Your Starter for 10
Using the data set
please have a go at these questions while you wait for the session to start

• Which is further North: Hurn or Cambourne?
• Which location in the UK is highest?
• What scales are used for measuring windspeed?
• How are snow or hail measured?
• Why can’t we test the hypothesis “It’s sunnier in the UK than in Perth”?
• Find some discrete data
• Find some categorical data
• Why might the Daily Mean Wind Direction not be a reliable indicator?
• What does a wind direction of 10 mean?
• What does “tr” mean?
• What is “normal air pressure” (not in the data set)
• Which is further North: Hurn or Cambourne? **Hurn**

• Which location in the UK is highest? **Cambourn 87m**

• What scales are used for measuring windspeed? **Beaufort conversion, knots**

• How are snow or hail measured? **Melted first and then measured as rain**

• Why can’t we test the hypothesis “It’s sunnier in the UK than in Perth”? **No sunshine data for overseas stations**

• Find some discrete data: **Cloud Cover, Wind Direction of 20**

• Find some categorical data: **Wind Cardinal Direction (NNW)**

• Why might the Daily Mean Wind Direction not be a reliable indicator? **Doesn’t give an indication of how variable the wind has been.**

• What does a wind direction of NE mean? **Wind is blowing FROM the North East and not towards the North East**

• What does “tr” mean? **A trace of rain but not enough to measure**

• What is “normal air pressure” (not in the data set) **1013 hPa**
Getting Ready to Teach Online course
The A level reforms

• All new AS and A levels will be assessed at the same standard as they are currently
• All new AS and A levels will be fully linear
• AS levels will be stand-alone qualifications
• The content of the AS level can be a sub-set of the A level content to allow co-teachability, but marks achieved in the AS will not count towards the A level
The A level reforms - content

A level Mathematics

• 100% core content
• Pure mathematics (broadly same as C1 to C4)
• Mechanics (mainly from M1 and M2)
• Statistics (mainly from S1 and S2)
• AS content shown in bold font
The A level reforms - content

- Requirement for the assessment of problem solving, communication, proof, modelling, application of techniques
- Requirement for a pre-release large data set (mathematics only)
- Requirement that candidates have a calculator with
  - the ability to compute summary statistics and access probabilities from standard statistical distributions
  - an iterative function
  - the ability to perform calculations with matrices up to at least order $3 \times 3$ (further mathematics only)
Our design principles

• Separate pure and applied papers
• Simple 2:1 ratio of pure to applied (A level Mathematics)
• Single large data set for the lifetime of the qualification – may be amended
• Further mathematics designed to aid parallel delivery with mathematics
• No non-calculator assessment
GCE Mathematics (2017)

Content
Overview of the specification

AS level Mathematics

<table>
<thead>
<tr>
<th>Paper 1: Pure Mathematics</th>
<th>62.5%</th>
<th>2 hours</th>
<th>100 marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper 2: Mechanics and Statistics</td>
<td>37.5%</td>
<td>1 hour 15 minutes</td>
<td>60 marks</td>
</tr>
</tbody>
</table>

[Image of pinecones]
Overview of the specification

A level Mathematics

<table>
<thead>
<tr>
<th>Paper 1: Pure Mathematics</th>
<th>33%</th>
<th>2 hours</th>
<th>100 marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper 2: Pure Mathematics</td>
<td>33%</td>
<td>2 hours</td>
<td>100 marks</td>
</tr>
<tr>
<td>Paper 3: Mechanics and Statistics</td>
<td>33%</td>
<td>2 hours</td>
<td>100 marks</td>
</tr>
</tbody>
</table>

Any pure content can be assessed on either paper
Applied Content Changes

- Standardised content across all boards
- More emphasis on Statistics at AS
- Statistics relies heavily on the use of calculators for distributions
<table>
<thead>
<tr>
<th>Unit</th>
<th>Title</th>
<th>Estimated hours</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Section A – Statistics</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Statistical sampling</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>Introduction to sampling terminology; Advantages and disadvantages of sampling</td>
<td>1</td>
</tr>
<tr>
<td>b</td>
<td>Understand and use sampling techniques; Compare sampling techniques in context</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Data presentation and interpretation</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>Calculation and interpretation of measures of location; Calculation and interpretation of measures of variation; Understand and use coding</td>
<td>4</td>
</tr>
<tr>
<td>b</td>
<td>Interpret diagrams for single-variable data; Interpret scatter diagrams and regression lines; Recognise and interpret outliers; Draw simple conclusions from statistical problems</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>Probability: Mutually exclusive events; Independent events</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Statistical distributions: Use discrete distributions to model real-world situations; Identify the discrete uniform distribution; Calculate probabilities using the binomial distribution (calculator use expected)</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>Statistical hypothesis testing</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>Language of hypothesis testing; Significance levels</td>
<td>2</td>
</tr>
<tr>
<td>b</td>
<td>Carry out hypothesis tests involving the binomial distribution</td>
<td>5</td>
</tr>
</tbody>
</table>

