MEI

Mathematics in Education and Industry

MEI Structured Mathematics

Module Summary Sheets

Statistics 1

(Version B: reference to new book)

Topic 1: Exploring Data

Topic 2: Data Presentation

Topic 3: Probability

Topic 4: The Binomial Distribution

Topic 5: Discrete Random Variables

MEI, Oak House, 9 Epsom Centre, White Horse Business Park, Trowbridge, Wiltshire. BA14 0XG.
Company No. 3265490 England and Wales Registered with the Charity Commission, number 1058911
Tel: 01225 776776. Fax: 01225 775755.

References:
Chapter 1
Pages 12-13

Types of data

Categorical data are data that are listed by a category or property rather than a number

Discrete data are data that can only take particular values
E.g. numbers of cars, marks in an examination

Continuous data are data that can take any value. These data will be given in a rounded form depending on the precision of measurement.
E.g. mass, length, temperature

E.g. colours of "Smarties" in a tube are categorical data.

E.g. Marks gained by students in an examination:

14, 17, 19, 19, 19, 20, 20, 21, 21, 21,
21, 25, 25, 26, 26, 27, 27, 30, 30, 32

These are discrete data.

E.g. heights of students measured to the nearest cm:

160, 162, 164, 165, 165, 165, 166, 168, 169, 169,
171, 172, 172, 173, 174, 175, 176, 177, 180, 182

These are continuous data.

The height recorded as 178 cm is in range
$177.5 \le 178 < 178.5$

References:
Chapter 1
Pages 6-8

Exercise 1A
Q. 2, 4

Stem and Leaf diagrams

A stem and leaf diagram is a way of writing out the data so that the shape of the distribution can be seen and so that outliers can be seen.
A key should always be given.

E.g. Stem and leaf diagram of the set of marks above.

1	4 7 9 9 9
2	0 0 1 1 1 1 5 5 6 6 7 7
3	0 0 2

Key: 3 | 2 means 32

References:
Chapter 1
Pages 13-16

Exercise 1B
Q. 2 (i), (iv)

Measures of Central Tendency

A measure of central location gives a typical ("representative") value for a data set.

Mean: $\bar{x} = \frac{1}{n}\sum x_i$

Median is the mid-value of the set, when arranged in ascending or descending order

i.e. the $\frac{n+1}{2}$ th value.

Mode is the most frequent value.

Midrange is the mid-point of the highest and lowest values.

E.g. for the marks gained in an examination:

Mean $= \bar{x} = \frac{14+17+.....+32}{20} = \frac{460}{20} = 23$

Median $= \frac{21+21}{2} = 21$

(Halfway between 10th and 11th values)

Mode = 21

Mid-range $= \frac{14+32}{2} = 23$

References:
Chapter 1
Pages 17-19

Exercise 1C
Q. 2, 4

Frequency distributions

There may be some elements of a data set that are equal. In this situation the data may be summarised in a frequency table.

E.g. for the marks gained in an examination:

x	14	17	19	20	21	25	26	27	30	32
f	1	2	2	2	4	2	2	2	2	1

References:
Chapter 1
Pages 22-29

Exercise 1D
Q. 2, 4

Grouped data

Data can be grouped into frequency distributions because:

- There may be a lot of data
- The data may be spread over a wide range
- Most of the values are different

E.g. Find the mean of the grouped heights

Height(cm)	Mid pt	f	fx
159.5 – 164.5	162	3	486
164.5 – 169.5	167	7	1169
169.5 – 174.5	172	5	860
174.5 – 179.5	177	3	531
179.5 – 189.5	184.5	2	369
Total:		20	3415

$\text{Estimate of Mean} = \frac{3415}{20} = 170.75$

You should make sure that you can use your calculator to carry out such calculations—the table is not necessary if you use it properly!

Modal class: 164.5 - 169.5

Summary S1 Topic 1: **Exploring Data—2**

References:
Chapter 1
Pages 31-40

Exercise 1E
Q. 1, 5

Measures of Spread

Indicate how widely spread are the data.

Range: Highest value – lowest value

For any item in the data set, deviation $= x_i - \bar{x}$

Sum of deviations $= \sum \left(x_i - \bar{x} \right) = 0$ always

Sum of absolute deviations $= \sum \left| \left(x_i - \bar{x} \right) \right|$

Mean Absolute Deviation $= \frac{1}{n} \sum \left| \left(x_i - \bar{x} \right) \right|$

Sum of squares of deviations, $S_{xx} = \sum \left(x_i - \bar{x} \right)^2$

$= \sum x_i^{\,2} - n\bar{x}^{\,2}$

Example 1.5
Page 39

Mean Square Deviation $= \frac{S_{xx}}{n}$

Root mean square deviation(rmsd) $= \sqrt{\frac{S_{xx}}{n}}$

Variance, $s^2 = \frac{S_{xx}}{n-1}$,

Standard deviation, $s = \sqrt{\frac{S_{xx}}{n-1}}$

References:
Chapter 1
Pages 40-41
Pages 73-74

An outlier is usually taken to be a value that is more than 2 standard deviations from the mean. It may also be a value that is more than 1.5 × IQR beyond the nearest quartile.

