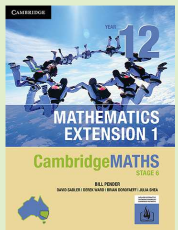
NORMANHURST BOYS HIGH SCHOOL

MATHEMATICS EXTENSION 1
(YEAR 12 COURSE)

Topic summary and exercises:

X1 X2 **Statistics/The Binomial Distribution**

With references to



Name: ....

Initial version by H. Lam, July 2020. Last updated September 30, 2023.
Various corrections by students & members of the Department of Mathematics at Normanhurst Boys High School.
With particular thanks to N. Zuber of Cammeraygal HS for his suggestions to Section 2 on page 27.

Acknowledgements Pictograms in this document are a derivative of the work originally by Freepik at `http://www.flaticon.com`, used under CC BY 2.0.

Symbols used

Beware! Heed warning.

Enrichment - not necessarily in the syllabus, or any formal assessment.

X1 Mathematics Extension 1 content.

X2 Mathematics Extension 2 content.

Literacy: note new word/phrase.

$\mathbb{N}$ the set of natural numbers

$\mathbb{Z}$ the set of integers

$\mathbb{Q}$ the set of rational numbers

$\mathbb{R}$ the set of real numbers

$\forall$ for all

Syllabus outcomes addressed

ME12-5 applies appropriate statistical processes to present, analyse and interpret data

Syllabus subtopics

ME-S1 The Binomial Distribution

Gentle reminder

- For a thorough understanding of the topic, *every* question in this handout is to be completed!
- Additional questions from *CambridgeMATHS Year 12 Extension* (Pender, Sadler, Ward, Dorofaeff, & Shea, 2019) will be completed at the discretion of your teacher.
- Remember to copy the question into your exercise book!

Contents

Section 1

The Binomial Distribution

1.1 Bernoulli trials

Learning Goal(s)

Knowledge
Bernoulli random variable

Skills
Discrete probability distributions

Understanding
Single experiment

By the end of this section am I able to:

33.1 Use a Bernoulli random variable as a model for two-outcome situations

33.2 Use Bernoulli random variables and their associated probabilities to solve practical problems

33.3 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

Definition 1

A *Bernoulli trial/experiment* is a single stage random experiment with ....two.... outcomes:

- "success" or "failure"
- "yes" or "no" etc

Laws/Results

- Probability of "success" usually denoted with lowercase p.
- Probability of "failure" usually denoted with lowercase $q = 1 - p$.

Definition 2

If X is a Bernoulli random variable which counts the number of successes in a single trial, then

$$p_X(X) = P(X = x) = \begin{cases} p & x = 1 \\ 1 - p & x = 0 \end{cases}$$

with $0 < p < 1$. Alternatively, as a discrete probability distribution table:

x	0	1
$P(X = x)$		

1.1.1 Mean and standard deviation

Laws/Results

For a Bernoulli random variable X:

- The **mean** (expected value):

$$\begin{aligned} E(X) &= \sum x_i p_i \\ &= 0(1-p) + 1(p) \\ &= p \end{aligned}$$

- The **variance**

$$\begin{aligned} \text{Var}(X) &= \sum {x_i}^2 p_i - \mu^2 \\ &= 0^2(1-p) + 1^2 p - p^2 \\ &= p - p^2 \\ &= p(1-p) \end{aligned}$$

History

Photo: Wikipedia

Jacob Bernoulli (1654-1705) was born in 1654 in Basel, Switzerland.

In 1683 Bernoulli discovered the constant e by studying a question about compound interest which required him to find the value of the following expression (which is in fact e):

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

Example 1

(Haese, Haese, & Humphries, 2016, p.205) A pen is randomly selected from a box containing 7 red pens and 3 blue pens. Let $X = 1$ if the selected pen is red, and $X = 0$ if the selected pen is blue.

(a) Explain why X is a Bernoulli random variable.

(b) Find the mean and standard deviation of X.

Answer: $E(X) = \frac{7}{10}$, $\sigma \approx 0.458$

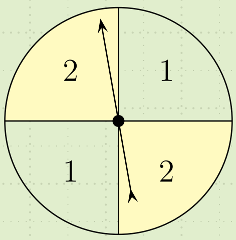
Example 2

[Ex 7F Q1] (Haese et al., 2016) Decide whether X is a Bernoulli random variable:

(a) A coin is tossed. $X = 1$ if the result is heads, $X = 0$ if the result is tails.

(b) Two coin are tossed. $X =$ the number of heads tossed.

(c) $X =$ the result when the following spinner is spun:



(d) A person enters a store. $X = 1$ if the person purchases something, $X = 0$ otherwise.

(e) A letter of the alphabet is chosen at random. $X =$ the number of vowels chosen.

Example 3

[Ex 7F Q3] (Haese et al., 2016) A regular six-sided die is rolled once. Let X represent the number of sixes rolled.

(a) Explain why X is a Bernoulli random variable.

(b) Find:

i. the mean of X.

ii. the variance of X.

Answer: $\mu = \frac{1}{6}$, $\sigma^2 = \frac{5}{36}$

1.2 Binomial distributions

Learning Goal(s)

Knowledge
Multiple Bernoulli trials results in Binomial Distribution

Skills
Discrete probability distributions, using random variables that follow the binomial distribution

Understanding
Situations arising with a random variable following the Binomial distribution

By the end of this section am I able to:

33.3 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

33.4 Use binomial distributions and their associated probabilities to solve practical problems

Definition 3

A *binomial experiment* occurs by repeating a Bernoulli experiment multiple times.

Laws/Results

If X is the *random variable* of a binomial experiment, then

$$X \sim \text{Bin}(n, p)$$

where

- n is the number of times the Bernoulli experiment is repeated
- p is the probability of success of one of such Bernoulli experiments.

Example 4

A regular six sided die is tossed 3 times.

(a) What is the probability of 0, 1, 2, or 3 sixes turning up?

(b) What is the expected number of 'sixes' to appear in the three tosses?

Steps

1. Let the random variable X be the number of '6' that turn up in three tosses, i.e.

$$X \sim \text{Bin}\left(3, \tfrac{1}{6}\right)$$

with $p = \frac{1}{6}$ is the probability of a six turning up in a single toss.

2. Draw and populate a probability tree.

3. Sum probabilities:

4. Hence

x	0	1	2	3
$P(X = x)$				

Important note

Draw picture! (Probability tree!)

Definition 4

Binomial probability distribution function In an n-stage Bernoulli experiment with k successes, and p being the probability of success of one of such experiments,

$$P(k \text{ successes}) = P(X = k) = {}^nC_k p^k q^{n-k}$$

- $P(X = k)$ is the ......binomial...... ......probability...... ......distribution...... function.
- The term in $p^k q^{n-k}$ is one of the terms in

$$\underline{(p+q)^n} = 1^n = 1$$

where $q = 1 - p$.

Laws/Results

For a Binomial random variable X:

- The **mean** (expected value):

$$E(X) = \mu = np \tag{1.1}$$

- The **variance**

$$\text{Var}(X) = \sigma^2 = np(1-p) \tag{1.2}$$

Example 5

[2022 Ext 1 HSC Q12] A game consists of randomly selecting 4 balls from a bag. After each ball is selected it is replaced in the bag. The bag contains 3 red balls and 7 green balls. For each red ball selected, 10 points are earned and for each green ball selected, 5 points are deducted. For instance, if a player picks 3 red balls and 1 green ball, the score will be $3 \times 10 - 1 \times 5 = 25$ points.

