 Year 12 Mathematics Extension 1

| ME-S1 The binomial distribution | Unit duration |
| --- | --- |
| The topic Statistical Analysis involves the exploration, display and interpretation of data via modelling to identify and communicate key information. A knowledge of statistical analysis enables careful interpretation of situations and an awareness of the contributing factors when presented with information by third parties, including its possible misrepresentation. The study of statistical analysis is important in developing students’ ability to consider the level of reliability that can be applied to the analysis of current situations and to predict future outcomes. It supports the development of understanding of how conclusions drawn from data can be used to inform decisions made by groups such as scientific investigators, business people and policy-makers. | 10 lessons |

| Subtopic focus | Outcomes |
| --- | --- |
| The principal focus of this subtopic is to develop an understanding of binomial random variables and their uses in modelling random processes involving chance and variation. Students develop an understanding of binomial distributions and associated statistical analysis methods and their use in modelling binomial events. Binomial probabilities and the binomial distribution are used to model situations where only two outcomes are possible. The use of the binomial distribution and binomial probability has many applications, including medicine and genetics. | A student:* applies appropriate statistical processes to present, analyse and interpret data ME12-5
* chooses and uses appropriate technology to solve problems in a range of contexts ME12-6
* evaluates and justifies conclusions, communicating a position clearly in appropriate mathematical forms ME12-7
 |

| Prerequisite knowledge | Assessment strategies |
| --- | --- |
| The subtopic MA-S3 Random Variables should be taught prior to this unit, as some of the skills learnt within it are then used here. Students would benefit from being familiar with using probability tables. | There are several opportunities for formative assessment through the activities embedded in this unit of work; including the opportunity for students to identify binomial distributions, with reasons, conducting binomial experiments and investigating sample proportions.  |

All outcomes referred to in this unit come from the [Mathematics Extension 1](http://educationstandards.nsw.edu.au/wps/portal/nesa/11-12/stage-6-learning-areas/stage-6-mathematics/mathematics-extension-1-2017) syllabus
© NSW Education Standards Authority (NESA) for and on behalf of the Crown in right of the State of New South Wales, 2017

Glossary of terms

| Term | Description |
| --- | --- |
| Bernoulli distribution **** | The Bernoulli distribution is the probability distribution of a random variable which takes the value 1 with ‘success’ probability $p$, and the value 0 with ‘failure’ probability $q=1-p$. The Bernoulli distribution is a special case of the binomial distribution, where $n=1$. |
| Bernoulli random variable | A Bernoulli random variable has two possible values, namely 0 representing failure and 1 representing success. The parameter associated with such a random variable is the probability $p$ of obtaining a 1. |
| Bernoulli trial | A Bernoulli trial is an experiment with only two possible outcomes, labelled ‘success’ and ‘failure’. |
| binomial coefficient | The coefficient of the term $x^{n-r}y^{r}$ in the expansion of $(x+y)^{n}$ is called a binomial coefficient. It is written as $ ^{n}C\_{r}$ or $\left(\begin{matrix}n\\r\end{matrix}\right)$ where $r=0, 1, ..., n$ and is given by: $ \frac{n!}{r!(n-r)!}$ |
| binomial distribution **** | The binomial distribution with parameters $n$ and $p$ is the discrete probability distribution of the number of successes in a sequence of $n$ independent Bernoulli trials, each of which yields success with probability $p$. |
| binomial expansion **** | A binomial expansion describes the algebraic expansion of powers of a binomial expression. |
| binomial random variable **** | A binomial random variable $X$ represents the number of successes in $n$ independent Bernoulli trials. In each Bernoulli trial, the probability of success is $p$ and the probability of failure is: $q=1-p$ |
| combination  | A combination is a selection of $r$ distinct objects from $n$ distinct objects, where order is not important. The number of such combinations is denoted by$ ^{n}C\_{r}$ or $ \left(\begin{matrix}n\\r\end{matrix}\right)$, and is given by: $ \frac{n!}{r!(n-r)!}$ |
| factorial | The product of the first $n$positive integers is called the factorial of $n$ and is denoted by $n!$. $n!=n\left(n-1\right)\left(n-2\right)\left(n-3\right)×…×3×2×1$By definition: $0!=1$ |
| fundamental counting principle | The fundamental counting principle states that if one event has $m$ possible outcomes and a second independent event has $n$ possible outcomes, then there are a total of $m×n$ possible outcomes for the two combined events. |
| permutation  | A permutation is an arrangement of $r$ distinct objects taken from $n$ distinct objects where order is important.The number of such permutations is denoted by$ ^{n}P\_{r}$ and is equal to: $ ^{n}P\_{r}=n(n-1)...(n-r+1)=\frac{n!}{(n-r)!}$The number of permutations of $n$ objects is $n!$.  |