**Total Estimated Hours:** 30
<table>
<thead>
<tr>
<th>Unit</th>
<th>Title</th>
<th>Estimated hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Regression and correlation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>Change of variable</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>Correlation coefficients</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Statistical hypothesis testing for zero correlation</td>
</tr>
<tr>
<td>2</td>
<td>Probability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>Using set notation for probability</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>Questioning assumptions in probability</td>
</tr>
<tr>
<td>3</td>
<td>The Normal distribution</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>Understand and use the Normal distribution</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>Use the Normal distribution as an approximation to the binomial distribution</td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>Statistical hypothesis testing for the mean of the Normal distribution</td>
</tr>
</tbody>
</table>

30 hours
GCE Mathematics (2017)

Standard Deviation and The Normal Distribution
What is the Normal Distribution (A Level)

These histograms show the distribution of heights of adult males in a particular city. As the class width reduces, the distribution gets smoother.
The Normal Distribution
A Level
Standard Deviation AS Level

A key measure of spread

Find the standard deviation of the following number: 3, 4, 6, 8, 8, 5

\[
\bar{x} = \frac{\Sigma fx}{n} = \frac{3 + 4 + 6 + 8 + 8 + 5}{n} = \frac{34}{6}
\]
<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>8</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{x}$</td>
<td>$\frac{34}{6}$</td>
<td>$\frac{34}{6}$</td>
<td>$\frac{34}{6}$</td>
<td>$\frac{34}{6}$</td>
<td>$\frac{34}{6}$</td>
<td>$\frac{34}{6}$</td>
</tr>
<tr>
<td>$x - \bar{x}$</td>
<td>$\frac{-8}{3}$</td>
<td>$\frac{-5}{3}$</td>
<td>$\frac{1}{3}$</td>
<td>$\frac{7}{3}$</td>
<td>$\frac{7}{3}$</td>
<td>$\frac{-2}{3}$</td>
</tr>
<tr>
<td>$(x - \bar{x})^2$</td>
<td>$\frac{64}{9}$</td>
<td>$\frac{25}{9}$</td>
<td>$\frac{1}{9}$</td>
<td>$\frac{49}{9}$</td>
<td>$\frac{49}{9}$</td>
<td>$\frac{4}{9}$</td>
</tr>
</tbody>
</table>

\[ \text{Var, } \sigma^2 = \frac{\sum (x-\bar{x})^2}{n} = \frac{\left(\frac{64}{3}\right)}{6} = \frac{32}{9} \]

\[ \text{SD, } \sigma = \sqrt{\text{Var}} = \sqrt{\frac{32}{9}} = 1.86 \text{ (to 2 dp)} \]
Easier Method

\[ \text{Var } \sigma^2 = \frac{\sum x^2}{n} - \frac{\sum x}{n} = \frac{\sum x^2}{n} - \bar{x}^2 \]

\[ \text{SD } \sigma = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2} \]

\[ \text{SD } \sigma = \sqrt{\frac{3^2 + 4^2 + 6^2 + 8^2 + 8^2 + 5^2}{6} - \left(\frac{34}{6}\right)^2} \]

\[ \text{SD } \sigma = \sqrt{\frac{214}{6} - \left(\frac{34}{6}\right)^2} = 1.86 \text{ (to 2dp)} \]

The mean of the squares minus the square of the mean.
# Grouped Data

<table>
<thead>
<tr>
<th>Length of telephone call (l min)</th>
<th>Frequency</th>
<th>Midpoint ( x )</th>
<th>( fx )</th>
<th>( fx^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 &lt; l ≤ 5</td>
<td>4</td>
<td>2.5</td>
<td>4 × 2.5 = 10</td>
<td>4 × 6.25 = 25</td>
</tr>
<tr>
<td>5 &lt; l ≤ 10</td>
<td>15</td>
<td>7.5</td>
<td>112.5</td>
<td>843.75</td>
</tr>
<tr>
<td>10 &lt; l ≤ 15</td>
<td>5</td>
<td>12.5</td>
<td>62.5</td>
<td>781.25</td>
</tr>
<tr>
<td>15 &lt; l ≤ 20</td>
<td>2</td>
<td>17.5</td>
<td>35</td>
<td>612.5</td>
</tr>
<tr>
<td>20 &lt; l ≤ 60</td>
<td>0</td>
<td>40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>60 &lt; l ≤ 70</td>
<td>1</td>
<td>65</td>
<td>65</td>
<td>4225</td>
</tr>
<tr>
<td><strong>total</strong></td>
<td><strong>27</strong></td>
<td></td>
<td><strong>285</strong></td>
<td><strong>6487.5</strong></td>
</tr>
</tbody>
</table>

\[ \sum fx^2 = 6487.5 \quad \sum fx = 285 \quad \sum f = 27 \]

\[ \sigma^2 = \frac{6487.5}{27} - \left( \frac{285}{27} \right)^2 = 128.85802 \]

\[ \sigma = \sqrt{128.85802} = 11.35 \]
Grouped Data

\[ \sigma = \sqrt{\frac{\sum fx^2}{n} - \left(\frac{\sum fx}{n}\right)^2} = \sqrt{\frac{\sum fx^2}{n} - \bar{x}^2} \]
Standard Notation

- $X$ is distributed normally
- Mean = 8
- SD = 0.2

8 is the mean of the distribution. The normal distribution is **symmetrical**, so for any normally distributed random variable $P(X > \mu) = 0.5$. 