What is the expected score in the game? **Answer:** -2

Example 6

[2008 VCE Mathematical Methods Paper 2 Q5] Let X be a discrete random variable with a binomial distribution. The mean of X is 1.2 and the variance of X is 0.72.
The values of n (the number of independent trials) and p (the probability of success in each trial) are

(A) $n = 4$, $p = 0.3$ (C) $n = 2$, $p = 0.6$ (E) $n = 3$, $p = 0.4$

(B) $n = 3$, $p = 0.6$ (D) $n = 2$, $p = 0.4$

Example 7

[2015 VCE Mathematical Methods Paper 2 Q9] The binomial random variable, X, has $E(X) = 2$ and $\text{Var}(X) = \dfrac{4}{3}$. $P(X = 1)$ is equal to

(A) $\left(\dfrac{1}{3}\right)^6$ (C) $\dfrac{1}{3} \times \left(\dfrac{2}{3}\right)^2$ (E) $6 \times \dfrac{2}{3} \times \left(\dfrac{1}{3}\right)^5$

(B) $\left(\dfrac{2}{3}\right)^6$ (D) $6 \times \dfrac{1}{3} \times \left(\dfrac{2}{3}\right)^5$

Example 8

A marksman, who keeps a record of his performances, finds that in the long run, he scores a bull's eye on 3 out of 4 occasions. He fires 5 rounds at a target. Assuming that each trial is independent of any other similar trial, find the probability of

(a) exactly 3 bull's eyes,

(b) at least 4 bull's eyes,

(c) a bull's eye in the second round only.

Answer: (a) $\frac{135}{512}$ (b) $\frac{81}{128}$ (c) $\frac{3}{1\,024}$

Example 9

[**1995 3U HSC**] In a Jackpot Lottery, 1 500 numbers are drawn from a barrel containing the 100 000 ticket numbers available.

After all the 1 500 prize-winning numbers are drawn, they are returned to the barrel and a jackpot number is drawn. If the jackpot number is the same as one of the 1 500 numbers that have already been selected, then the additional jackpot prize is won.

The probability that the jackpot prize is won in a given game is thus

$$p = \frac{1\,500}{100\,000} = 0.015$$

(i) Calculate the probability that the jackpot prize will be won *exactly* once in 10 independent lottery games. **1**

(ii) Calculate the probability that the jackpot prize will be won *at least* once in 10 independent lottery games. **2**

(iii) The jackpot prize is initially \$8 000, and it increases by \$8 000 each time the prize is NOT won. **3**

Calculate the probability that the jackpot prize will exceed \$200 000 when it is finally won.

Example 10

[**2009 Ext 1 HSC**] A test consists of five multiple-choice questions. Each question has four alternative answers. For each question only one of the alternative answers is correct.

Huong randomly selects an answer to each of the five questions.

(i) What is the probability that Huong selects three correct and two incorrect answers? **2**

(ii) What is the probability that Huong selects three or more correct answers? **2**

(iii) What is the probability that Huong selects at least one incorrect answer? **1**

Example 11

[**2018 VCE Mathematical Methods NHT Paper 2 Q2**] *Rebecca's Robotics* manufactures three types of components for robots: sensors, motors and controllers. The manufacturing processes for each type of component are independent.

It is known that 8% of all of the sensors manufactured are defective.

A random sample of five sensors is selected. Find, correct to four decimal places, the probability that

i. Exactly two of these selected sensors are defective. **2**

ii. Exactly two of these selected sensors are defective, given that at most two sensors in the sample are defective. **2**

Answer: i. 0.0498 ii. 0.0501

Additional questions

1. **[2019 VCE Mathematical Methods NHT Paper 1 Q8]** A fair standard die is rolled 50 times. Let W be a random variable with binomial distribution that represents the number of times the face with a six on it appears uppermost.

(a) Write down the expression for $P(W = k)$, where $k \in \{0, 1, 2, \cdots, 50\}$. **1**

(b) Show that $\dfrac{P(W = k+1)}{P(W = k)} = \dfrac{50-k}{5(k+1)}$. **2**

(c) Hence or otherwise, find the value of k for which $P(W = k)$ is the greatest. **2**

2. A light bulb is classed as 'defective' if it burns out in under 1 000 hours. A company making light bulbs finds, after careful testing, that 1% of its bulbs are defective. If it packs its bulbs in boxes of 50, find, correct to three significant figures:

(a) the probability that a box will contain no defective bulbs,

(b) the probability that at least two bulbs in a box are defective.

3. **[2004 Ext 1 HSC]** Katie is one of ten members of a social club. Each week one member is selected at random to win a prize.

i. What is the probability that in the first 7 weeks Katie will win at least 1 prize? **1**

ii. Show that in the first 20 weeks Katie has a greater chance of winning exactly 2 prizes than of winning exactly 1 prize. **2**

iii. For how many weeks must Katie participate in the prize drawing so that she has a greater chance of winning exactly 3 prizes than of winning exactly 2 prizes? **2**

Answers

1. (a) ${}^{50}C_k p^k q^{50-k}$ (b) Show (c) $k = 7$ **2.** (a) 0.605 (b) 0.0894 **3.** i. 0.5217 ii. Show $P(X = 2) > P(X = 1)$ iii. $n = 30$

Further exercises

Ex 17A	Ex 17B
• Q1-20	• Q1-6

1.3 Normal approximation to the binomial distribution

1.3.1 Why use a normal approximation to the binomial distribution?

Theorem 1

Normal approximation to the binomial distribution The normal distribution's probability density function approximates the frequency polygon of the binomial distribution when n is sufficiently large.

GeoGebra

Use this GeoGebra applet to verify the probability density for a binomial random variable $X \sim \text{Bin}(n, p)$ such that:

- $n = 1$, $p = 0.5$

  Observations:

  ....

- $n = 5$, $p = 0.5$

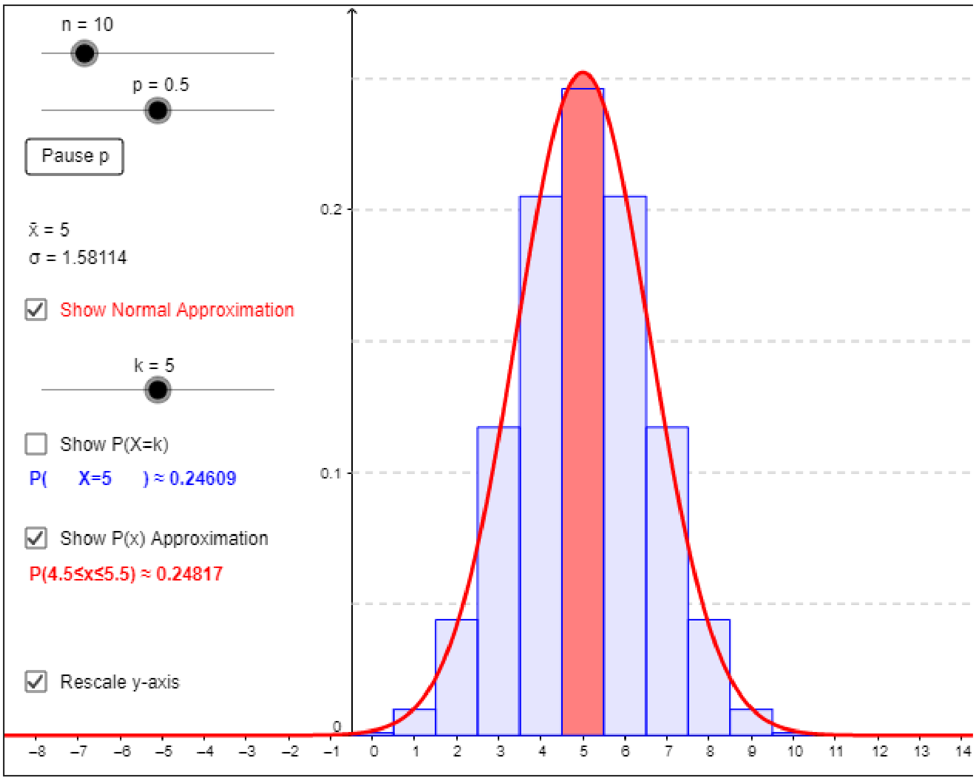
  Observations:

