

The Poisson distribution

A-Level Maths 9709 · Topic 6.1 revision notes

Everything from the 2 lessons of this topic on a few pages: the key idea of each lesson, the diagram to remember, the points to learn and the traps examiners set. Read it, cover it, test yourself.

1. The Poisson distribution

2. Approximations: binomial to Poisson,
Poisson to normal

thelucidstem.com

1 The Poisson distribution

Syllabus 6.1

If events occur singly, independently and at random at a constant mean rate, the number X in a fixed interval satisfies $X \sim \text{Po}(\lambda)$ with $P(X = r) = e^{-\lambda} \frac{\lambda^r}{r!}$ for $r = 0, 1, 2, \dots$. The mean and the variance are both λ , and λ scales in proportion to the length of the interval.

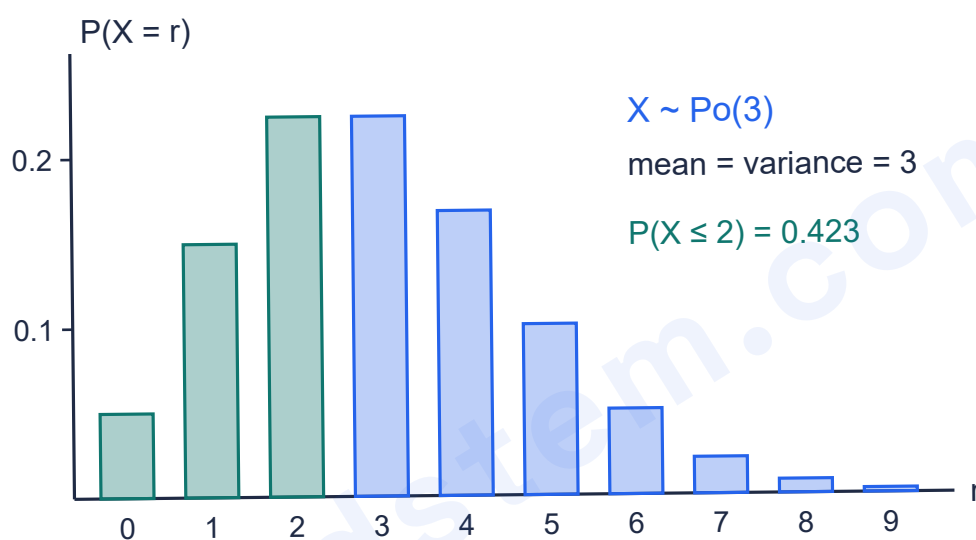


Fig. 1: the Poisson distribution with mean 3, drawn to scale. The teal bars add to $P(X \leq 2) = 0.423$; the bars for $r = 2$ and $r = 3$ are equal because $P(X = 3) = P(X = 2) \times 3/3$

Worked example: Is a Poisson model suitable?

In 100 one-minute intervals, the numbers of cars passing a point were recorded. The sample mean was 1.96 and the sample variance was 1.82. A second road, next to a set of traffic lights, gave mean 2.0 and variance 0.4. Comment on each.

1. For a Poisson distribution the mean equals the variance.
2. First road: 1.96 and 1.82 are close, so a Poisson model with $\lambda \approx 1.96$ is reasonable.
3. Second road: the variance 0.4 is much smaller than the mean 2.0.
4. Traffic lights release cars in bunches at regular times, so the cars do not arrive independently at random.

Answer: Poisson is suitable for the first road but not for the second (variance far below the mean; arrivals are not independent).

Define X , fix λ for the interval, then calculate.

- ✓ **Define:** "Let X be the number of ... in ..." and write $X \sim \text{Po}(\lambda)$.
- ✓ **Scale:** multiply the rate by the length of the interval in the question.
- ✓ **Boundaries:** write the event as an inequality ($X \geq 3$, $X \leq 2$) before calculating.
- ✓ **Complements:** $P(X \geq k) = 1 - P(X \leq k - 1)$.
- ✓ **Context:** state conditions and conclusions in terms of the situation.
- ✓ **Accuracy:** keep full calculator values; give answers to 3 s.f.

Watch out!

Do not just recite "independent and random". Cambridge wants the condition in context: "the calls arrive independently of each other" or "the mean rate of calls is constant during the day". A condition with no reference to the situation usually earns nothing.