\[ X \sim N(8, 0.2^2) \]
Key Cases

a. $P(X < 33) = 0.7734$ (4 d.p.)

b. $P(X \geq 24) = P(X > 24) = 0.9332$ (4 d.p.)

c. $P(33.5 < X < 38.2) = 0.1706$ (4 d.p.)

d. $P(X < 27 \text{ or } X < 32) = 1 - P(27 < X < 32)$
   
   $= 1 - 0.4648$
   
   $= 0.5352$ (4 d.p.)
Standard Normal Distribution

Useful for finding missing mean and SD values

\[ Z \sim N(\mu, \sigma^2) \]

where \[ Z \sim N(0, 1^2) \]

If \[ X \sim N(50, 4^2) \] find \( P(X < 53) \)

We “could” use coding to relate it to the Standard Normal Distribution
Coding

\[ X \sim N(50, 4^2) \]
\[ P(X < 53) \]

\[ Z \sim N(0, 1^2) \]
\[ P(Z < 0.75) \]
Coding

If \( X \sim N(50, 4^2) \) find \( P(X < 53) \)

\[
P(x < 53) = P \left( Z < \frac{53 - 50}{4} \right)
\]

\[
= P(z < 0.75)
\]

\[
= 0.7734 \text{(to 4dp)}
\]

Coding formula is \( P \left( Z < \frac{a - \mu}{\sigma} \right) = \alpha \)
The random variable $X \sim N(\mu, 3^2)$. Given that $P(X > 20) = 0.20$, find the value of $\mu$.

We need to use $P(X < 20) = 0.8$

From your calculator $P(Z < 0.8416) = 0.80$

Using $P \left( Z < \frac{a-\mu}{\sigma} \right) = \alpha$ gives $P \left( Z < \frac{20-\mu}{3} \right) = 0.20$

So $\frac{20-\mu}{3} = 0.8416$

Therefore $\mu = 20 - 3 \times 0.8146 = 17.4$ (to 3 sf)
GCE Mathematics (2017)

Binomial Distribution

AS Level
(examined also at
A Level)
Binomial Distribution

**Pure-related question:**
In the following expansion, how many different ways could one obtain $a^3b^2$ terms?
Expand: $(a + b)^5$
**Pure-related question:**
In the following expansion, how many different ways could one obtain $a^3b^2$ terms?

Expand:

$$(a + b)(a + b)(a + b)(a + b)(a + b)$$
**Pure-related question:**
In the following expansion, how many different ways could one obtain \(a^3b^2\) terms?

Expand:

\[(a + b)(a + b)(a + b)(a + b)(a + b)\]
Binomial Distribution

**Pure-related question:**
Expand:

\[(a + b)(a + b)(a + b)(a + b)(a + b)\]

Bracket 1: \(a\)
Bracket 2: \(a\)
Bracket 3: \(b\)
Bracket 4: \(b\)
Bracket 5: \(a\)

Result: \(a^3b^2\)
Binomial Distribution

**Pure-related question:**
Expand:

$$(a + b)(a + b)(a + b)(a + b)(a + b)$$

Bracket 1: a
Bracket 2: a
Bracket 3: b
Bracket 4: b
Bracket 5: a

How many more ways?
Pure-related question:
Expand:
\[(a + b)(a + b)(a + b)(a + b)(a + b)\]
\[a^3b^2\]
We know from Binomial Expansion (in pure) that the coefficient of \(a^3b^2\) would be: \(\binom{5}{3}\)
Binomial Distribution

**Pure-related question:**
Expand:

$$(a + b)(a + b)(a + b)(a + b)(a + b)$$

Binomial: two parts
Binomial Distribution

Binomial expansion and Binomial distribution:
Variables which have: only two possible (mutually exclusive) outcomes are distributed binomially.

The Binomial distribution can be used to model repeated trials of a variable with two possible outcomes.

Example: We throw a biased coin 5 times and evaluate the probability of obtaining a head at least 3 times when the probability of throwing a head is 0.7
Binomial Distribution

**Binomial expansion and Binomial distribution:**

**Example:** We throw a biased coin 5 times and evaluate the probability of obtaining a head at least 3 times when the probability of throwing a head is 0.7

**Key information:**
Number of trials \((n) = 5\)
Probability of success \((p) = 0.7\)
Calculate \(P(X \geq 3)\)
Binomial expansion and Binomial distribution:

**Key information:**
Number of trials \( (n) = 5 \)
Probability of success \( (p) = 0.7 \)

Calculate \( P(X \geq 3) \)

\[
P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5)
\]

We can work out each of these individually...
Binomial Distribution

Binomial expansion and Binomial distribution:

Key information:
Number of trials (n) = 5
Probability of success (p) = 0.7
P(X = 3) + P(X = 4) + P(X = 5)
Binomial expansion and Binomial distribution:

\[ P(X = 3) + P(X = 4) + P(X = 5) \]
Binomial expansion and Binomial distribution:

\[ P(X = 3) \]

Tree diagram skeleton:
Binomial expansion and Binomial distribution:

P(X = 3)

Tree diagram skeleton:
Binomial Distribution

Binomial expansion and Binomial distribution:

\[ P(X = 3) \]

Tree diagram skeleton:

But we want exactly three heads if we are considering \( P(X = 3) \)...
Binomial Distribution

**Binomial expansion and Binomial distribution:**

\[ P(X = 3) \]

Tree diagram skeleton:

How many other final outcomes would result in exactly 3 heads?
Binomial expansion and Binomial distribution:

\[ P(X = 3) \]

Tree diagram skeleton:

If I could have either “H” or “T” each time out of five, how many ways could I end up with exactly 3 Hs?
Binomial Distribution

*Binomial expansion and Binomial distribution:*

\[ P(X = 3) \]

Tree diagram skeleton:

If I expand \((H + T)^5\).