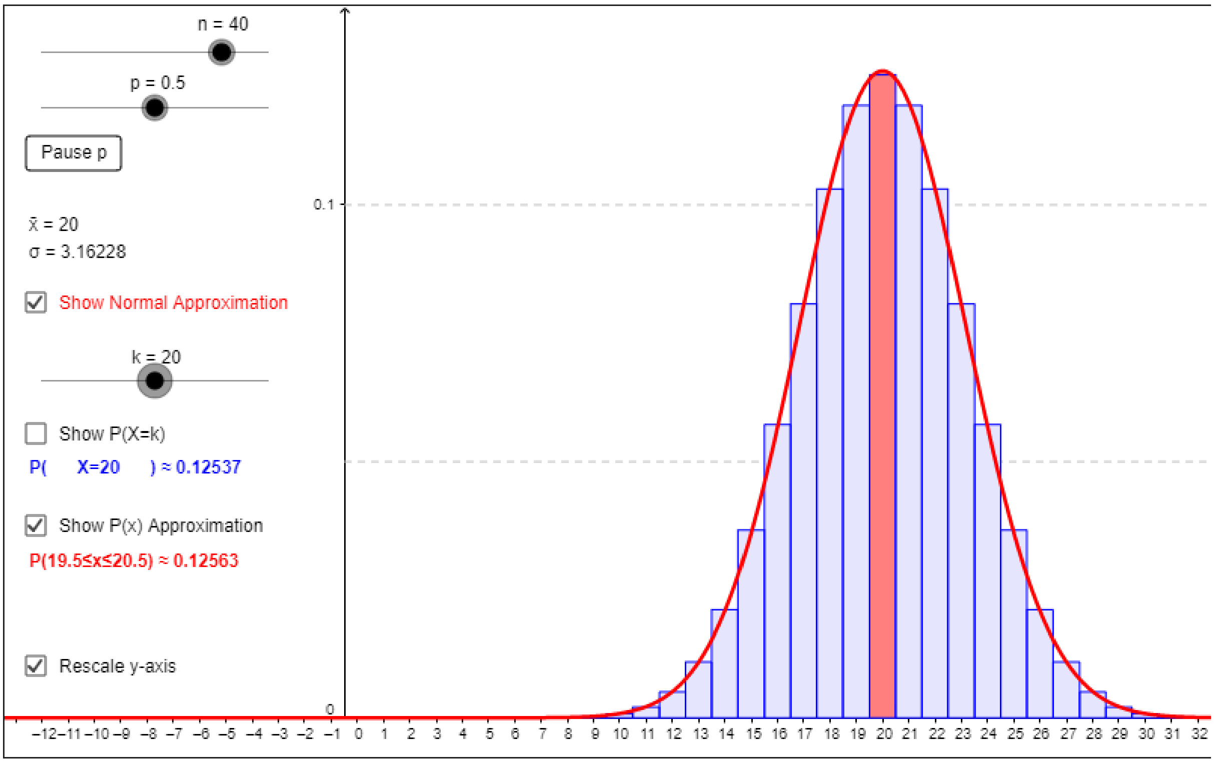
  ....

- n as large as possible

  Observations:

  ....


n = 10
p = 0.5
Pause p
x̄ = 5
σ = 1.58114
Show Normal Approximation
k = 5
Show P(X=k)
P(X=5) ≈ 0.24609
Show P(x) Approximation
P(4.5≤x≤5.5) ≈ 0.24817
Rescale y-axis


n = 40
p = 0.5
Pause p
x̄ = 20
σ = 3.16228
Show Normal Approximation
k = 20
Show P(X=k)
P(X=20) ≈ 0.12537
Show P(x) Approximation
P(19.5≤x≤20.5) ≈ 0.12563
Rescale y-axis

Laws/Results

If $X \sim \text{Bin}(n, p)$,

Mean $\mu = np$ (See Equation 1.1 on page 10)

Variance $\sigma^2 = np(1-p)$ (See Equation 1.2 on page 10)

Its *normal distribution approximation* will be

$$\begin{aligned} X &\sim N\left(\mu, \sigma^2\right) \\ &\sim N\left(np, np(1-p)\right) \end{aligned}$$

1.3.2 When the normal approximation may be inappropriate

Important note

Avoid using the normal approximation to the Binomial distribution unless n gets large such that

$$\begin{cases} np > 10 \\ nq > 10 \end{cases}$$

and $q = 1 - p$. Other sources may use $np > 5$ and $nq > 5$.

Example 12

[2023 JRAHS Ext 1 Trial Q15] It is known that the probability of obtaining a head when tossing a particular biased coin is 0.2. A random 100 flips of this biased coin is conducted and the number of heads are observed.

i. Justify why the proportions of heads can be approximated using the normal distribution. **1**

ii. Determine the mean and standard deviation of the proportion of heads. **1**

iii. Hence or otherwise, approximate the probability that the proportion of heads obtained will be between 0.1 and 0.3 inclusive, to 4 decimal places. **2**

Answer: i. Check np and nq. ii. $\mu = 0.2$, $\sigma = 0.04$ iii. 0.9876

Example 13

[**2023 JRAHS Ext 1 Trial Q14**] In a large pool of people, it is known that 10% of them take less than 12 minutes to complete a particular test.

i. 15 people are selected at random. What is the probability that fewer than 2 people will take less than 12 minutes to complete the test, correct to 3 decimal places? **2**

ii. A random sample of n people is taken, and the probability that fewer than 9 of these n people will take less than 12 minutes to complete the test is 0.3446, correct to 4 decimal places. **4**

Use normal approximation to solve for n, correct to the nearest integer.

Answer: i. 0.549 ii. 102.13

Steps

ii. ...Use normal approximation to solve for n, correct to the nearest integer. **4**

1. Write down a short statement about the random variable involved:

Let Y be the random variable for the number of people who take less than 12 minutes to complete the test.

