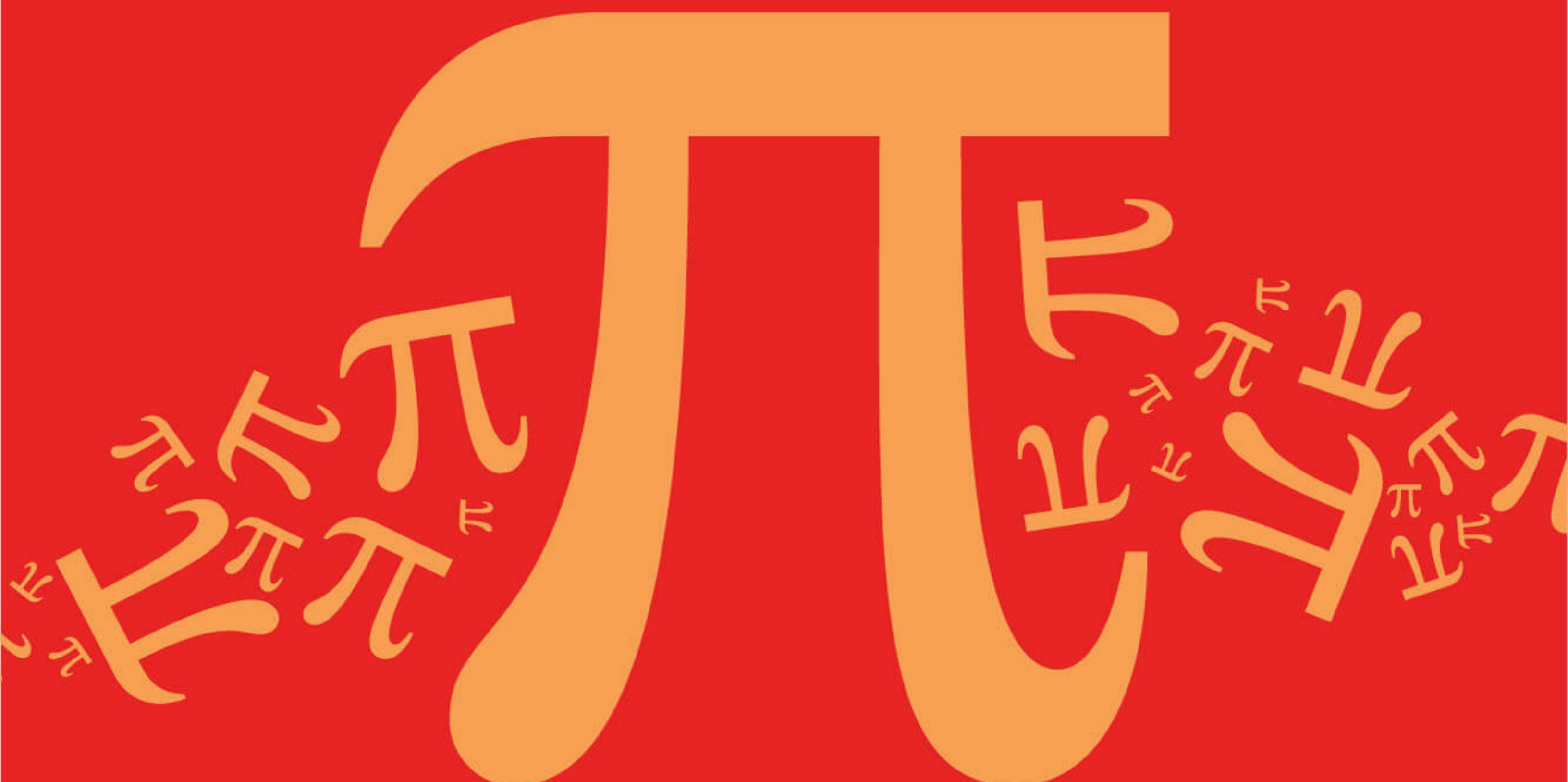
EXAM PAPERS PRACTICE

4.10 Poisson Distribution



IB Maths - Revision Notes

4.10.1 Poisson Distribution

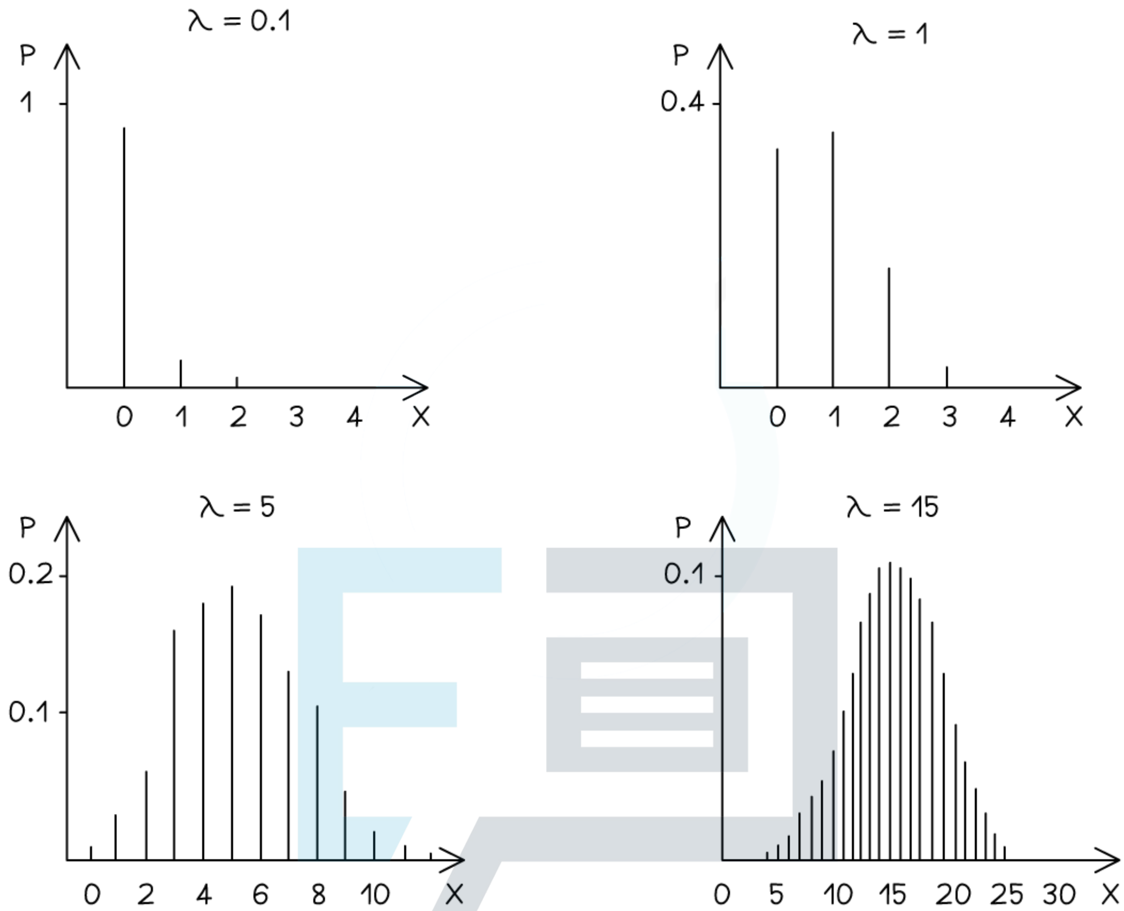
Properties of Poisson Distribution

What is a Poisson distribution?

- A Poisson distribution is a **discrete probability distribution**
- A **discrete random variable** X follows a **Poisson distribution** if it **counts the number of occurrences** in a fixed time period given the following conditions:
  - Occurrences are **independent**
  - Occurrences occur at a **uniform average rate** for the time period **(*m*)**
- If X follows a Poisson distribution then it is denoted $X \sim \mathrm{Po}(m)$
  - *m* is the average rate of occurrences for the time period
- The formula for the probability of *r* occurrences is given by:
  - $\mathrm{P}(X = r) = \dfrac{\mathrm{e}^{-m} m^r}{r!}$ for $r = 0, 1, 2, \ldots$
    - e is Euler's constant 2.718...
    - $r! = r \times (r - 1) \times \ldots \times 2 \times 1$ and $0! = 1$
    - There is no upper bound for the number of occurrences
  - You will be expected to use the distribution function on your GDC to calculate probabilities with the Poisson distribution

What are the important properties of a Poisson distribution?

