STEP Support Programme

STEP 2 Specification Probability Notes

These notes are designed to help students in preparing for STEP 2. They cover the "bold and italic" sections of the STEP 2 specification which are not covered in the A-level single Mathematics specifications, or AS Further Maths Common Core. Many of these topics will be covered in A Level Further Mathematics, and will be covered in some AS Further Mathematics modules.

There are more notes on the various sections of the specification in the STEP 2 modules.

Poisson Distribution

The *Poisson distribution* measures the number of occurrences of an event in a given time interval. It was first used by Ladislaus Josephovich Bortkiewicz to model the number of deaths of Prussian cavalry-men by horse kicks in a year.

A Poisson random variable satisfies the following conditions:

I Occurrences are independent.

II The mean (or expected) number of occurrences during a time interval is proportional to the length of the time interval.

As well as modelling the number of occurrences in a given time interval it can be used to model the number of occurrences in a given space interval. Some applications are the number of car accidents in a mile of road, the number of people joining a queue every 5 minutes and the number of hairs in a burger.

The number of occurrences in a given time interval is given by:

$$\mathrm{P}(X = r) = \frac{\mathrm{e}^{-\lambda}\lambda^r}{r!}$$

where r is an integer, with $r \geqslant 0$, λ is the mean number of occurrences in the given interval and (by convention) $0! = 1$.

Note that the sum of all the probabilities is given by:

$$\sum_{r=0}^{\infty} \frac{\mathrm{e}^{-\lambda}\lambda^r}{r!} = \mathrm{e}^{-\lambda} \times \sum_{n=0}^{\infty} \frac{\lambda^n}{r!} = \mathrm{e}^{-\lambda} \times \mathrm{e}^{\lambda} = 1\,.$$

For the last equality, we used the exponential series expansion $\mathrm{e}^x = 1 + x + \frac{x^2}{2!} + \cdots$. The fact that the probabilities sum to 1 should be reassuring!

The expectation of the Poisson distribution is given by:

$$\begin{aligned}
\mathrm{E}(X) &= \sum_{r=0}^{\infty} \frac{\mathrm{e}^{-\lambda}\lambda^r}{r!} \times r \\
&= 0 + \left(1 \times \mathrm{e}^{-\lambda}\lambda\right) + \left(2 \times \frac{\mathrm{e}^{-\lambda}\lambda^2}{2!}\right) + \left(3 \times \frac{\mathrm{e}^{-\lambda}\lambda^3}{3!}\right) + \left(4 \times \frac{\mathrm{e}^{-\lambda}\lambda^4}{4!}\right) + \cdots \\
&= \lambda\mathrm{e}^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots\right) \\
&= \lambda\mathrm{e}^{-\lambda} \times \mathrm{e}^{\lambda} \\
&= \lambda
\end{aligned}$$

Similarly, it can be shown that $\mathrm{Var}(X) = \lambda$ (you might like to do this). If X has a Poisson distribution then we write $X \sim \mathrm{Po}(\lambda)$.

The proportionality principle means that if the number of successes in a time interval can be modelled as a Poisson distribution with a mean of λ, then the number of successes in a time interval twice as long can be modelled as a Poisson distribution with mean 2λ, and so on.

Example: The number of calls to "Kalculus Kids" is Poisson distributed with, on average, 8 calls an hour. Find the probabilities of:

(a) 5 calls in one hour

here $X \sim \text{Po}(8)$ and so $\text{P}(X = 5) = \dfrac{\text{e}^{-8} \times 8^5}{5!}$

(b) 10 calls in two hours

here $X \sim \text{Po}(16)$ and so $\text{P}(X = 10) = \dfrac{\text{e}^{-16} \times 16^{10}}{10!}$

(c) 2 calls in 30 mins

here $X \sim \text{Po}(4)$ and so $\text{P}(X = 2) = \dfrac{\text{e}^{-4} \times 4^2}{2!} = 8\text{e}^{-4}$

(d) Fewer than 2 calls in 15 mins

here $X \sim \text{Po}(2)$ and so $\text{P}(X < 2) = \text{P}(X = 0 \text{ or } 1) = \dfrac{\text{e}^{-2} \times 2^0}{0!} + \dfrac{\text{e}^{-2} \times 2^1}{1!} = 3\text{e}^{-2}$

Note that I have not written these as decimals, and have only simplified the expressions if it was straightforward and useful to do so.

If you have two *independent*[1] Poisson distributions, X with mean λ and Y with mean μ, then the sum $X + Y$ is also a Poisson distribution with mean $\lambda + \mu$.

Example: A company making gizmos accepts orders on-line or by phone. Both of these follow a Poisson distribution, telephone orders with a mean of 2 per day and on-line orders with a mean of 5 per day. The telephone and on-line orders are independent of each other. What is the probability that they receive 10 orders on one day?

The total number of orders per day is distributed as $X \sim \text{Po}(7)$ and so we have
$\text{P}(X = 10) = \dfrac{\text{e}^{-7} \times 7^{10}}{10!}$.

[1] *Independent* means that the outcome of one of the events has no affect on the outcome of the other. If two events are independent then we have $\text{P}(X = x \text{ and } Y = y) = \text{P}(X = x) \times \text{P}(Y = y)$.

Approximating Distributions

The Normal, Binomial and Poisson distributions model different situations:

- The *Normal* distribution models a continuous situation (such as height), and can take values in the range $(-\infty, \infty)$. The Normal distribution is symmetrical about the mean.
- The *Binomial* distribution is a discrete distribution which can take the values $0, 1, 2, 3, \cdots, n$. It can be thought of as the number of "successes" out of n "trials".
- The *Poisson* distribution is also a discrete distribution, but in this case it can take the values $0, 1, 2, 3, \cdots$ (i.e. there is no upper bound). It can be thought of as the number of "occurrences" in a given interval (which might be a time interval, space interval etc.).