2. Write down the distribution which Y follows:

$$Y \sim \text{Bin}\left(n, 0.1\right)$$

3. It is known that $P(Y < 9) \approx 0.3446$. However,

$$P(Y<9) = P(Y=0) + P(Y=1) + P(Y=2) + \cdots + P(Y=7) + P(Y=8)$$

which will be too tedious to find. Hence, the normal approximation is required. Rewriting Y's distribution expression:

$$Y \sim \text{N}\left(\mu, \sigma\right)$$

- However, μ and σ are not yet known.

$$\mu = E(Y) \qquad \text{Var}(Y) = np(1-p)$$

$$= np \qquad = 0.09n$$

$$= 0.1n \qquad \therefore \sigma = 0.03\sqrt{n}$$

Hence,

$$Y \sim \text{N}\left(0.1n, 0.3\sqrt{n}\right)$$

4. Rewrite the probability expression, changing the Y random variable into the standard normal random variable to utilise the approximation:

$$P(Y<9) = P\left(\frac{Y-0.1n}{0.3\sqrt{n}} < \frac{9-0.1n}{0.3\sqrt{n}}\right)$$

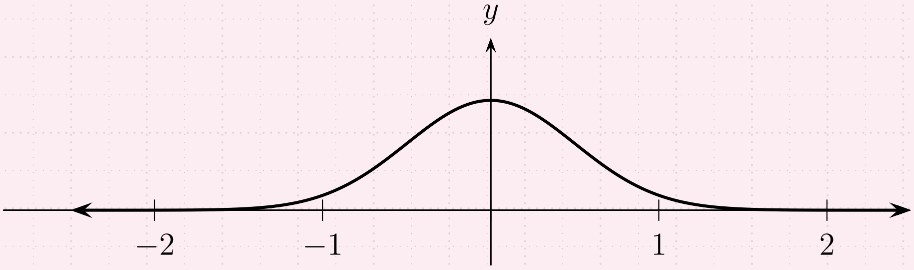
$$= P\left(Z < \frac{9-0.1n}{0.3\sqrt{n}}\right)$$

$$= 0.3446$$

Steps

ii. ...Use normal approximation to solve for n, correct to the nearest integer. **4**

5. On a standard normal curve, find the corresponding value of z such that $P\left(Z < \dfrac{9-0.1n}{0.3\sqrt{n}}\right) = 0.3446$:



From Table 1.1 on page 26: $z = 0.4$ will result in $P(Z < 0.4) = 0.6554$. Hence by symmetry,

$$P\left(Z < \frac{9-0.1n}{0.3\sqrt{n}}\right) = 1 - P\left(Z < 0.4\right) = 0.3446$$

occurs when

$$\frac{9-0.1n}{0.3\sqrt{n}} = -0.4$$

6. Solve for n:

7. However, $n = 79.31$ and $n = 102.13$ both seem to be valid. Check for the correct value by subtituting into $\dfrac{9-0.1n}{0.3\sqrt{n}}$:

Additional questions

(Use Table 1.1 on page 26 to answer these questions)

1. **[2022 Ext 1 HSC Q14]** An airline company that has empty seats on a flight is not maximising its profit. **4**

An airline company has found that there is a probability of 5% that a passenger books a flight but misses it. The management of the airline company decides to allow for overbooking, which means selling more tickets than the number of seats available on each flight.

To protect their reputation, management makes the decision that no more than 1% of their flights should have more passengers showing up for the flight than available seats.

Given management's decision and using a suitable approximation, find the maximum number of tickets that can be sold for a flight which has 350 seats.

Answer: $n = 358$

1.3.3 Continuity correction

Important note

- For small values of n, apply the *continuity correction* and integrate between half intervals corresponding to the boundaries of the cdf. Check the wording of the question to see whether this is necessary or not.
- Not explicitly specified in the syllabus, but can be handy to use in some cases.

- To approximate $P(X = 8, 9, 10, 11, 12)$, treat X as a continuous normal variable and find
$$P(7.5 \le X \le 12.5)$$

Example 14

[2022 Ext 1 HSC Q13] (Use Table 1.1 on page 26 to answer this question)

A chocolate factory sells 150-gram chocolate bars. There has been a complaint that the bars actually weigh less than 150 grams, so a team of inspectors was sent to the factory to check. They randomly selected 16 bars, weighed them and noted that 8 bars weighed less than 150 grams. The factory manager claims 80% of the chocolate bars produced by the factory weigh 150 grams or more.

i. The inspectors used the normal approximation to the binomial distribution to calculate the probability, $\mathcal{P}$, of having at least 8 bars weighing less than 150 grams in a random sample of 16, assuming the factory manager's claim is correct. **2**

Calculate the value of $\mathcal{P}$.

ii. The factory manager disagrees with the method used by the inspectors as described in part (i). **1**

Explain why the method used by the inspectors might not be valid.

Answer: 0.15% (0.36% if continuity correction is applied)

Further exercises

Ex 17C

- Q1-9
- Q10 onwards

	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Table 1.1 – Standard normal distribution values - $Z \sim N(0,1),\ P(Z < a) = \displaystyle\int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}\, dz$

Section 2

Sample proportions

Learning Goal(s)

Knowledge
Graphs in generalising behaviour of sample proportions, and how the normal distribution can approximate the binomial distribution as number of experiments increases

Skills
Calculate expected value and variance

Understanding
Normal distribution to approximate sample proportion random variables

By the end of this section am I able to:

33.5 Use appropriate graphs to explore the behaviour of the sample proportion on collected or supplied data

33.6 Explore the behaviour of the sample proportion using simulated data and its limitations

Important note

Motivation: survey only select groups to make ....predictions........ about the population.

- A *census* is when ....every.... person is asked.
- A *survey* is taken when only a selection of people are asked. Economically much more feasible than a census.

2.1 Actual observations and sample proportions

Definition 5

A *sample* is a ....collection........ of successes of Bernoulli trials

Definition 6

A *sample proportion* is the number of successes in the Bernoulli trials within that sample, divided by the total number surveyed.

Definition 7

The *sample proportion* ....random........ ....variable........ $\widehat{p}$, takes on values from 0 to 1 inclusive.

Important note

The Australian Curriculum uses $\widehat{P}$ as the random variable symbol.

Definition 8

The observed value of $\widehat{p}$, is the *point estimate* of p.

Example 15

[AMSI ESA] (See AMSI ESA website) Suppose a random sample of 500 voters is obtained, and 227 prefer Labor.

Fill in the spaces

- p is the proportion of all Australian voters who prefer Labor.
- Sample size: ......$n = 500$......
- The ......random...... ......variable...... X is the *number* of voters who prefer Labor in a random sample of 500 voters.
- The ......random...... ......variable...... $\widehat{p} = \dfrac{X}{500}$ is the ......proportion...... of voters who prefer Labor in a random sample of 500 voters.
- ......$x = 227$...... is an ......observation...... of X.
- ......$\dfrac{227}{500} \approx 0.454\cdots$...... is the corresponding observation of $\widehat{p}$.

Watch multimedia

Year 12 Methods: Sample Proportions (Chris Simpson, Aranmore Catholic College, WA)

- In NSW, ignore *sample mean*.
- 'Spec' implies Specialist Mathematics, 'Methods' implies Mathematical Methods.
- *Central Limit Theorem* not in NSW Syllabus, but in fact serves as a bridge to understanding this content.