How many terms will have \(H^3\)?
Binomial expansion and Binomial distribution:
P(X = 3)

Tree diagram skeleton:

If I expand \((H + T)^5\). How many terms will have \(H^3 T^2\)?
**Pure-related question:**

Expand:

\[(a + b)(a + b)(a + b)(a + b)(a + b)\]

We know from Binomial Expansion (in pure) that the coefficient of \(a^3b^2\) would be: \({5 \choose 3}\)
Binomial Distribution

Binomial expansion and Binomial distribution:

\[ P(X = 3) \]

Tree diagram skeleton:

There are \( \binom{5}{3} \) number of ways of having exactly 3 heads.

\[
\binom{5}{3} = \frac{5!}{3! (5 - 3)!}
\]
**Binomial expansion and Binomial distribution:**

\[ P(X = 3) \]

Tree diagram skeleton:

There are \( \binom{5}{3} \) number of ways of having exactly 3 heads.

\[
\binom{5}{3} = \frac{5!}{3! \cdot 2!}
\]
Binomial Distribution

Binomial expansion and Binomial distribution:

\[ P(X = 3) \]

Tree diagram skeleton:

There are \( \binom{5}{3} \) number of ways of having exactly 3 heads.

\[ \binom{5}{3} = \frac{5 \times 4}{2} \]
Binomial expansion and Binomial distribution:

\[ P(X = 3) \]

Tree diagram skeleton:

There are \( \binom{5}{3} \) number of ways of having exactly 3 heads.

\[ \binom{5}{3} = 10 \]
Binomial expansion and Binomial distribution:

P(X = 3)

Tree diagram skeleton:

Let’s look at the first two...

There are \( \binom{5}{3} \) number of ways of having exactly 3 heads.

\[
\binom{5}{3} = 10
\]
Binomial Distribution

Binomial expansion and Binomial distribution:

P(X = 3)

Tree diagram skeleton:

Let’s look at the first two...

P(HHHTT) = P(H)P(H)P(H)P(T)P(T)
= 0.7 \times 0.7 \times 0.7 \times P(T)P(T)
= 0.7 \times 0.7 \times 0.7 \times 0.3 \times 0.3
= 0.7^3 \times 0.3^2

P(HHTHT) = P(H)P(H)P(T)P(H)P(T)
= P(H)P(H)P(H)P(T)P(T)
= 0.7^3 \times 0.3^2
**Binomial expansion and Binomial distribution:**

P(X = 3)

Tree diagram skeleton:

Let’s look at the first two...

There are \( \binom{5}{3} \) number of ways of having exactly 3 heads.

\[
\binom{5}{3} = 10
\]
Binomial distribution:

\[ P(X = 3) = 10 \times 0.7^3 \times 0.3^2 \]

\[ = \binom{5}{3} P(H)^3 P(T)^2 \]

\[ = \binom{5}{3} P(H)^3 \left[ 1 - P(H) \right]^2 \]

\[ = \binom{5}{3} p^3 (1 - p)^2 \]

\[ = \binom{n}{3} p^3 (1 - p)^{n-3} \]

Number of different ways to obtain exactly 3 heads:

\[ P(X = r) = \binom{n}{r} p^r (1 - p)^{n-r} \]
Binomial Distribution

**Binomial distribution:**

\[ P(X = 3) = 10 \times 0.7^3 \times 0.3^2 \]

\[ = \binom{5}{3} P(H)^3 P(T)^2 \]

\[ = \binom{5}{3} P(H)^3 [1 - P(H)]^2 \]

\[ = \binom{5}{3} p^3 (1 - p)^2 \]

\[ = \binom{n}{3} p^3 (1 - p)^{n-3} \]

\[ P(X = r) = \binom{n}{r} p^r (1 - p)^{n-r} \]

- Number of different ways to obtain exactly \( r \) heads
- \( P(H) \times P(H) \times \ldots \times P(H) \)
- \( P(T) \times \ldots \times P(T) \)
Binomial Distribution

**Binomial distribution:**

More generally for anything distributed binomially:

\[ P(X = r) = \binom{n}{r} p^r (1 - p)^{n-r} \]

