**Lesson 5**

**The Poisson Distribution**
**Further Statistics 1**

## 5.1 Approximations of The Binomial Distribution

Historically, the Normal and Poisson distributions were used as approximations to the binomial distribution because, for large sample sizes, calculating a binomial distribution became awkward. However, modern computing power has removed this reason, and the fact that some distributions can model by other distributions is now primarily considered because it's mathematically interesting.

## 5.2 The Normal Approximation of The Binomial Distribution

Previously we have looked at the conditions under which the Normal distribution was a satisfactory approximation of the binomial distribution.
Here are the criteria,

---

If $n$ is large and $p$ is close to 0.5, then the binomial distribution $X \sim B\,(\,n, p\,)$ can be approximated by the normal distribution $N\,(\,\mu, \sigma^2\,)$ where

$$\mu = np$$

$$\sigma = \sqrt{np\,(\,1 - p\,)}$$

---

## 5.3 The Poisson Approximation of The Binomial Distribution

Written in the same style, here are the criteria for when the Poisson distribution can be used to approximate the binomial distribution,

---

If $n$ is large and $p$ is small, then the binomial distribution $X \sim B\,(\,n, p\,)$ can be approximated by the Poisson distribution $Po\,(\lambda)$ where

$$\lambda = np$$

---

Typically,

- the phrase "$n$ is large" suggests $n \geqslant 100$
- the phrase "$p$ is small" suggests $p \leqslant 0.2$
- These two suggestions combine to give the requirement that $np \leqslant 10$

Outside of these limits the approximation becomes inaccurate.

**5.4 Example**

Given the random variable,

$$X \sim B(100,\ 0.1)$$

**( i )** Find $p\,( X \leqslant 4 )$

Give your answer to six decimal places.

**( ii )** Calculate the mean, $\mu$ and the variance, $\sigma^2$

Explain how your answers justify the use of a Poisson distribution to approximate $X$

**( iii )** How else could you justify the use of a Poisson distribution to approximate $X$ ?

**( iv )** Using an appropriate Poisson distribution, $Y$, to approximate $X$, calculate, $p\,( Y \leqslant 4 )$

Give your answer to six decimal places.

**( v )** What is the percentage error in the approximation ?

This is close to a 'worst case' situation; it's typically far better than this !

## 5.5 Exercise

**Question 1**

A discrete random variable, $X$, has a binomial distribution given by,

$$X \sim B(40,\ 0.1)$$

**( a )** Calculate, accurate to four decimal places,

**( i )** $p\ (X = 8)$ **( ii )** $p\ (X \geqslant 6)$

**( b )** Give a reason why it is not appropriate to approximate $X$ with the use of the Poisson distribution.

**Question 2**

A discrete random variable, $W$, has a binomial distribution given by,

$$W \sim B(400,\ 0.6)$$

**( a )** Calculate, accurate to four decimal places,

**( i )** $p\ (W \leqslant 245)$ **( ii )** $p\ (230 \leqslant W \leqslant 250)$

**( b )** Give a reason why it is not appropriate to approximate $W$ with the use of the Poisson distribution.

**Question 3**

A discrete random variable, $Y$, has a Poisson distribution given by,

$$Y \sim Po(40)$$

Calculate, accurate to four decimal places,

**( i )** $p\ (Y = 29)$ **( ii )** $p\ (Y < 35)$

**Question 4**

There were 792 pupils at Shrewsbury School for the academic year 2018-2019.

**( a )** Find the probability that exactly 4 of them have a birthday on 21st January by,

**( i )** using a binomial distribution

**( ii )** using a Poisson approximation

**( b )** Explain why answers given to three decimal places are not sufficiently accurate to calculate the percentage error in using the Poisson approximation in part ( a )

**( c )** Calculate the percentage error, correct to three significant figures in using the Poisson distribution rather than the binomial distribution in part ( a )

**Question 5**

Tony has a dice rolling obsession.
He insists on rolling his twenty faced dice, numbered with the integers from 1 to 20 every hour, on the hour, day and night.
(If truth be told, Tony's not much fun to be with)

He only eats a meal when his dice rolls a 'lucky' 7.

**( i )** Write down a description for $X$, the binomial distribution for the number of meals that Tony eats in a 168 hour week. $(7 \times 24$ hours)

**( ii )** Calculate the probability that Tony has no more than 7 meals in a 168 hour week.

**( iii )** Explain why the binomial distribution, $X$, can be approximated by a Poisson distribution, $Y$, where

$$Y \sim Po(8.4)$$

**( iv )** Repeat the part ( ii ) calculation but this time using $Y$

**( v )** What is the percentage error in using the Poisson distribution, $Y$, rather than the binomial distribution, $X$ ?

**Question 6**

The random variable $X \sim B$ (200, 0.98)

**( a )** Calculate, accurate to four decimal places,

**( i )** $p\ (X = 197)$

**( ii )** $p\ (X \geqslant 195)$

**( b )** Although 0.98 is not a small probability, the part ( a ) answers can be found using a Poisson distribution.

Create a variable $W \sim B$ (200, 0.02).

Show how to how obtain the part (a) answers using a Poisson approximation to $W$

**Question 7**

**( i )** Is the Poisson distribution a discrete or a continuous distribution ?

**( ii )** What is the relationship between the mean and the variance for the Poisson distribution ?

**( iii )** Give three conditions that must be satisfied for the Poisson distribution to be used to approximate a binomial distribution.

**Question 8**

An angler is known to catch fish at a mean rate of 2 per hour.
The number of fish caught by the angler in an hour follows a Poisson distribution.
The angler takes 5 fishing trips, each lasting 2 hours.
Find the probability that the angler catches at least 5 fish on exactly 3 of these trips.

HINT : Find $p\ (X \geqslant 5)$ using a Poisson variable.
Then use this value as the parameter $p$ in a binomial model.