- The **expected number (mean)** of occurrences is *m*
  - You are given this in the **formula booklet**
- The **variance** of the number of occurrences is *m*
  - You are given this in the **formula booklet**
  - Square root to get the **standard deviation**
- The **mean** and **variance** for a Poisson distribution are **equal**
- The distribution can be represented visually using a vertical line graph
  - The graphs have **tails to the right** for all values of *m*
  - As ***m* gets larger** the graph gets **more symmetrical**
- If $X \sim \mathrm{Po}(m)$ and $Y \sim \mathrm{Po}(\lambda)$ are **independent** then $X + Y \sim \mathrm{Po}(m + \lambda)$
  - This extends to *n* independent Poisson distributions $X_i \sim \mathrm{Po}(m_i)$
    - $X_1 + X_2 + \ldots + X_n \sim \mathrm{Po}(m_1 + m_2 + \ldots + m_n)$


λ = 0.1
P
1
0
1
2
3
4
X
λ = 1
P
0.4
0
1
2
3
4
X
λ = 5
P
0.2
0.1
0
2
4
6
8
10
X
λ = 15
P
0.1
0
5
10
15
20
25
30
X

Modelling with Poisson Distribution

How do I set up a Poisson model?

- **Identify** what an **occurrence** is in the scenario
  - For example: a car passing a camera, a machine producing a faulty item
- Use **proportion** to find the **mean number of occurrences** for the given time period
  - For example: 10 cars in 5 minutes would be 120 cars in an hour if the Poisson model works for both time periods
- Make sure you **clearly state** what your **random variable** is
  - For example: let X be the number of cars passing a camera in 10 minutes

What can be modelled using a Poisson distribution?

- Anything that satisfies the **two conditions**
- For example, Let C be the number of calls that a helpline receives within a 15-minute period: $C \sim \text{Po}(m)$
  - An occurrence is the helpline receiving a call and can be considered independent
  - The helpline receives calls at an average rate of m calls during a 15-minute period
- Sometimes a **measure of space** will be used instead of a time period
  - For example, the number of daisies that exist on a square metre of grass
- If the **mean** and **variance** of a discrete variable are **equal** then it might be possible to use a Poisson model

Exam Tip

- An exam question might involve different types of distributions so make it clear which distribution is being used for each variable

Worked example

Jack uses $\text{Po}(6.25)$ to model the number of emails he receives during his hour lunch break.

a) Write down two assumptions that Jack has made.

Jack has assumed that:

- the emails that he receives are independent
- he receives emails at a uniform average rate of 6.25 emails per hour during his lunch breaks

b) Calculate the standard deviation for the number of emails that Jack receives during his hour lunch breaks.

Formula booklet

Poisson distribution $X \sim \text{Po}(m)$	
Mean	$\text{E}(X) = m$
Variance	$\text{Var}(X) = m$

$\sigma^2 = 6.25$

$\sigma = \sqrt{6.25}$

Standard deviation = 2.5 emails

4.10.2 Calculating Poisson Probabilities

Calculating Poisson Probabilities

Throughout this section we will use the random variable $X \sim \mathrm{Po}(m)$. For a Poisson distribution *X*, the probability of *X* taking a non-integer or negative value is always zero. Therefore, any values mentioned in this section for *X* will be assumed to be non-negative integers. The value of *m* can be any real positive value.

How do I calculate P(*X* = *x*): the probability of a single value for a Poisson distribution?

- You should have a **GDC** that can calculate **Poisson probabilities**
- You want to use the "**Poisson Probability Distribution**" function
  - This is sometimes shortened to PPD, Poisson PD or Poisson Pdf
- You will need to enter:
  - The *'x'* value - the value of *x* for which you want to find $\mathrm{P}(X = x)$
  - The 'λ' value - the **mean number of occurrences (*m*)**
- Some calculators will give you the option of **listing the probabilities** for **multiple values of *x* at once**
- There is a formula that you can use but you are expected to be able to use the distribution function on your GDC
  - $\mathrm{P}(X = x) = \dfrac{\mathrm{e}^{-m} m^x}{x!}$
    - where e is Euler's constant
    - $x! = x \times (x-1) \times \ldots \times 2 \times 1$ and $0! = 1$

How do I calculate P(*a* ≤ *X* ≤ *b*): the cumulative probabilities for a Poisson distribution?