- Number of different ways to obtain exactly \( r \) successes
- \( p \times p \times p \times \ldots \times p \)
- \( (1 - p) \times (1 - p) \times \ldots \times (1 - p) \)

where \( p = P(\text{success}) \)

where \( 1 - p = P(\text{failure}) \)
Binomial Distribution

Back to the question... Binomial distribution:

\[ P(X = 3) + P(X = 4) + P(X = 5) \]

\[ = 10 \times 0.7^3 \times 0.3^2 + P(X = 4) + P(X = 5) \]

\[ = 10 \times 0.7^3 \times 0.3^2 + \binom{5}{4}0.7^4 \times 0.3^1 + \binom{5}{5}0.7^5 \times 0.3^0 \]

\[ = 10 \times 0.7^3 \times 0.3^2 + 5 \times 0.7^4 \times 0.3^1 + 1 \times 0.75 \times 1 \]

\[ = 10 \times 0.7^3 \times 0.3^2 + 5 \times 0.7^4 \times 0.3^1 + 0.7^5 \]

\[ = 0.3087 + 0.36015 + 0.16807 \]

\[ = 0.83692 \]

Remember this – you’ll need it (much) later
Expected Value and Variance of a Binomially distributed variable:

Since the Binomial distribution has two possible outcomes (usually success and failure) we can represent by coding using the binary system i.e.:

Success = 1; Failure = 0

If we repeat the trial 20 times and the probability of success is 0.8 then our expected value would be:

20 x 0.8

Why?
Binomial Distribution

**Expectation Value and Variance of a Binomially distributed variable:**

Discrete Random Variable, $X$, with distribution as follows:

<table>
<thead>
<tr>
<th>Outcome</th>
<th>T</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coded outcome ($x$)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$P(X = x)$</td>
<td>0.2</td>
<td>0.8</td>
</tr>
<tr>
<td>$x P(X = x)$</td>
<td>0</td>
<td>0.8</td>
</tr>
<tr>
<td>$x^2 P(X = x)$</td>
<td>0</td>
<td>0.8</td>
</tr>
</tbody>
</table>

$$E(X) = \sum_{i=1}^{2} x_i P(X = x_i) = 0.8$$

$$E(X^2) = \sum_{i=1}^{2} x_i^2 P(X = x_i) = 0.8$$

$$Var(X) = E(X^2) - [E(X)]^2 = 0.8 - 0.64 = 0.16$$
Expectation Value and Variance of a Binomially distributed variable:

In general, for any Binomially distributed variable, $X$...

$$E(X) = \sum_{i=1}^{2} x_i P(X = x_i)$$

$$= p$$

$$E(X^2) = \sum_{i=1}^{2} x_i^2 P(X = x_i)$$

$$= p$$

$$Var(X) = E(X^2) - [E(X)]^2$$

$$= p - p^2$$

$$= p(1 - p)$$

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Fail</th>
<th>Success</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coded outcome ($x$)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$P(X = x)$</td>
<td>$1 - p$</td>
<td>$p$</td>
</tr>
<tr>
<td>$x P(X = x)$</td>
<td>0</td>
<td>$p$</td>
</tr>
<tr>
<td>$x^2 P(X = x)$</td>
<td>0</td>
<td>$p$</td>
</tr>
</tbody>
</table>
Expectation Value and Variance of a Binomially distributed variable:

2 repetitions...

\[ E(X) = \sum_{i=1}^{3} x_i p(X = x_i) \]
\[ = 2p - 2p^2 + 2p^2 \]
\[ = 2p \]

\[ E(X^2) = \sum_{i=1}^{3} x_i^2 p(X = x_i) \]
\[ = 2p - 2p^2 + 4p^2 \]
\[ = 2p + 2p^2 \]

\[ Var(X) = E(X^2) - [E(X)]^2 \]
\[ = 2p + 2p^2 - (2p)^2 \]
\[ = 2p - 2p^2 \]
\[ = 2p(1 - p) \]
**Binomial Distribution**

*Expectation Value and Variance of a Binomially distributed variable:*

\[
E(X) = \sum_{i=1}^{n} x_i P(X = x_i) = np
\]

\[
Var(X) = np(1 - p)
\]

Showing the Variance of the binomial is beyond the scope of the course but does not require an awful lot more. You will need to think about how many ways, having repeated \(n\) times, you could obtain a total of, say, 5 to complete the table above for possible values.

It may be helpful to consider \(n = 3\) and \(n = 4\) initially as well as the scenarios \(n = 1\) and \(n = 2\) that I have shown here in order to then be able to obtain the result...
Binomial when $n$ is ‘large’

**Binomial distribution:**
Example: Throwing a biased coin 500 times to evaluate the probability of obtaining a head **exactly** 300 times when the probability of throwing a head is 0.56

\[
P(X = 300) = \binom{500}{300} \times 0.56^{300} \times 0.44^{200}
\]

Try it on a calculator...
Binomial distribution:
Example: Throwing a biased coin 500 times to evaluate the probability of obtaining a head exactly 300 times when the probability of throwing a head is 0.56

\[ P(X = 300) = \binom{500}{300} \times 0.56^{300} \times 0.44^{200} \]

The combination calculation at the start of the question is too big for the calculator to compute but...
Binomial distribution & Central Limit Theorem:
If $p \approx 0.5$ and the trial is repeated ‘many’ times then we can use the Normal distribution to approximate the Binomial.