Fill in the spaces

- $n = 1000$ in the survey.
- Let L be the random variable for the number of left handers, and $\widehat{p}$ be the random variable for the sample proportion of left handers.

$$\widehat{p} = \frac{L}{1000}$$

- The 50 samples of 20 trials each, the mean position of all samples will approximate the actual population proportion.
  - The sample proportion will approximate the population proportion.
- Ensure sampling is not biased.

Laws/Results

As the number of samples taken increases, the symmetry of the sample proportion distributions increases and the spread of the distributions decreases.
(See AMSI ESA website)

2.2 Sample proportion random variable

Theorem 2

If $X \sim \text{Bin}(n, p)$ is a binomial random variable consisting of n independent Bernoulli trials, each with a probability of p of success, the **sample proportion** (random variable) $\widehat{p}$ is the proportion of successes when the experiment is run, such that

$$\widehat{p} = \frac{X}{n}$$

i.e. a horizontal stretch factor of the corresponding distribution by $\dfrac{1}{n}$ (compressed inward by factor of n)

Important note

Keyword: sample proportions. Look for a 'hat' in the random variable.

Important note

- $\widehat{p}$ is a *compressed* version of a corresponding random variable X.
- If the random variable X is the number of '6's showing up in 300 tosses, then the result $X = 15$ will correspond to

$$\widehat{p} = \frac{15}{300} = \frac{1}{20}$$

- The probability of each value of $\widehat{p}$ has the same probability as the corresponding value of X.
  - For a sample proportion random variable with corresponding random variable X and 300 tosses,

$$\begin{aligned} P(\widehat{p} \geq 0.25) &= P\left(\frac{X}{300} \geq 0.25\right) \\ &= P(X \geq 75) \end{aligned}$$

2.2.1 Expected value and variance

Laws/Results

If $X \sim \text{Bin}(n, p)$,

Mean $\mu = np$ (See Equation 1.1 on page 10)

Variance $\sigma^2 = np(1-p)$ (See Equation 1.2 on page 10)

Its *sample proportion mean* and *sample proportion variance*

Sample proportion mean $E(\widehat{p}) = E\left(\frac{X}{n}\right)$

$$\begin{aligned} E(\widehat{p}) &= E\left(\frac{X}{n}\right) \\ &= \frac{E(X)}{n} \\ &= \frac{np}{n} \\ &= p \end{aligned}$$

Sample proportion variance $\text{Var}(\widehat{p}) = \text{Var}\left(\frac{X}{n}\right)$

$$\begin{aligned} \text{Var}(\widehat{p}) &= \text{Var}\left(\frac{X}{n}\right) \\ &= \frac{\text{Var}(X)}{n^2} \\ &= \frac{np(1-p)}{n^2} \\ &= \frac{p(1-p)}{n} \end{aligned}$$

Example 16

[2017 VCE Mathematical Methods NHT Paper 1 Q6] At a large sporting arena there are a number of food outlets, including a cafe. Let $\widehat{P}$ represent the sample proportion of men rostered to work on a particular day.

(a) The cafe employs five men and four women. Four of these people are rostered at random to work each day.

i. List the possible values that $\widehat{P}$ can take. **1**

ii. Find $P\left(\widehat{P}=0\right)$. **1**

⚠ Do not look for a binomial expansion!

(b) There are over 80 000 spectators at a sporting match at the arena. Five in nine of these spectators support the Goannas team. A simple random sample of 2 000 spectators is selected. **1**

What is the standard deviation of the distribution of $\widehat{P}$, the sample proportion of spectators who support the Goannas team?

Answer: (a) i. $0, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, 1$ ii. $\frac{1}{126}$ (b) $\frac{1}{90}$

Example 17

[2017 VCE Mathematical Methods Paper 2 Q16] For random samples of five Australians, $\widehat{P}$ is the random variable that represents the proportion who live in a capital city.

Given that $P\left(\widehat{P}=0\right)=\dfrac{1}{243}$, then $P\left(\widehat{P}>0.6\right)$, correct to 4 decimal places is

(A) 0.0453 (C) 0.4609 (E) 0.7901

(B) 0.3209 (D) 0.5390

Steps

1. Write the description of the corresponding random variable X:

Let X be the ....random........ ....variable........ for the ....number........ of Australians in a ....sample........ of five Australians that live in a capital city.

2. What distribution does X follow?

- ....Binomial........
- Case of success: lives in a capital city
- Case of failure: does not live in a capital city

3. Relate $\widehat{P}$ to X, and $P\left(\widehat{P}=0\right)$ to the random variable $P(X=0)$:

$$\widehat{P}=\frac{X}{5}\quad\Rightarrow\quad P\left(\widehat{P}=0\right)=P(X=0)$$

4. Relate $P\left(\widehat{P}>0.6\right)$ to $P(X>k)$ for some value of k and use the appropriate distribution to find the correct probability:

$$\begin{aligned}P\left(\widehat{P}>0.6\right)&=P\left(\tfrac{X}{5}>0.6\right)\\&=P(X>3)\\&=\text{[terms in } p^4 \text{ and } p^5\text{]}\\&=\binom{5}{4}\left(\tfrac{2}{3}\right)^4\left(\tfrac{1}{3}\right)+\binom{5}{5}\left(\tfrac{2}{3}\right)^5\\&=\tfrac{112}{243}\approx 0.4609\end{aligned}$$

Example 18

[**2016 VCE Mathematical Methods Paper 2 Q17**] Inside a container there are one million coloured building blocks. It is known that 20% of the blocks are red. A sample of 16 blocks is taken from the container. For samples of 16 blocks, $\widehat{P}$ is the random variable of the distribution of sample proportions of red blocks. (Do not use a normal approximation.)

$P\left(\widehat{P} \geq \dfrac{3}{16}\right)$ is closest to

(A) 0.6482 (C) 0.7543 (E) 0.3211

(B) 0.8593 (D) 0.6542

Answer: 0.6482

Example 19

[2019 VCE Mathematical Methods Paper 1 Q16] Fred owns a company that produces thousands of pegs each day. He randomly selects 41 pegs that are produced on one day and finds eight faulty pegs.

(a) What is the proportion of faulty pegs in this sample? **1**

(b) Pegs are packed each day in boxes. Each box holds 12 pegs. Let $\widehat{P}$ be the random variable that represents the proportion of faulty pegs in a box. **2**

The actual proportion of faulty pegs produced by the company each day is $\dfrac{1}{6}$.

Find $P\left(\widehat{P} < \dfrac{1}{6}\right)$. Express your answer in the form $a \times b^n$ where a and b are positive rational numbers and n in a positive integer.

Answer: $\frac{17}{6} \times \left(\frac{5}{6}\right)^{11}$

Example 20

[2016 VCE Mathematical Methods Paper 2 Q3] A school has a class set of 22 new laptops kept in a recharging trolley. Provided each laptop is correctly plugged into the trolley after use, its battery recharges.

On a particular day, a class of 22 students uses the laptops. All laptop batteries are fully charged at the start of the lesson. Each student uses and returns exactly one laptop. The probability that a student does not correctly plug their laptop into the trolley at the end of the lesson is 10%. The correctness of any student's plugging-in is independent of any other student's correctness.

(a) Determine the probability that at least one of the laptops is not correctly plugged into the trolley at the end of the lesson. Give your answer correct to four decimal places. **2**

(b) A teacher observes that at least one of the returned laptops is not correctly plugged into the trolley. **2**

Given this, find the probability that fewer than five laptops are not correctly plugged in. Give your answer correct to four decimal places.

