Chapter 2: Poisson Distributions

Pearson Edexcel International A Level Statistics 2 Student Book

Chapter 2 of the "Pearson Edexcel International A Level Statistics 2 Student Book" is dedicated to **Poisson Distributions** [1, 2]. This chapter covers various aspects of this probability distribution, including its definition, conditions for its use, how to model real-world situations, its additive property, and its mean and variance [2, 3].

Here are the main ideas discussed in Chapter 2, presented with solved examples from the sources:

1. The Poisson Distribution Definition and Basic Probability Calculation

- **Idea**: This section introduces the Poisson distribution, denoted as $X \sim Po(\lambda)$, where λ is the average rate of occurrence [4, 5]. It explains that a Poisson distribution is suitable when events occur:
  - **Singly** in space or time [4].
  - **Independently** of each other [4].
  - At a **constant average rate** within a given interval [4].
- **Formula**: The probability of x occurrences is given by $P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$ [5].
- **Solved Example (from Example 1, page 19)** [5]: Let a random variable X follow a Poisson distribution $X \sim Po(2.1)$.

  a $P(X = 3)$: Using the formula $P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$ with $x = 3$ and $\lambda = 2.1$:

  $$P(X = 3) = \frac{e^{-2.1}(2.1)^3}{3!} = \frac{0.12245 \times 9.261}{6} \approx \mathbf{0.1890} \quad \text{(4 d.p.)[5].}$$

  b $P(X > 1)$: This is calculated as $P(X > 1) = 1 - P(X \leq 1)$ [5]. From a calculator or Poisson cumulative distribution table, $P(X \leq 1) \approx 0.3796$. Therefore,

  $$P(X > 1) = 1 - 0.3796 = \mathbf{0.6204} \quad \text{(4 d.p.)[5].}$$

  c $P(1 \leq X < 4)$: This refers to the sum of probabilities for $X = 1, X = 2$, and $X = 3$ [5].

  $$\begin{aligned} P(X = 1) &= \frac{e^{-2.1}(2.1)^1}{1!} \approx 0.2571 \\ P(X = 2) &= \frac{e^{-2.1}(2.1)^2}{2!} \approx 0.2700 \\ P(X = 3) &= \frac{e^{-2.1}(2.1)^3}{3!} \approx 0.1890 \\ P(1 \leq X < 4) &= P(X = 1) + P(X = 2) + P(X = 3) \\ &= 0.2571 + 0.2700 + 0.1890 = \mathbf{0.7161} \quad \text{(4 d.p.)[5].} \end{aligned}$$

2. Modelling with the Poisson Distribution (Real-World Applications)

- **Idea**: This section focuses on identifying real-world scenarios where the Poisson distribution can be used to model the number of times a particular event occurs within a fixed interval of time or space [6]. It emphasizes checking the underlying assumptions of singleness, independence, and constant average rate [6].

- **Solved Example (from Example 4, page 22)** [7]: An internet service provider observes that, on average, 4 users every hour fail to connect on their first attempt.

  a Give two reasons why a Poisson distribution might be a suitable model for the number of failed connections every hour.

  - Failed connections occur **singly** (one at a time) and randomly in a given timeframe [7].
  - The average rate of failed connections is **constant** (4 per hour) [7].
  - The occurrences of failed connections are **independent** [7].

  b Find the probability that in a randomly chosen hour: Let X be the number of failed connections in one hour. Since the average is 4 per hour, $X \sim Po(4)$ [7].

  i 2 users fail to connect on their first attempt ($P(X = 2)$):

$$\begin{aligned} P(X = 2) &= \frac{e^{-4}4^2}{2!} \\ &= \frac{0.018316 \times 16}{2} \approx \mathbf{0.1465} \quad \text{(4 d.p.)[7].} \end{aligned}$$

  ii more than 6 users fail to connect on their first attempt ($P(X > 6)$):

$$\begin{aligned} P(X > 6) &= 1 - P(X \leq 6) \\ &= 1 - 0.8893 = \mathbf{0.1107} \quad \text{(4 d.p.)[7].} \end{aligned}$$

  c Find the probability that in a randomly chosen 90-minute period: A 90-minute period is 1.5 hours. The new average rate (λ) for this period is $4 \times 1.5 = 6$. Let Y be the number of failed connections in 90 minutes. $Y \sim Po(6)$ [7].