In our example, $n = 500$, $p = 0.56$ (i.e. $P(1) = 0.56$)
\[ E(X) = 500 \times 0.56 \]

From the first session today (with Mark) you will know that a Normal Distribution is symmetric which would reflect a probability of 0.5 and that...

the Central Limit Theorem indicates that when $n$ is large, the Normal Distribution provides a good approximation.
Binomial distribution & Central Limit Theorem:
If $p \approx 0.5$ and the trial is repeated ‘many’ times then we can use the Normal distribution to approximate the Binomial.
In our example, $n = 500, p = 0.56$ (i.e. P(1) = 0.56)
$E(X) = 500 \times 0.56$
$= 280$
From earlier, we know the variance of $n$ trials of a binomial distribution is $np(1 - p)$ so for our distribution, $Var(X) = 500 \times 0.56 \times 0.44$
$= 123.2
Binomial when $n$ is ‘large’

Normal approximation of Binomial distribution:

$X \sim B(500, 0.56)$

$E(X) = 280$

$\text{Var}(X) = 123.2$

$Y \sim N(280, 123.2)$

Our example asked for $P(X = 300)$...

$P(X = 300) \approx P(299.5 < Y < 300.5)$

$= P\left(\frac{299.5-280}{\sqrt{123.2}} < Z < \frac{300.5-280}{\sqrt{123.2}}\right)$

$= P(Z < \frac{300.5-280}{\sqrt{123.2}}) - P(Z < \frac{299.5-280}{\sqrt{123.2}})$

$= P(Z < \frac{300.5-280}{\sqrt{123.2}}) - P(Z < \frac{299.5-280}{\sqrt{123.2}})$

$= P(Z < 1.8469 \ldots) - P(Z < 1.7568 \ldots)$

$\approx P(Z < 1.85) - P(Z < 1.76)$

$\approx 0.9678 - 0.0509$
Binomial when $n$ is ‘large’

Normal approximation of Binomial distribution:

$$X \sim \text{Binomial}(5, 0.2)$$

$$E(X) = 3.5$$

$$\text{Var}(X) = 1.05$$

$$X \sim \text{Normal}(3.5, 1.05)$$

How good would the Normal approximation have been for the example where $n$ was small and $P(\text{success})$ was not as close to $0.5$...?

Our example asked for $P(X \geq 3)$...

$$P(X \geq 3) \approx P(Y \geq 2.5)$$

$$= 1 - P(Y < 2.5)$$

$$= 1 - P(Z < \frac{2.5 - 3.5}{\sqrt{1.05}})$$

$$= 1 - P(Z < -0.9759 \ldots)$$

$$= 1 - P(Z > 0.9759 \ldots)$$

$$\approx P(Z < 0.9759 \ldots)$$

$$= 0.8365$$

Our answer earlier was... 0.83692

The approximation is pretty good but... the more trials (larger $n$) and the more symmetric (closer $p$ is to 0.5), the better the approximation...
GCE Mathematics (2017)

Hypothesis Testing

AS and A Level
Hypothesis...

$H_0$: All school subjects are equally brilliant

$H_1$: Maths is the greatest subject ever

Not really testable within the scope of the course but here’s a quick intro to hypothesis testing from the Statistics part of the new Maths A-Level specification..
Hypothesis Testing

- Types of hypothesis
  - Null Hypothesis \((H_0)\)
  - Alternate Hypothesis \((H_1)\)
    - One-tailed or two-tailed test
- Significance level
- Drawing a conclusion
Hypothesis Testing

• **Types of hypothesis**
  ➔ **Null Hypothesis (H₀)**
  The ‘boring’ or ‘everything is as it should be’ hypothesis
  ➔ **Alternate Hypothesis (H₁)**
  The ‘interesting’ or ‘we speculate a theory’ hypothesis

**Example:** A manufacturing plant bottles 500ml of Lemonade on industrial scale. The process aims to ensure the volume of lemonade in each bottle is 503ml but there is speculation from the public that the bottles regularly have less than this...
**Example:** A manufacturing plant bottles 500ml of Lemonade on industrial scale. The process aims to ensure the volume of lemonade in each bottle is 503ml but there is speculation from the public that the bottles regularly have less than this...

In this case, our hypotheses would be as follows:

- $H_0$: The mean volume is 503ml
- $H_1$: The mean volume is less than 503ml
Hypothesis Testing

- **Types of hypothesis**
  - Null Hypothesis \((H_0)\)
  - Alternate Hypothesis \((H_1)\)
  - One-tailed or two-tailed test

If the Alternate Hypothesis explicitly puts forward whether the test statistic is either greater than or less than in the Null Hypothesis then the test is... 1-tailed.

If, on the other hand, it puts forward simply that the test statistic is ‘wrong’ then the test is being conducted both at a higher and lower level than stated so it is 2-tailed.
Hypothesis Testing

If the Alternate Hypothesis explicitly puts forward whether the test statistic is either greater than or less than in the Null Hypothesis then the test is 1-tailed.

If, on the other hand, it puts forward simply that the test statistic is ‘wrong’ then the test is being conducted both at a higher and lower level than stated so it is 2-tailed.

In the Lemonade example, the Alternate Hypothesis put forward that the mean volume of lemonade was less than 503ml

1-tailed
Hypothesis Testing

In the Lemonade example, the Alternate Hypothesis put forward that the mean volume of lemonade was less than 503ml.

Here’s an adaptation... Instead of the public speculating that the bottles had insufficient lemonade, the chairman is not convinced that the foreman is doing a rigorous enough job and believes that the mean volume is not 503ml...