The time for which a laptop will work without recharging (the battery life) is normally distributed, with a mean of three hours and 10 minutes and standard deviation of six minutes. Suppose that the laptops remain out of the recharging trolley for three hours.

(c) For any one laptop, find the probability that it will stop working by the end of these three hours. Give your answer correct to four decimal places. **2**

A supplier of laptops decides to take a sample of 100 new laptops from a number of different schools. For samples of size 100 from the population of laptops with a mean battery life of three hours and 10 minutes and standard deviation of six minutes, $\widehat{P}$ is the random variable of the distribution of sample proportions of laptops with a battery life of less than three hours.

(d) Find the probability $P\left(\widehat{P} \geq 0.06 \,\middle|\, \widehat{P} \geq 0.05\right)$. Give your answer correct to three decimal places. Do not use a normal approximation. **3**

It is known that when laptops have been used regularly in a school for six months, their battery life is still normally distributed but the mean battery life drops to three hours. It is also known that only 12% of such laptops work for more than three hours and 10 minutes.

(e) Find the standard deviation for the normal distribution that applies to the battery life of laptops that have been used regularly in a school for six months, correct to four decimal places. **2**

The laptop supplier collects a sample of 100 laptops that have been used for six months from a number of different schools and tests their battery life. The laptop supplier wishes to estimate the proportion of such laptops with a battery life of less than three hours.

(f) Suppose the supplier tests the battery life of the laptops one at a time. **1**

Find the probability that the first laptop found to have a battery life of less than three hours is the third one.

Important note

⚠ Some calculations here can be very tedious as VCE students are able to use other assistive technology.

Answers

(a) $P(X \geq 1) = 0.9015$ (b) $P(X < 5|X > 1) \approx 0.9311$ (4 dp) (c) $P(Y \leq 180) \approx 0.0478$ (4 dp) (d) $P\left(\widehat{P} \geq 0.06 \,\middle|\, \widehat{P} \geq 0.05\right) \approx$ 0.658 (3 dp) (e) $\sigma \approx 8.5107$ (4 dp) (f) $\frac{1}{8}$

2.2.2 Approximating sample proportions via the normal distribution

Important note

- Identify n, p.
- Check the sample size! Is it binomial, or can the normal approximation be used?
  - General rule of thumb: see page 19
- If the normal approximation can be used,
  - Find $E(\widehat{p})$ and $\mathrm{Var}(\widehat{p})$.
  - The $\widehat{p}$ score becomes the x score. Find the z score by transforming the $\widehat{p}$ score.

Example 21

[**2020 Mathematics Extension 1 Sample HSC Q12**] A recent census showed that 20% of the adults in a city eat out regularly.

i. A survey of 100 adults in this city is to be conducted to find the proportion who eat out regularly. Show that the mean and standard deviation for the distribution of sample proportions of such surveys are 0.2 and 0.04 respectively. **2**

ii. Use the extract shown from a table giving values of $P(Z < z)$, where z has a standard normal distribution, to estimate the probability that a survey of 100 adults will find that at most 15 of those surveyed eat out regularly. **2**

(Use Table 1.1 on page 26)

Example 22

[2019 WACE Mathematics Methods Section 2 Q13] The proportion of working adults who miss breakfast on week days is estimated to be 40%. A study takes a random sample of 400 working adults.

For this sample:

i. What is the (approximate) distribution of the sample proportion of workers who miss breakfast? **2**

ii. What is the probability that the sample proportion of workers who miss breakfast is greater than 44%? **2**

Answer: i. $\widehat{p} \sim N(0.4, 0.0006)$ ii. $P(\widehat{p} > 0.44) \approx 0.0512$

Example 23

[**2018 WACE Mathematics Methods Section 2 Q17**] Tina believes that approximately 60% of the mangoes she produces on her farm are large. She takes a random sample of 500 mangoes from a day's picking

i. Assuming Tina is correct and 60% of the mangoes her farm produces are large, what is the approximate probability distribution of the sample proportion of large mangoes in her sample? **3**

ii. What is the probability that the sample proportion of large mangoes is less than 0.58? **2**

Answer: i. $\widehat{p} \sim N(0.6, 0.02191^2)$ ii. $P(\widehat{p} < 0.58) \approx 0.18066$

Example 24

[2020 Caringbah HS Ext 1 Trial Q12] Records show that 64% of students at a school travelled to and from school by bus. Samples of 100 students at the school are taken to determine the proportion who travel to and from school by bus.

i. Show that the distribution of such sample proportions has mean 0.64 and standard deviation 0.048. **2**

ii. Use the z score table on page 26 to estimate the probability that a sample of 100 students will contain at least 58 and at most 64 students who travel to and from school by bus. **3**

Answer: 0.3944

Example 25

[2021 Ext 1 HSC Q14] At a certain factory, the proportion of faulty items produced by a machine is $p = \dfrac{3}{500}$, which is considered to be acceptable. To confirm that the machine is working to this standard, a sample of size n is taken and the sample proportion $\widehat{p}$ is calculated.

It is assumed that $\widehat{p}$ is approximately normally distributed with $\mu = p$ and $\sigma^2 = \dfrac{p(1-p)}{n}$.

Production by this machine will be shut down if $\widehat{p} \geq \dfrac{4}{500}$.

The sample size is to be chosen so that the chance of shutting down the machine unnecessarily is less than 2.5%.

Find the approximate sample size required, giving your answer to the nearest thousand.

Answer: $n = 5\,964 \approx 6\,000$

Further exercises

Ex 17D

- Q1-10
- ⏭ Q11 onwards

Additional questions

(Use Table 1.1 on page 26 to answer these questions)

Source for Q3 onwards: Haese et al. (2016, Ex 9E)

1. [**2023 Killara HS Ext 1 HSC Q13**] NSW has been running a campaign to reduce smoking in young adults. The proportion of smokers in the general population is 21%.

In a survey of 1 200 randomly selected young adults, it was found that 226 people smoked regularly.

i. Find the sample proportion and sample standard deviation. **2**

ii. NSW are comparing their sample for young adults with the general population. **3**

By finding the probability that no more than 226 people in a sample of 1 200 would be regular smokers, determine whether the program has been successful. The program is considered successful if the probability is less than 5%.

2. [**2022 Mathematical Methods VCE NHT Paper 2 Q15**] The probability distribution for the discrete random variable X is shown in the table below.

x	0	1	2	3	4
$P(X=x)$	k	$2k$	$3k$	$4k$	$5k$

When a sample size of $n=4$ is taken, $P\left(\widehat{P}\leq\frac{1}{3}\right)$ is equal to

(A) 0 (B) 1 (C) $\frac{7}{3}$ (D) $\frac{1}{15}$ (E) $\frac{1}{5}$

3. Given a population proportion $p=0.2$ and sample size $n=100$, find the mean and standard deviation of the sample proportion $\widehat{p}$.

4. (a) Find the mean and standard deviation of the sample proportion $\widehat{p}$ given a population proportion $p=0.7$ and sample size:

i. $n=20$ ii. $n=50$ iii. $n=80$

(b) Discuss what happens to the standard deviation of $\widehat{p}$ as the sample size n increases.

5. 40% of fish in a lake are male. A sample of 60 fish are taken from the lake. Find the mean and standard deviation of $\widehat{p}$, the sample proportion of male fish.

