S2 Cheat Sheet

1 Binomial and Poisson

1.1 Be able to list the assumptions

Binomial	Poisson
• Fixed number of trials • Probability of success constant • Each trial is independent • Each trial has two outcomes	• Events occur independently • No simultaneous events • A fixed rate at which events occur

1.2 Geometric distribution questions - Binomial

Where X is the distribution for the number of successes and Y is the distribution for the number of failures, and n is taken from the distribution

$$P(X < k) = P(Y > n - k)$$

$$P(X \geqslant k) = P(Y \leqslant n - k)$$

1.3 Finding an unknown n or p from a context

Set up the probability described in terms of the binomial distribution formula, then solve.

Example

I play a game for which the probability of winning is 0.7. If I win every game, what is the smallest number of times I play such that the probability of winning every game is less than 0.01?

$$P(X = n) = 0.7^n < 0.01$$

$$n \log(0.7) < \log(0.01)$$

$$n > \frac{\log(0.01)}{\log(0.7)} = 12.9$$

At least 13 games

1.4 Solving double inequalities

Given $X \sim B(50, 0.6)$ *find the smallest value of* k *such that* $P(X < k) \geqslant 0.9$

$$X \sim B(50, 0.6) \quad Y \sim B(50, 0.4)$$

$$P(X < k) = P(Y > 50 - k) \geqslant 0.9$$

$$1 - P(Y \leqslant 50 - k) \geqslant 0.9$$

$$P(Y \leqslant 50 - k) \leqslant 0.1$$

$$50 - k \leqslant 15$$

$$k \geqslant 35 \therefore k = 35$$

1.5 Feeding a poisson into a binomial

Sometimes the probability calculated by possion can then be fed into binomial for a further calculation, for example: *Defects occur at a rate of 0.5 per 10cm. If bob boys 6 planks each of length 100cm, find the probability fewer than 2 planks contain at most 3 defects*

$$X \sim P_o(5) \quad P(X \leqslant 3) = 0.2560$$

$$Y \sim B(6, 0.256) \quad P(Y < 2) = 0.4987$$

2 Continuous random variables

2.1 Find the probability of a range of values

$$P(a < X < b) = \int_a^b f(x)\ dx$$

$$P(X > a) = \int_a^\infty f(x)\ dx$$

For the second term, the probability will become zero after a certain value of x, use this as the upper limit

2.2 Point probabilities

The probability at any given value of x for a continuous random variable is zero

2.3 State the PDF given a graph

The PDF from a graph is the equations of the lines over the range, but it is important to include the statement that the probability is zero otherwise, for example

$$f(x) = \begin{cases} \frac{1}{2}(x-3), & \text{for } 3 \leq x \leq 5 \\ 0, & \text{otherwise} \end{cases}$$

2.4 Calculate measures of location and spread

Mean and variance and $E(X^2)$ are given on the formula book.
To calculate quartiles/deciles set $F(x)$ equal to the location through the distribution and solve. For example the first quartile would be found using $F(x) = 0.25$ and the third quartile would be found using $F(x) = 0.75$.
The mode can be found either graphically or where $f'(x) = 0$, making sure that $f''(x) < 0$

2.5 Converting between f(x) and F(x)

To find $F(x)$, integrate f(x), the +c should be used to make sure that the start of the next range is equal to the previous range

2.6 Finding greater than probabilities

As the probability of any given value:

$$P(X \geqslant 10) = 1 - P(X \leqslant 10)$$

Unlike for discrete distributions

3 Continuous uniform distribution

3.1 Proof of formulas on the data sheet

3.1.1 Mean

$f(x) = \frac{1}{b-a}$

$$E(x) = \frac{1}{b-a}\int_a^b x\ dx = \frac{1}{b-a}\times\frac{1}{2}\left[x^2\right]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{1}{2}(a+b)$$

3.1.2 Variance

$f(x) = \frac{1}{b-a}$

$$E(x^2) = \frac{1}{b-a}\int_a^b x^2\ dx = \frac{1}{b-a}\times\frac{1}{3}\left[x^3\right]_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{(b-a)(a^2+ab+b^2)}{3(b-a)} = \frac{a^2+ab+b^2}{3}$$

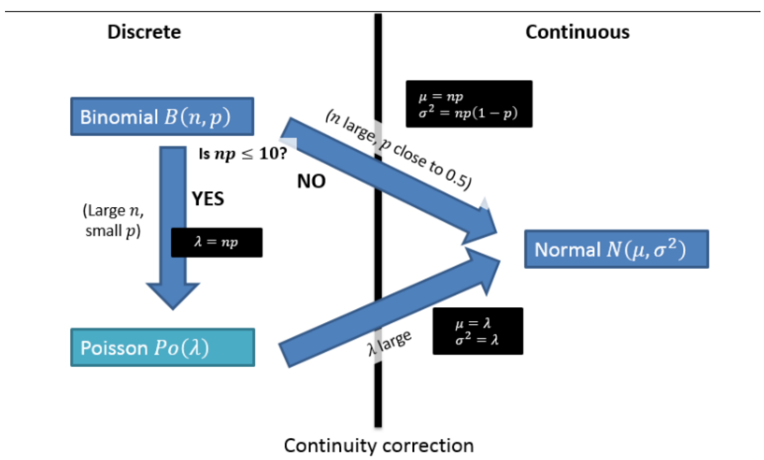
$$Var(x) = E(x^2) - (E(x))^2$$

$$Var(x) = \frac{a^2+ab+b^2}{3} - \frac{(a+b)^2}{4} = \frac{4a^2+4ab+4b^2}{12} - \frac{3a^2+6ab+3b^2}{12} = \frac{a^2-2ab+b^2}{12} = \frac{(b-a)^2}{12}$$

3.2 Finding probability over a range

When finding the probability over a range, multiply the PDF by the width, but remember to check that the range does not extend beyond the range of the distribution, if so, multiply by the width that overlaps with the distribution.

4 Approximations



All the values for approximation are given on the formula book

4.1 Continuity correction

Continuity correction is required when converting from a discrete to a continuous distribution.
Method:

- Make sure the inequality uses $\geqslant$ or $\leqslant$
- "Extend" the range by 0.5 at each end

4.2 Method for a normal approximation

1. Determine the correct approximation to use
2. Identify the original distribution
3. Write the approximation
4. Carry out continuity correction
5. Find the probability

5 Population and samples

5.1 Definitions

- **Statistic -** A random variable which is some function of the sample and not dependent on any population parameters
- **Population -** A collection of all items
- **Sample -** Some subset of the population which is intended to be representative of the population
- **Census -** When the entire population is samples
- **Sampling unit -** Individual member or element of the population or sampling frame
- **Sampling frame -** A list of all sampling units or all the population
- **Sampling distribution -** All possible samples that are chosen from a population, the values of a statistic and the associated probabilities.