"More than 2" means $X \geq 3$, so it is $1 - P(X \leq 2)$, not $1 - P(X \leq 3)$. Write the inequality in symbols before you calculate; most lost marks here are a boundary slipped by one.

In a two-stage question, do not add the rates to $\lambda = 6$ for the week and find $P(X = 0)$; that is the probability of no faults all week, a different event. Identify what is being counted at each stage.

2 Approximations: binomial to Poisson, Poisson to normal

Syllabus 6.1

If $X \sim B(n, p)$ with $n > 50$ and $np < 5$, use $X \approx \text{Po}(np)$. If $X \sim \text{Po}(\lambda)$ with $\lambda > 15$, use $X \approx N(\lambda, \lambda)$: the variance equals the mean, and a continuity correction of 0.5 is needed because a discrete count is replaced by a continuous variable.

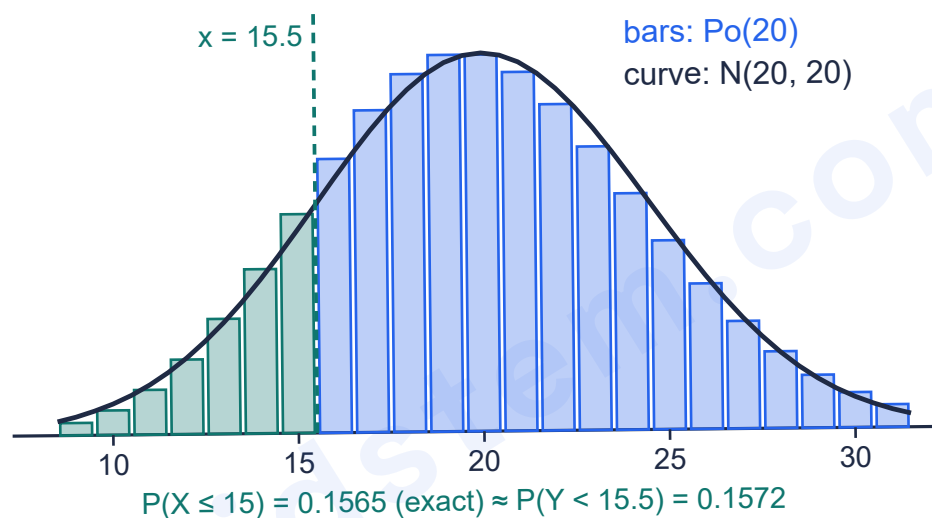


Fig. 1: $\text{Po}(20)$ (bars, drawn to scale) and the normal curve $N(20, 20)$. The teal bars are $X \leq 15$; the approximating area stops at 15.5, half-way through the next bar

Worked example: A rare defect

1.5% of bolts made by a machine are faulty. A random sample of 200 bolts is taken. Use a suitable approximation to find the probability that at most 2 are faulty.

1. $X \sim B(200, 0.015)$. Here $n = 200 > 50$ and $np = 3 < 5$, so use $X \approx \text{Po}(3)$.

2. $P(X \leq 2) \approx e^{-3} \left(1 + 3 + \frac{3^2}{2} \right) = 8.5e^{-3}$.

3. $= 0.4232$. (The exact binomial value is 0.4215, so the approximation is good.)

Answer: 0.423 (3 s.f.)

Check the conditions, choose, then correct.

- ✓ **Binomial, rare successes:** $n > 50$, $np < 5$: use $Po(np)$, no continuity correction.
- ✓ **p close to 1:** count the failures instead.
- ✓ **Poisson, large mean:** $\lambda > 15$: use $N(\lambda, \lambda)$, standard deviation $\sqrt{\lambda}$.
- ✓ **Continuity correction:** list the integers included, then go half a unit beyond the last one.
- ✓ **Justify:** quote the numerical conditions in your answer.

Watch out!

State the conditions with numbers: " $n = 200 > 50$ and $np = 3 < 5$ ". Writing " n is large and p is small" with no values often loses the mark, and using $Po(np(1-p))$ is wrong: the Poisson parameter is the mean, np .

The standard deviation of $N(20, 20)$ is $\sqrt{20}$, not 20. Dividing by 20 is the commonest error in this topic, closely followed by forgetting the continuity correction altogether.

Do not apply a continuity correction when approximating binomial by Poisson. Both are discrete, so $P(X \leq 2)$ stays $P(X \leq 2)$.