| **Lesson sequence** | **Content** | **Suggested teaching strategies and resources**  | **Date and initial** | **Comments, feedback, additional resources used** |
| --- | --- | --- | --- | --- |
| Introduction to Bernoulli random variables(2 lessons) | **S1.1: Bernoulli and binomial distributions*** use a Bernoulli random variable as a model for two-outcome situations (ACMMM143)
	+ identify contexts suitable for modelling by Bernoulli random variables (ACMMM144)
* use Bernoulli random variables and their associated probabilities to solve practical problems (ACMMM146) **AAM**
	+ understand and apply the formulae for the mean, $E\left(X\right)=\overbar{x}=p$, and variance,$Var\left(X\right)=p(1-p)$, of the Bernoulli distribution with parameter $p$, and $X$ defined as the number of successes (ACMMM145)
 | **Introducing Bernoulli random variables*** A Bernoulli trial is an experiment with only two possible outcomes, labelled ‘success’ and ‘failure’.
* Consider scenarios that can be represented as two-outcome situations. Students should consider the probability of each of the two outcomes, and ensure the sum of the outcomes is 1. In each case, one outcome is identified as a ‘success’ denoted as $p$ and the other outcome as a ‘failure’ (or ‘not success’) and denoted as $q$. i.e. $q=1-p$.
* **Note:** ‘Success’ is the occurrence of ‘the event of interest’, and is not necessarily a positive attribute. e.g. The number of defective earbuds from a sample of 200, where a ‘success’ is a defective earbud being found.
* Consider the scenario of a single experiment, or trial, that has only two possible outcomes, a success or a failure, and the probability of a success is $p$. If the random variable $X$ is defined so that it takes the value 1 when the experiment results in a success and a 0 if the experiment results in a failure, then the variable is described as a **Bernoulli variable** and is said to have a Bernoulli distribution.

$X$ ~ Bernoulli ($p$) if $$X=\left\{\begin{array}{c}0, \&1-p\\1, \&p \end{array}\right.$$* Explore an example that has a specific value for $p$, and calculate the mean and variance of a Bernoulli distribution.

**Example:** Rolling a 6 on an unbiased die:Probability of success: $p =\frac{1}{6}$ , Probability of failure: $q=\frac{5}{6}$$X$ ~ Bernoulli ($p$) if $$X=\left\{\begin{array}{c}0, \&\frac{5}{6} \\1, \&\frac{1}{6} \end{array}\right.$$**Mean or expected value,** $E\left(X\right)$$E\left(X\right)=0×\frac{5}{6}+1×\frac{1}{6}$ $$E(X)=\frac{1}{6}$$ **Variance,** $Var\left(X\right)$$Var\left(X\right)=E(X^{2})-\left[E\left(X\right)\right]^{2}$ $Var\left(X\right)=\left(0^{2}×\frac{5}{6}+1^{2}×\frac{1}{6}\right)-\left(\frac{1}{6}\right)^{2}$ $$Var\left(X\right)=\frac{1}{6}-\left(\frac{1}{6}\right)^{2}$$$$Var\left(X\right)=\frac{1}{6}\left(1-\frac{1}{6}\right)$$$$Var\left(X\right)=\frac{1}{6}\left(\frac{5}{6}\right)$$$$Var\left(X\right)=\frac{5}{36}$$**Standard deviation,** $σ=\sqrt{Var\left(X\right)}$$$σ=\sqrt{\frac{5}{36}}$$$$σ=\frac{\sqrt{5}}{6}$$* Generalise the results for mean and varianceas $E\left(X\right)=\overbar{x}=p$, and $Var\left(X\right)=p\left(1-p\right)$, for the Bernoulli distribution with parameter $p$, and $X$ defined as the number of successes. They also note that $Var\left(X\right)=pq$ where $q=1-p$.