6. 7% of Australians have blood type A^-. Let $\widehat{p}$ be the sample proportion of Australians with blood type A^- from a sample of 40.

(a) Find the mean and standard deviation of $\widehat{p}$.

(b) Is $\widehat{p}$ normally distributed? Explain your answer.

7. Five sevenths of households in a country town are known to own computers. Let $\widehat{p}$ be the sample proportion of households in the town which own a computer from a sample of size 200.

(a) Find the mean and standard deviation of $\widehat{p}$.

(b) Is $\widehat{p}$ normally distributed? Explain your answer.

(c) Estimate the probability that, in a sample of 200, between 72% and 73% (inclusive) of households own a computer.

8. A regular pentagon has sectors numbered 1, 1, 2, 3, 4. Estimate the probability that when the pentagon is spun 400 times, the result of a '1' occurs

(a) more than 37% of the time

(b) less than 43.75% of the time

9. 85% of the plum trees grown in a particular area produce more than 700 plums.

(a) State the conditions under which the sample proportion $\widehat{p}$ can be approximated by the normal distribution.

(b) If a random sample of 200 plum trees is selected, estimate the probability that:

i. less than 75% produce more than 700 plums.

ii. between 80% and 90% produce more than 700 plums.

10. A pre-election poll is run to estimate the proportion of voters who favour the Labor Party. The poll is based on one random sample of 2 500 voters. The true proportion of voters who support the Labor Party is $p = 0.465$.

(a) For the sample proportion $\widehat{p}$, find:

i. the mean   ii. the standard deviation

(b) Is $\widehat{p}$ normally distributed?

(c) Estimate the probability that:

i. the sample proportion is less than 0.46

ii. between 45% and 47% of voters in the sample favour the Labor Party

iii. the proportion $\widehat{p}$ of Labor Party supporters in the sample differs by more than 0.035 from p.

(d) Interpret your answer to (c)iii.

Answers

1. i. $\widehat{p} = \frac{113}{600}$, $\text{Var}(\widehat{p}) = 0.00013825$ ii. 0.003269 **2.** (E) **3.** mean $= 0.2$, s.d. $= 0.04$ **4.** (a) i. mean $= 0.7$, s.d. ≈ 0.102 ii. mean $= 0.7$, s.d. ≈ 0.0648 iii. mean $= 0.7$, s.d. ≈ 0.0512 (b) As sample size increases, the standard deviation decreases. **5.** mean $= 0.4$, s.d. ≈ 0.0632 **6.** (a) mean $= 0.07$, s.d. ≈ 0.0403 (b) No, as $np < 5$ **7.** (a) mean $= \frac{5}{7}$, s.d. ≈ 0.0319 (b) Yes, as $np > 5$ and $n(1-p) > 5$. (c) ≈ 0.118 **8.** (a) ≈ 0.846 (b) ≈ 0.937 **9.** (a) Sample size $n \geq 34$ (b) i. $\approx 0.000\,037\,4$ ii. ≈ 0.952 **10.** (a) i. 0.465 ii. $\approx 0.009\,98$ (b) Yes (c) i. ≈ 0.308 ii. ≈ 0.626 iii. $\approx 0.000\,451$ (d) It is very likely that the sample proportion of Labor Party supporters will lie within 3.5% of the population proportion.

NESA Reference Sheet – calculus based courses

NSW Education Standards Authority

2020 HIGHER SCHOOL CERTIFICATE EXAMINATION

Mathematics Advanced
Mathematics Extension 1
Mathematics Extension 2

REFERENCE SHEET

Measurement

Length

$$l = \frac{\theta}{360} \times 2\pi r$$

Area

$$A = \frac{\theta}{360} \times \pi r^2$$

$$A = \frac{h}{2}(a + b)$$

Surface area

$$A = 2\pi r^2 + 2\pi rh$$

$$A = 4\pi r^2$$

Volume

$$V = \frac{1}{3}Ah$$

$$V = \frac{4}{3}\pi r^3$$

Functions

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

For $ax^3 + bx^2 + cx + d = 0$:

$$\alpha + \beta + \gamma = -\frac{b}{a}$$

$$\alpha\beta + \alpha\gamma + \beta\gamma = \frac{c}{a}$$

and $\alpha\beta\gamma = -\frac{d}{a}$

Relations

$$(x - h)^2 + (y - k)^2 = r^2$$

Financial Mathematics

$$A = P(1 + r)^n$$

Sequences and series

$$T_n = a + (n - 1)d$$

$$S_n = \frac{n}{2}[2a + (n - 1)d] = \frac{n}{2}(a + l)$$

$$T_n = ar^{n-1}$$

$$S_n = \frac{a(1 - r^n)}{1 - r} = \frac{a(r^n - 1)}{r - 1}, \; r \neq 1$$

$$S = \frac{a}{1 - r}, \; |r| < 1$$

Logarithmic and Exponential Functions

$$\log_a a^x = x = a^{\log_a x}$$

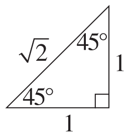
$$\log_a x = \frac{\log_b x}{\log_b a}$$

$$a^x = e^{x \ln a}$$

Trigonometric Functions

$\sin A = \dfrac{\text{opp}}{\text{hyp}}, \quad \cos A = \dfrac{\text{adj}}{\text{hyp}}, \quad \tan A = \dfrac{\text{opp}}{\text{adj}}$

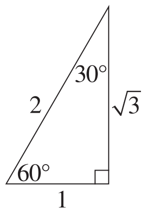
$A = \dfrac{1}{2}ab\sin C$



$\dfrac{a}{\sin A} = \dfrac{b}{\sin B} = \dfrac{c}{\sin C}$

$c^2 = a^2 + b^2 - 2ab\cos C$

$\cos C = \dfrac{a^2 + b^2 - c^2}{2ab}$



$l = r\theta$

$A = \dfrac{1}{2}r^2\theta$

Trigonometric identities

$\sec A = \dfrac{1}{\cos A}, \ \cos A \neq 0$

$\operatorname{cosec} A = \dfrac{1}{\sin A}, \ \sin A \neq 0$

$\cot A = \dfrac{\cos A}{\sin A}, \ \sin A \neq 0$

$\cos^2 x + \sin^2 x = 1$

Compound angles

$\sin(A + B) = \sin A\cos B + \cos A\sin B$

$\cos(A + B) = \cos A\cos B - \sin A\sin B$

$\tan(A + B) = \dfrac{\tan A + \tan B}{1 - \tan A\tan B}$

If $t = \tan\dfrac{A}{2}$ then $\sin A = \dfrac{2t}{1 + t^2}$

$\cos A = \dfrac{1 - t^2}{1 + t^2}$

$\tan A = \dfrac{2t}{1 - t^2}$

$\cos A\cos B = \dfrac{1}{2}\left[\cos(A - B) + \cos(A + B)\right]$

$\sin A\sin B = \dfrac{1}{2}\left[\cos(A - B) - \cos(A + B)\right]$

$\sin A\cos B = \dfrac{1}{2}\left[\sin(A + B) + \sin(A - B)\right]$

$\cos A\sin B = \dfrac{1}{2}\left[\sin(A + B) - \sin(A - B)\right]$

$\sin^2 nx = \dfrac{1}{2}(1 - \cos 2nx)$

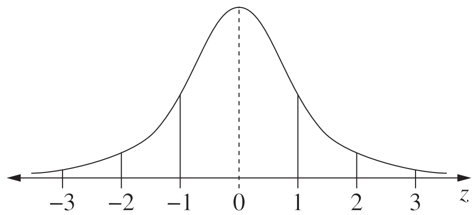
$\cos^2 nx = \dfrac{1}{2}(1 + \cos 2nx)$