  i 5 users fail to connect on their first attempt ($P(Y = 5)$):

$$P(Y = 5) = \frac{e^{-6}6^5}{5!} \approx \mathbf{0.1606} \quad \text{(4 d.p.)[7].}$$

  ii fewer than 7 users fail to connect on their first attempt ($P(Y < 7)$): This means $P(Y \leq 6)$.

$$P(Y < 7) = P(Y \leq 6) \approx \mathbf{0.6063} \quad \text{(4 d.p.)[7].}$$

3. Adding Poisson Distributions (Additive Property)

- **Idea**: This property states that if X and Y are two **independent** random variables following Poisson distributions with parameters λ and μ respectively (i.e., $X \sim$

$Po(\lambda)$ and $Y \sim Po(\mu)$), then their sum $X + Y$ also follows a Poisson distribution with parameter $\lambda + \mu$ (i.e., $X + Y \sim Po(\lambda + \mu)$) [8]. This is applicable when both distributions model events occurring in the same interval of time or space [8].

- **Solved Example (from Example 6, page 26)** [8]: Given $X \sim Po(3.6)$ and $Y \sim Po(4.4)$.

  a Find $P(X + Y = 7)$: Since X and Y are independent Poisson variables, their sum $X + Y$ follows a Poisson distribution with parameter $\lambda = 3.6 + 4.4 = 8$. So, $X + Y \sim Po(8)$ [8]. Let $Z = X + Y$. We need $P(Z = 7)$.

  $$P(Z = 7) = \frac{e^{-8}8^7}{7!} \approx \mathbf{0.1396} \quad \text{(4 d.p.)[8].}$$

  b Find $P(X + Y \leq 5)$: Using a cumulative distribution table or calculator for $Z \sim Po(8)$,

  $$P(Z \leq 5) \approx \mathbf{0.1912} \quad \text{(4 d.p.)[8].}$$

4. Mean and Variance of a Poisson Distribution

- **Idea**: A key characteristic of the Poisson distribution is that its mean (expected value) is equal to its variance [9]. For a random variable $X \sim Po(\lambda)$:
  - Mean $E(X) = \lambda$ [9].
  - Variance $Var(X) = \lambda$ [9].
  - The standard deviation is $\sqrt{\lambda}$.

- **Solved Example (from Example 8, page 29)** [9, 10]: A botanist counts the number of daisies, x, in each of 80 randomly selected squares. Summarized results: $\Sigma x = 295$, $\Sigma x^2 = 1386$.

  a Calculate the mean and the variance of the number of daisies per square. Give your answers to 2 decimal places.
  - **Mean ($E(X)$)**:

  $$E(X) = \frac{\Sigma x}{n} = \frac{295}{80} = 3.6875 \approx \mathbf{3.69} \quad \text{(2 d.p.)[10].}$$

  - **Variance ($Var(X)$)**:

  $$\begin{aligned} Var(X) &= \frac{\Sigma x^2}{n} - (\text{mean})^2 \\ &= \frac{1386}{80} - (3.6875)^2 \\ &= 17.325 - 13.59140625 \approx 3.73359375 \approx \mathbf{3.73} \quad \text{(2 d.p.)[10].} \end{aligned}$$

  b Explain how the answers from part a support the choice of a Poisson distribution as a model. The mean (≈ 3.69) and the variance (≈ 3.73) are **approximately equal** [10]. This equality (or near-equality) between the mean and variance is a fundamental property of the Poisson distribution, making it a suitable model [9].

c Using a suitable value for λ, estimate the probability that exactly 3 daisies will be found in a randomly selected square. Using the estimated mean as the parameter for the Poisson distribution, $\lambda = 3.7$ [10]. Let $X \sim Po(3.7)$. We need to find $P(X = 3)$. Using tables or a calculator for $Po(3.7)$,

$$P(X = 3) \approx \mathbf{0.2087} \quad \text{(4 d.p.)[10].}$$

Chapter 2 concludes with a Chapter Review section, offering further practice problems related to these concepts [11].