**Proof**$X$ ~ Bernoulli ($p$) if $$X=\left\{\begin{array}{c}0, \&1-p \\1, \&p\end{array}\right.$$**Mean or expected value,** $E\left(X\right)$$E\left(X\right)=0×(1-p)+1×p$ $$E(X)=p$$ **Variance,** $Var\left(X\right)$$Var\left(X\right)=E(X^{2})-\left[E\left(X\right)\right]^{2}$ $Var\left(X\right)=\left(0^{2}×(1-p)+1^{2}×p\right)-\left(p\right)^{2}$ $$Var\left(X\right)=p-\left(p\right)^{2}$$$$Var\left(X\right)=p\left(1-p\right)$$$$Var\left(X\right)=pq$$where $q=1-p$* Consider a variety of Bernoulli random variables, and calculate the mean and variance for each distribution.
 |  |  |
| Trials, experiments and random variables(1 lesson) | * understand the concepts of Bernoulli trials and the concept of a binomial random variable as the number of ‘successes’ in $n$ independent Bernoulli trials, with the same probability of success $p$ in each trial (ACMMM147)
	+ calculate the expected frequencies of the various possible outcomes from a series of Bernoulli trials
* use binomial distributions and their associated probabilities to solve practical problems (ACMMM150) **AAM** Critical and creative thinking icon
	+ identify contexts suitable for modelling by binomial random variables (ACMMM148)
	+ identify the binomial parameter $p$ as the probability of success
	+ understand and use the notation X∼Bin(*n,p*) to indicate that the random variable X is distributed binomially with parameters *n* and *p*
 | **Linking Bernoulli trials to binomial random variables** * Explore the link between Bernoulli trials and the binomial distribution. The clip [Bernoulli, Binomial and Poisson Random Variables](https://www.youtube.com/watch?v=mOkvzPkWBTc) (duration 25:11) may be useful for making this connection.
* Define a binomial random variable: A binomial random variable$ X$ represents the number of successes in$ n$ independent Bernoulli trials. In each Bernoulli trial, the probability of success is $p $and the probability of failure is: $q=1-p$

**Binomial experiment**Students to conduct a binomial experiment with the class. **Resource:** conducting-a-binomial-experiment.DOCX (Part A)* Formalise definition of a binomial experiment as one for which each of the following must apply:
	+ there are $n$ repeated trials
	+ all trials are independent
	+ there are only two outcomes, a ‘success’ which has a probability of $p$ and a ‘failure’ which has a probability $q=1-p$
	+ the value of $p$ remains unchanged for each trial.
* Introduce notation $X\~Bin\left(n,p\right)$ to indicate that the random variable $X$ is distributed binomially, that $X$ represents the number of successes in $n$ trials where the probability of a success in each trial is $p$. The teacher may also discuss alternate notations for the binomial distribution. **Resource:** identifying-binomial-distributions.DOCX
 |  |  |
| Binomial probabilities(2-3 lessons) | * + apply the formulae for probabilities $P\left(X=r\right)= ^{n}C\_{r}p^{r}(1-p)^{n-r}$ associated with the binomial distribution with parameters $n$ and $p$ and understand the meaning of $ ^{n}C\_{r}$ as the number of ways in which an outcome with $r$ successes can occur  Information and communication technology capability icon
	+ understand and apply the formulae for the mean, $E\left(X\right)=\overbar{x}=np$, and the variance, $Var\left(X\right)=np(1-p)$, of a binomial distribution with parameters $n$ and $p$
 | **Binomial probability*** The teacher explains that the notation $P\left(X=r\right)$ in the context of a binomial experiment is used to denote the probability that there are $r$ successes in $n$ trials.
* Student investigation: With teacher guidance as necessary, students link the expressions for the probabilities with the terms of the expansion of $\left(p+q\right)^{n}$. This result is then generalised to $P\left(X=r\right)= ^{n}C\_{r}p^{r}(1-p)^{n-r}$.

**Resource:** discovering-binomial-probability.DOCX* Complementary relationships:

Pose the question: If we rolled a die 50 times, what is the probability of getting at least 3 sixes?Students discuss the meaning of the notations $P\left(X\leq r\right)$ and $P\left(X>r\right)$ and establish their complementary relationship.* Calculate different probabilities given a series of problems that involve binomial distributions.**Resource:** binomial-probability-problems.DOCX (Part A)

**Modelling binomial probability*** Students create and answer a series of binomial probability questions using the binomial data collected earlier in the unit (see, ‘Binomial experiment’)**Resource:** conducting-a-binomial-experiment.DOCX (Part B)
 |  |  |
|  |  | **Mean and variance of a binomial distribution*** Generalise the results for mean and variancein a Bernoulli experiment with number of trials $n$.