2-tailed
Hypothesis Testing

• Types of hypothesis
  → Null Hypothesis (H₀)
  → Alternate Hypothesis (H₁)
  → One-tailed or two-tailed test

• Significance level
Ahead of conducting the test, the significance level should be set. It is a threshold probability for the test statistic.

Example: Consider the one-tailed bottling example from earlier. Suppose the process is believed to be distributed Normally with mean 503, standard deviation 13 and we set the significance level at 5%.
Hypothesis Testing

**Significance level**
Ahead of conducting the test, the significance level should be set. It is a threshold probability for the test statistic.

**Example:** Consider the one-tailed bottling example from earlier. Suppose the process yields bottles whose volume, $X$, is believed to be distributed Normally with mean 503, variance 169 and we set the significance level at 5%.

If we take a sample of 10 bottles and test its mean, we know the sample mean of the bottles, $\bar{X}$, should be distributed Normally with mean 503, variance $\frac{169}{10}$. 


Hypothesis Testing

**Example:** Consider the one-tailed bottling example from earlier. Suppose the process yields bottles whose volume, $X$, is believed to be distributed Normally with mean 503, variance 169 and we set the significance level at 5%.

If we take a sample of 10 bottles and test its mean, we know the sample mean of the bottles, $\bar{X}$, should be distributed Normally with mean 503, variance $\frac{169}{10}$.

If the probability of the sample mean from this sample, $\bar{x}$, being less than or equal to the value it has is less than 5% then the threshold (significance level) criterion has been met and we have sufficient evidence to reject the null hypothesis.
Hypothesis Testing

Adaptation: Consider the **two-tailed** bottling example from earlier. Suppose the process yields bottles whose volume, \( X \), is believed to be distributed Normally with mean 503, variance 169 and we set the significance level at 5%.

If we take a sample of 10 bottles and test its mean, we know the sample mean of the bottles, \( \bar{X} \), should be distributed Normally with mean 503, variance \( \frac{169}{10} \).
Hypothesis Testing

• Types of hypothesis
  ➔ Null Hypothesis \((H_0)\)
  ➔ Alternate Hypothesis \((H_1)\)
  ➔ One-tailed or two-tailed test

• Significance level

• Drawing a conclusion
If the probability of your test statistic is not significant (ie it is more likely than your significance level), then you do not have sufficient evidence to reject the Null Hypothesis.

[Note: there is not enough evidence to accept the Null Hypothesis]...
Think back to the first Binomial from earlier. We had a biased coin, \( P(H) = 0.7 \), the coin was thrown 5 times and the probability of obtaining at least three heads was 0.83692

**Example:**

Sam throws the above coin 5 times and obtains 4 heads. He conjectures that the \( P(H) \) is, in fact, higher than 0.7 and decides to test his claim with a 10% significance level.
Hypothesis Testing

Example:
Sam throws the above coin 5 times and obtains 4 heads. He conjectures that the P(H) is, in fact, higher than 0.7 and decides to test his claim with a 10% significance level.

**Set up the hypothesis test:**

\[ H_0: P(H) = 0.7 \]
\[ H_1: P(H) > 0.7 \]
Significance level = 10%
Hypothesis Testing

Example:
Sam throws the above coin 5 times and obtains 4 heads. He conjectures that the $P(H)$ is, in fact, higher than 0.7 and decides to test his claim with a 10% significance level.

**Set up the hypothesis test:**
$H_0$: $P(H) = 0.7$; $H_1$: $P(H) > 0.7$; **Significance level** = 10%

**Test the hypothesis test:**
We know from earlier that:

$P(X \geq 4) = P(X = 4) + P(X = 5)$

$= \binom{5}{4}0.7^4 \times 0.3^1 + \binom{5}{5}0.7^5 \times 0.3^0$

$= 0.36015 + 0.16807$

$= 0.52822$

$= 52.822\%$

$P(X \geq 4) > 10\%$ (the significance level)

- There is insufficient evidence to reject the Null Hypothesis.
GCE Mathematics (2017)

Calculator Support
Calculator Support

Casio fx-991EX (scientific)

Menu based system

Scientific calculator keys have similar layout but added functionality
Calculator Support

**Casio fx-991EX (scientific)**

**Menu options...**

1. **Calculate**
2. **Complex**
3. **Base-N**
4. **Matrix**
5. **Vector**
6. **Statistics**
7. **Distribution**
8. **Spreadsheet**
9. **Table**

**A. Equation/Func** (solve simultaneous and polynomials)

**B. Inequality**

**C. Ratio**
Find the probability of **exactly** three heads out of five throws of a biased coin $P(H) = 0.7$

Binomial PD

<table>
<thead>
<tr>
<th>$x$</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>5</td>
</tr>
<tr>
<td>$p$</td>
<td>0.7</td>
</tr>
</tbody>
</table>

$P = 0.3087$
Find the probability of up to and including five heads out of fifteen throws of a fair coin.
If $X \sim N(50, 4^2)$ find $P(X < 53)$
Inverse Normal Calculations

\( X \sim N(30, 5^2) \), if \( P(X > a) = 0.4 \), find \( a \)

We need to use \( P(X < a) = 0.6 \)

\[
\begin{array}{l}
1: \text{Normal PD} \\
2: \text{Normal CD} \\
3: \text{Inverse Normal} \\
4: \text{Binomial PD}
\end{array}
\]

\[
\begin{array}{l}
\text{Inverse Normal} \\
\text{Area} : 0.6 \\
\sigma : 5 \\
\mu : 30
\end{array}
\]

\[
\text{xInv=} \\
31.26673547
\]
Storing answers
Recalling Answers
Helen models the number of hours of sunshine each day, for the month of July at Heathrow by $N(6.6, 3.7^2)$.