Under certain situations the Binomial distribution can be approximated by a Poisson or Normal distribution, and the Poisson distribution can be approximated by a Normal distribution.

If n "is large" and p "is small" then the Binomial distribution $\mathrm{B}(n, p)$ can be approximated with a Poisson distribution, $\mathrm{Po}(np)$.

"Large" and "small" are not well defined. One rule of thumb is that we need $np \leqslant 10$, and the larger the value of n the better the approximation.

- If $X \sim \mathrm{B}(n, p)$ and n is "large" and/or p is "close to $\frac{1}{2}$" then X can be approximated by a normal distribution, $X \sim \mathrm{N}(np, np(1-p))$.
- If $X \sim \mathrm{Po}(\lambda)$ and λ is "large" then X can be approximated by a normal distribution, $X \sim \mathrm{N}(\lambda, \lambda)$.

If you are approximating a discrete distribution, X, (e.g. Binomial or Poisson) by a continuous Normal distribution, Y, then you will need to apply a continuity correction:

- $\mathrm{P}(X < 3) \implies \mathrm{P}(Y < 2.5)$
- $\mathrm{P}(X \leqslant 3) \implies \mathrm{P}(Y < 3.5)$
- $\mathrm{P}(X \geqslant 3) \implies \mathrm{P}(Y > 2.5)$
- $\mathrm{P}(X > 3) \implies \mathrm{P}(Y > 3.5)$
- $\mathrm{P}(1 < X \leqslant 4) \implies \mathrm{P}(1.5 < Y < 4.5)$

Note that it doesn't matter if we use strict or non-strict inequalities for Y — as Y is continuous we have $\mathrm{P}(Y < 2.5) = \mathrm{P}(Y \leqslant 2.5)$.

Continuous Distributions

The *Normal distribution* is one example of a continuous distribution.

A continuous distribution is defined by a *probability density function* (or p.d.f.) and a range of possible values. The p.d.f. for the Standard Normal distribution is $\mathrm{f}(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}x^2}$.

For a continuous distribution with p.d.f. $\mathrm{f}(x)$ we have:

- $\mathrm{P}(a \leqslant X \leqslant b) = \displaystyle\int_a^b \mathrm{f}(x)\,\mathrm{d}x$
- $\displaystyle\int_{-\infty}^{\infty} \mathrm{f}(x)\,\mathrm{d}x = 1$ (as the total probability must be 1). This means that the total area under the curve $y = \mathrm{f}(x)$ must be 1.

  If $\mathrm{f}(x)$ is equal to 0 for some ranges of x then you will be able to change the limits accordingly. For example if $\mathrm{f}(x) = kx$ for $0 \leqslant x \leqslant 10$, and is equal to 0 elsewhere, then we could write $\displaystyle\int_0^{10} \mathrm{f}(x)\,\mathrm{d}x = 1$ (and hence find the value of k).
- The *expectation* (or mean) is given by:
  $$\mu = \int_{-\infty}^{\infty} x\mathrm{f}(x)\,\mathrm{d}x$$
- The *variance* is given by:
  $$\int_{-\infty}^{\infty} x^2\mathrm{f}(x)\,\mathrm{d}x - \mu^2$$
- The *cumulative distribution function* is defined by:
  $$\mathrm{F}(x) = \mathrm{P}(X \leqslant x) = \int_{-\infty}^{x} \mathrm{f}(t)\,\mathrm{d}t$$
  Here we have taken the lower limit as $-\infty$. It may be that $\mathrm{f}(x) = 0$ for $-\infty < x < a$ say, in which case we could write the lower limit as a. Note the use of "dummy variable" t inside the integral — we cannot use x inside the integral as it is used as a limit.
- $\mathrm{f}(x) = \mathrm{F}'(x)$
- $\mathrm{P}(a \leqslant X \leqslant b) = \mathrm{F}(b) - \mathrm{F}(a)$
- The *median*, m satisfies $P(X \leqslant m) = P(X \geqslant m) = \frac{1}{2}$, i.e. $\displaystyle\int_{-\infty}^{m} \mathrm{f}(x)\,\mathrm{d}x = \int_{m}^{\infty} \mathrm{f}(x)\,\mathrm{d}x = \frac{1}{2}$. Equivalently we have $\mathrm{F}(m) = \frac{1}{2}$.
- The *mode* is where the probability distribution function has a maximum (there may be more than one!). With piece-wise p.d.f.s you might have to consider the boundary points separately.

Note that for a Normal distribution we have mean = median = mode.

Continuous Uniform Distribution

The *continuous uniform distribution* or *rectangular* distribution is defined between two end points a and b has a p.d.f. which is a horizontal line segment. Since the total area under this line segment must be equal to 1 then the line is at a height of $\dfrac{1}{b-a}$ and so the p.d.f. is given by:

$$\mathrm{f}(x) = \frac{1}{b-a}, \quad x \in [a, b]$$

.

The mean of the uniform distribution is given by:

$$\begin{aligned}\int_a^b \frac{x}{b-a}\,\mathrm{d}x &= \left[\frac{x^2}{2(b-a)}\right]_a^b \\ &= \frac{b^2-a^2}{2(b-a)} \\ &= \frac{(b+a)(b-a)}{2(b-a)} \\ &= \tfrac{1}{2}(a+b)\end{aligned}$$

This makes sense as you would expect the mean to be half way through the rectangle.

Using $\mathrm{Var}(X) = \displaystyle\int_a^b \frac{x^2}{b-a}\,\mathrm{d}x - [\mathrm{E}(X)]^2$ gives the variance as $\dfrac{(b-a)^2}{12}$.

You might like to do the integration and show that the variance is actually as given above!