**Proof**If $X$ ~$Bin(n,p$), **Mean or expected value,** $E\left(X\right)$Given $E(X)=p$ for each trial, if $n$ trials are conducted, $E(X)=np$**Variance,** $Var\left(X\right)$Given $Var\left(X\right)=pq$ for each trial, if $n$ trials are conducted, $Var(X)=npq$**Teacher resource:** For an in depth explanation see the Khan academy videos on [E(X)](https://www.khanacademy.org/math/statistics-probability/random-variables-stats-library/binomial-mean-standard-dev-formulas/v/expected-value-of-binomial-variable) and [Var(X)](https://www.khanacademy.org/math/statistics-probability/random-variables-stats-library/binomial-mean-standard-dev-formulas/v/variance-of-binomial-variable).* Calculate different probabilities, means and variances given a series of problems that involve binomial and Bernoulli distributions. **Resources:** binomial-probability-problems.DOCX (Part B), conducting-a**-**binomial-experiment.DOCX (Part C)
 |  |  |
| Normal approximation for the sample proportion(3-4 lessons) | **S1.2: Normal approximation for the sample proportion*** use appropriate graphs to explore the behaviour of the sample proportion on collected or supplied data **AAM**
	+ understand the concept of the sample proportion $\hat{p}$ as a random variable whose value varies between samples (ACMMM174)
* explore the behaviour of the sample proportion using simulated data **AAM**
	+ examine the approximate normality of the distribution of $\hat{p}$ for large samples (ACMMM175)
* understand and use the normal approximation to the distribution of the sample proportion and its limitations **AAM**
 | **Investigating normal approximations for a sample*** Explain the meaning of the sample proportion and how it can be calculated.

For example, take a random sample of 20 scores from a set of data and calculate the proportion of scores that are 80 or higher. This will be a fractional value $\frac{k}{20}$, where k is the number of scores that are 80 or higher. This value is called a sample proportion,$ \hat{p}$.Use Microsoft Excel to explore sample proportions. **Resources:** sample-proportions.XLSM, refer to teacher notes: sample-proportion-teacher-notes.DOCX* Students can explore and reinforce their discoveries by watching:
* Estimating population proportions using  [sample proportions](https://www.youtube.com/watch?v=MWff4_ORZjM) (duration 3.08)
	+ [The sampling distribution of the sample proportion](https://www.youtube.com/watch?v=fuGwbG9_W1c) (duration 9:48)
* Key questions to consider when viewing:
	+ What values can $\hat{p}$ take in this specific experiment?
	+ Are the values of $\hat{p}$ different between samples? If so, why?
	+ What is the relationship between p, the proportion in the population, and $\hat{p}$, the sample proportion?
	+ What are the limitations of the sample proportion that may affect its use in solving problems?
* Teacher explanation of key points:
	+ $\hat{p}$ takes a value between 0 and 1, as a fraction $\frac{k}{n}$ where k is the number of scores satisfying the condition in the sample and n is the sample size.
	+ It will vary due to the random nature of the sample.
	+ $\hat{p}$ is an estimator of p

If a random sample size $n $is large and $p$ is not too close to 0 or 1 (if $np\geq 10$ and $n\left(1-p\right)\geq 10$) then the distribution of $\hat{p}$ is approximately normal with:$$μ\_{\hat{p}}=p$$$$σ\_{\hat{p}}=\sqrt{\frac{p(1-p)}{n}}$$In general this means that the binomial distribution can reasonably be approximated by a normal distribution.The Khan Academy video, [sampling-distribution-of-sample-proportion part 1](https://www.khanacademy.org/math/ap-statistics/sampling-distribution-ap/sampling-distribution-proportion/v/sampling-distribution-of-sample-proportion-part-1) (duration 9:56), has an in-depth explanation of the mean and standard deviation.Note: Some sources use 15 as the test but NESA’s sample unit uses 10.The teacher can use the Khan Academy, [Candy Sample Distribution](https://www.khanacademy.org/computer-programming/candy-sampling-distribution/5180356611473408) applet to demonstrate the skewed nature of the distribution if the results do not hold.**Student practice*** Students apply the tests, $np\geq 10$ and $n\left(1-p\right)\geq 10$, to determine the appropriate sample size for the binomial situations they explored earlier in the unit (see above).
* Students practise transforming values of $\hat{p}$ into values of the standardised normal variable $z$.
* Students apply their knowledge and skills by solving problems related to the sample proportion and binomial probability. These are based on NESA’s sample questions. **Resource:** sample-proportion-and-binomial-probabiliy-problems.DOCX
 |  |  |

Reflection and evaluation

Please include feedback about the engagement of the students and the difficulty of the content included in this section. You may also refer to the sequencing of the lessons and the placement of the topic within the scope and sequence. All ICT, literacy, numeracy and group activities should be recorded in the ‘Comments, feedback, additional resources used’ section.