Statistical Analysis

$z = \dfrac{x - \mu}{\sigma}$

An outlier is a score
less than $Q_1 - 1.5 \times IQR$
or
more than $Q_3 + 1.5 \times IQR$

Normal distribution



- approximately 68% of scores have z-scores between −1 and 1
- approximately 95% of scores have z-scores between −2 and 2
- approximately 99.7% of scores have z-scores between −3 and 3

$E(X) = \mu$

$\operatorname{Var}(X) = E\left[(X - \mu)^2\right] = E\left(X^2\right) - \mu^2$

Probability

$P(A \cap B) = P(A)P(B)$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A|B) = \dfrac{P(A \cap B)}{P(B)}, \ P(B) \neq 0$

Continuous random variables

$P(X \le x) = \displaystyle\int_a^x f(x)\,dx$

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

Binomial distribution

$P(X = r) = {}^nC_r p^r (1 - p)^{n - r}$

$X \sim \operatorname{Bin}(n, p)$

$\Rightarrow \quad P(X = x)$

$= \dbinom{n}{x} p^x (1 - p)^{n - x}, \ x = 0, 1, \ldots, n$

$E(X) = np$

$\operatorname{Var}(X) = np(1 - p)$

Differential Calculus

Function	Derivative
$y = f(x)^n$	$\dfrac{dy}{dx} = nf'(x)\left[f(x)\right]^{n-1}$
$y = uv$	$\dfrac{dy}{dx} = u\dfrac{dv}{dx} + v\dfrac{du}{dx}$
$y = g(u)$ where $u = f(x)$	$\dfrac{dy}{dx} = \dfrac{dy}{du} \times \dfrac{du}{dx}$
$y = \dfrac{u}{v}$	$\dfrac{dy}{dx} = \dfrac{v\dfrac{du}{dx} - u\dfrac{dv}{dx}}{v^2}$
$y = \sin f(x)$	$\dfrac{dy}{dx} = f'(x)\cos f(x)$
$y = \cos f(x)$	$\dfrac{dy}{dx} = -f'(x)\sin f(x)$
$y = \tan f(x)$	$\dfrac{dy}{dx} = f'(x)\sec^2 f(x)$
$y = e^{f(x)}$	$\dfrac{dy}{dx} = f'(x)e^{f(x)}$
$y = \ln f(x)$	$\dfrac{dy}{dx} = \dfrac{f'(x)}{f(x)}$
$y = a^{f(x)}$	$\dfrac{dy}{dx} = (\ln a)f'(x)a^{f(x)}$
$y = \log_a f(x)$	$\dfrac{dy}{dx} = \dfrac{f'(x)}{(\ln a)f(x)}$
$y = \sin^{-1} f(x)$	$\dfrac{dy}{dx} = \dfrac{f'(x)}{\sqrt{1-\left[f(x)\right]^2}}$
$y = \cos^{-1} f(x)$	$\dfrac{dy}{dx} = -\dfrac{f'(x)}{\sqrt{1-\left[f(x)\right]^2}}$
$y = \tan^{-1} f(x)$	$\dfrac{dy}{dx} = \dfrac{f'(x)}{1+\left[f(x)\right]^2}$

Integral Calculus

$$\int f'(x)\left[f(x)\right]^n dx = \frac{1}{n+1}\left[f(x)\right]^{n+1} + c$$

where $n \neq -1$

$$\int f'(x)\sin f(x)\,dx = -\cos f(x) + c$$

$$\int f'(x)\cos f(x)\,dx = \sin f(x) + c$$

$$\int f'(x)\sec^2 f(x)\,dx = \tan f(x) + c$$

$$\int f'(x)e^{f(x)}dx = e^{f(x)} + c$$

$$\int \frac{f'(x)}{f(x)}dx = \ln\left|f(x)\right| + c$$

$$\int f'(x)a^{f(x)}dx = \frac{a^{f(x)}}{\ln a} + c$$

$$\int \frac{f'(x)}{\sqrt{a^2-\left[f(x)\right]^2}}dx = \sin^{-1}\frac{f(x)}{a} + c$$

$$\int \frac{f'(x)}{a^2+\left[f(x)\right]^2}dx = \frac{1}{a}\tan^{-1}\frac{f(x)}{a} + c$$

$$\int u\frac{dv}{dx}dx = uv - \int v\frac{du}{dx}dx$$

$$\int_a^b f(x)\,dx$$

$$\approx \frac{b-a}{2n}\left\{f(a)+f(b)+2\left[f\left(x_1\right)+\cdots+f\left(x_{n-1}\right)\right]\right\}$$

where $a = x_0$ and $b = x_n$

Combinatorics

$$^{n}P_{r} = \frac{n!}{(n-r)!}$$

$$\binom{n}{r} = {}^{n}C_{r} = \frac{n!}{r!\,(n-r)!}$$

$$(x+a)^n = x^n + \binom{n}{1}x^{n-1}a + \cdots + \binom{n}{r}x^{n-r}a^r + \cdots + a^n$$

Vectors

$$\left|\underset{\sim}{u}\right| = \left|x\underset{\sim}{i} + y\underset{\sim}{j}\right| = \sqrt{x^2 + y^2}$$

$$\underset{\sim}{u} \cdot \underset{\sim}{v} = \left|\underset{\sim}{u}\right|\left|\underset{\sim}{v}\right|\cos\theta = x_1x_2 + y_1y_2,$$

where $\underset{\sim}{u} = x_1\underset{\sim}{i} + y_1\underset{\sim}{j}$

and $\underset{\sim}{v} = x_2\underset{\sim}{i} + y_2\underset{\sim}{j}$

$$\underset{\sim}{r} = \underset{\sim}{a} + \lambda\underset{\sim}{b}$$

Complex Numbers

$$\begin{aligned} z = a + ib &= r(\cos\theta + i\sin\theta) \\ &= re^{i\theta} \end{aligned}$$

$$\begin{aligned} \left[r(\cos\theta + i\sin\theta)\right]^n &= r^n(\cos n\theta + i\sin n\theta) \\ &= r^n e^{in\theta} \end{aligned}$$

Mechanics

$$\frac{d^2x}{dt^2} = \frac{dv}{dt} = v\frac{dv}{dx} = \frac{d}{dx}\left(\frac{1}{2}v^2\right)$$

$$x = a\cos(nt + \alpha) + c$$

$$x = a\sin(nt + \alpha) + c$$

$$\ddot{x} = -n^2(x - c)$$

References

Haese, M., Haese, S., & Humphries, M. (2016). *Mathematics for Australia 12 Mathematical Methods* (2nd ed.). Haese Mathematics.

Pender, W., Sadler, D., Ward, D., Dorofaeff, B., & Shea, J. (2019). *CambridgeMATHS Stage 6 Mathematics Extension 1 Year 12* (1st ed.). Cambridge Education.