(e) Use Helen’s model to predict the number of days in July at Heathrow when the number of hours of sunshine is more than 1 standard deviation above the mean.

<table>
<thead>
<tr>
<th>(e)</th>
<th>$[H = \text{no. of hours}] \quad P(H &gt; 10.3) \text{ or } P(Z &gt; 1) = [0.15865... \ ]$</th>
<th>M1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predict</td>
<td>$31 \times 0.15865... = \textbf{4.9 or 5 days}$</td>
<td>A1</td>
</tr>
</tbody>
</table>

(2)
5. (a) The discrete random variable \( X \sim B(40, 0.27) \)

\[
\text{Find } \quad P(X \geq 16)
\]

\[
P(X \geq 16) = 1 - P(X \leq 15)
\]

\[
= 1 - 0.949077\ldots = \text{awrt } 0.0509
\]

\( \therefore \) A1
GCE Mathematics (2017)

Working with the Large Data Set
General Points

• Don’t try and upload the entire large data set into Geogebra or Autograph: They can’t cope with multiple page spreadsheets
  – Copy and paste any data you need into them

• Sample your data in Excel first before pasting into a graphing programme
  – Use filters in Excel to order or select your data

• Be careful when plotting histograms
Random Sampling

Insert Random Sampling Video
When was the best time to go on holiday in 2015?

Choose a location in the UK: Cambourne, Heathrow, Hurn, Leeming, Leuchars
Highlight the date column and the temperature column
Plot a time series graph (use a line graph)
  – Insert a title and axis labels on your graph
How could we improve the graph?
Copy and paste your graph into Word
What is your conclusion in relation to the question?

Extension: is there a common pattern across the UK?
1987? is there a pattern between 2015 and
• Is this pattern local to Cambourne or is it consistent around the UK?
• Pretty meaningless really
• How could we make any trend clearer?
Patterns are clearer with a moving average
Is there a common pattern across the UK?

What about across two different years?
2015 Daily Max Temperature 7 Day Average

1987 Daily Max Temperature 7 Day Average
Comparing Distributions

Where’s the sunniest place to live in the UK

Copy the Total Hours of Sunshine Data for two of the UK locations onto a new spreadsheet.

Open Geogebra.

- Select the spreadsheet option
- Paste the data into the Geogebra spreadsheet
- Check you have more than 100 rows
- Highlight the columns and create box and whisker plots.
- Export the box plot as a picture and insert into Word.
Total Hours of Sunshine

- Cambourne
- Heathrow
- Hurn
- Leeming
- Leuchars

Hours of Sunshine (hrs)
Using Geogebra

Insert Geogebra Video
Categorical Data

Hypothesis: The prevailing wind in the UK is from the South West

Count the frequency of N, NNE, NE etc for one of the locations in the UK

- Use filters to help you count up the data
- More advanced "=countif(M7:M203,NNE)" will count up the frequency of NE

Select an appropriate graph for categorical data from Excel
Are there any alternative graphs we could use?
Extension:
Does the prevailing wind vary around the UK?
Is it consistent between 2015 and 1987?
Hypothesis: Increased cloud cover results in fewer hours of sunshine

Test if there is a correlation by plotting a scatter graph of the Daily Mean Total Cloud Cover against Daily Total Sunshine

- Hint: Excel will use the left hand column for the x axis- you may need to adjust your spreadsheets before plotting graphs
- Remember to add axis labels and a title
Hypothesis: More cloud cover means higher rainfall
Adjusted Hypothesis:
On days where there is “significant rainfall” higher cloud cover means there is more rain

Data is messy and doesn’t fit nice trends but that doesn’t mean we can’t draw conclusions.
Example Questions

1. Sara is investigating the variation in daily maximum gust, $t$ kn, for Camborne in June and July 1987.

She used the large data set to select a sample of size 20 from the June and July data for 1987. Sara selected the first value using a random number from 1 to 4 and then selected every third value after that.

(a) State the sampling technique Sara used.

(1)

(b) From your knowledge of the large data set, explain why this process may not generate a sample of size 20.

(1)
GCE Mathematics (2017)

Support and Resources
Free support

- **Plan:**
  - schemes of work and course planners to help you deliver the qualifications in the best way for your centre
  - content mapping documents
  - Getting started guide

- **Teach:**
  - topic-based resources to use in the classroom, particularly for the new and unfamiliar topics
  - content exemplification

- **Track and Assess:**
  - specimen papers
  - secure mock papers
  - practice papers
  - assessment guide
  - exemplar solutions
  - Exam Wizard & Results Plus

- **Develop:**
  - a full programme of launch and training events – live and online
  - pre-recorded getting ready to teach sessions
  - our collaborative network events
  - the famous Mathematics Emporium, led by Graham Cumming