- You should have a **GDC** that can calculate **cumulative Poisson probabilities**
  - Most calculators will find $\mathrm{P}(a \le X \le b)$
  - Some calculators can only find $\mathrm{P}(X \le b)$
    - The identities below will help in this case
- You should use the "**Poisson Cumulative Distribution**" function
  - This is sometimes shortened to PCD, Poisson CD or Poisson Cdf
- You will need to enter:
  - The lower value - this is the **value *a***
    - This can be zero in the case $\mathrm{P}(X \le b)$
  - The upper value - this is the **value *b***

- This can be a very large number (9999...) in the case $P(X \geq a)$
- The 'λ' value - the **mean number of occurrences (*m*)**

How do I find probabilities if my GDC only calculates P(*X*≤ *x*)?

- To calculate P(*X*≤ *x*) just enter *x* into the cumulative distribution function

- To calculate P(*X* < *x*) use:
  - $P(X < x) = P(X \leq x - 1)$ which works when *X* is a Poisson random variable
    - P(*X* < 5) = P(*X*≤ 4)

- To calculate P(*X*> *x*) use:
  - $P(X > x) = 1 - P(X \leq x)$ which works for any random variable *X*
    - P(*X*> 5) = 1 - P(*X*≤ 5)

- To calculate P(*X*≥ *x*) use:
  - $P(X \geq x) = 1 - P(X \leq x - 1)$ which works when *X* is a Poisson random variable
    - P(*X*≥ 5) = 1 - P(*X*≤ 4)

- To calculate P(*a*≤ *X*≤ *b*) use:
  - $P(a \leq X \leq b) = P(X \leq b) - P(X \leq a - 1)$ which works when *X* is a Poisson random variable
    - P(5 ≤ *X*≤ 9) = P(*X*≤ 9) - P(*X*≤ 4)

What if an inequality does not have the equals sign (strict inequality)?

- For a Poisson distribution (as it is discrete) you could **rewrite all strict inequalities** (< and >) as **weak inequalities** (≤ and ≥) by using the identities for a Poisson distribution
  - $P(X < x) = P(X \leq x - 1)$ and $P(X > x) = P(X \geq x + 1)$
  - For example: P(*X* < 5) = P(*X*≤ 4) and P(*X*> 5) = P(*X*≥ 6)
- It helps to think about the **range of integers** you want
  - Identify the smallest and biggest integers in the range
- If your range has no minimum then use 0
  - $P(X \leq b) = P(0 \leq X \leq b)$

- $P(a < X \leq b) = P(a + 1 \leq X \leq b)$
  - P(5 < *X*≤ 9) = P(6 ≤ *X*≤ 9)

- $P(a \leq X < b) = P(a \leq X \leq b - 1)$
  - P(5 ≤ *X* < 9) = P(5 ≤ *X*≤ 8)

- $P(a < X < b) = P(a + 1 \leq X \leq b - 1)$
  - P(5 < *X* < 9) = P(6 ≤ *X*≤ 8)

Worked example

The random variables $X \sim \text{Po}(6.25)$ and $Y \sim \text{Po}(4)$ are independent. Find:

i) $\text{P}(X = 5)$,

Use Poisson probability distribution on GDC

$m = 6.25$ (λ) $x = 5$

$P(X = 5) = 0.15341...$

$P(X = 5) = 0.153$ (3sf)

ii) $\text{P}(Y \le 5)$,

Identify upper and lower bounds

$P(Y \le 5) = P(0 \le Y \le 5)$

Use Poisson cumulative distribution on GDC $m = 4$ (λ)

$P(Y \le 5) = 0.78513...$

$P(Y \le 5) = 0.785$ (3sf)

iii) $\text{P}(X + Y > 7)$.

Form the distribution $m = 6.25 + 4 = 10.25$

$X + Y \sim Po(10.25)$

Identify lower bound - no upper bound so use a large number 999...

$P(X + Y > 7) = P(8 \le X + Y)$

Use Poisson cumulative distribution on GDC

$P(8 \le X + Y) = 0.80146...$

$P(X + Y > 7) = 0.801$ (3sf)