E.g. Find the measures of spread of examination marks on the previous page.

Note: You need to know how to use your calculator—different makes work in different ways. They will have different buttons available in the statistical mode.

$\text{Range} = 32 - 14 = 18$

$S_{xx} = (14-23)^2 + 2\times(17-23)^2 + ...(32-23)^2 = 436$

Or: $\sum fx^2 = 11016$

So $S_{xx} = \sum fx^2 - n\bar{x}^{\,2} = 11016 - 20\times 23^2 = 436$

$msd = \frac{436}{20} = 21.8 \Rightarrow rmsd = 4.669$

$s^2 = \frac{436}{19} = 22.95 \Rightarrow s = 4.790$

E.g. Find the mean of the following values

x	f
33	4
35	5
37	7
39	4

$y = x - 33$
$z = y/2$ gives the following table

x	y	z	f	zf
33	0	0	4	0
35	2	1	5	5
37	4	2	7	14
39	6	3	4	12

$\sum zf = 31 \Rightarrow \bar{z} = \frac{31}{20} = 1.65$

$\Rightarrow \bar{y} = 3.1 \Rightarrow \bar{x} = 36.1$

References:
Chapter 1
Pages 46-48

Exercise 1F
Q. 2

Linear Coding

If $y = a + bx$ then $\bar{y} = a + b\bar{x}$; $s_y^{\,2} = b^2 s_x^{\,2}$

$s_y = |b| s_x$

E.g. If the examination marks are doubled and then decreased by 20:

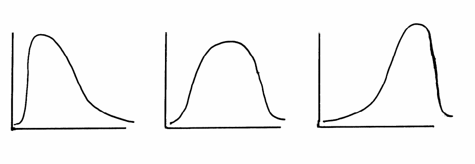
$\bar{y} = 2\bar{x} - 20 = 2 \times 23 - 20 = 26$

$s_y = 2 s_x = 9.581$

References:
Chapter 1
Pages 1-6

Skewness

When the mode of a frequency distribution is off to one side then the data may exhibit positive or negative skewness.

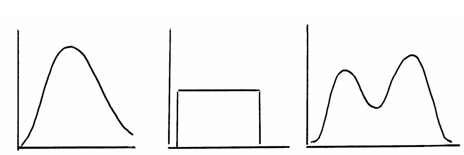


Positive Symmetrical Negative

N.B. The "skewness" goes "with the tail" rather then the “hump”.

Shapes of Distributions

Unimodal Uniform Bimodal



N.B. "Bimodal" does not mean that the peaks are of the same height.

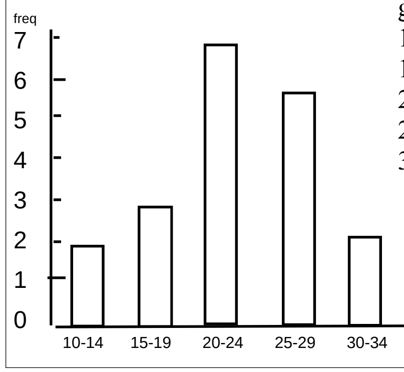
References:
Chapter 2
Pages 56-59

Exercise 2A
Q. 6

Bar Charts

A Bar Chart should be used for categorical data. Bar Charts may also be used for discrete data. There should be equal widths between the bars, which should also be of equal widths.

Bar Chart



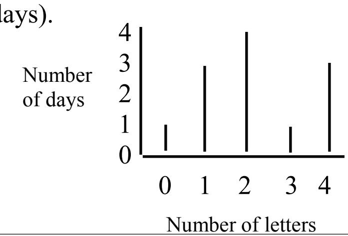
group	frequency
10-14	2
15-19	3
20-24	7
25-29	6
30-34	2

Vertical Line Charts

A vertical line chart should be used for discrete data that can only take integer values. It is like a bar chart with minimum width!

E.g. the number of letters received by a householder in a 2 week period (12 days).

No of letters	Days
0	1
1	3
2	4
3	1
4	3



References:
Chapter 2
Pages 59-60

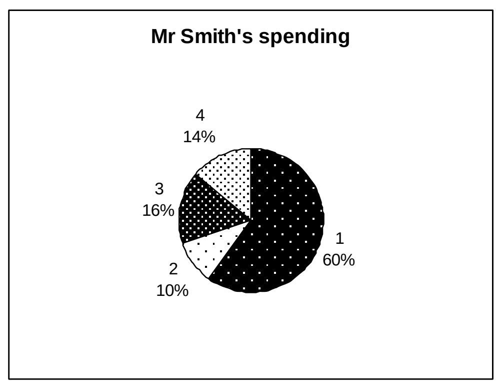
Exercise 2A
Q. 2

Pie Charts

Pie charts are used to illustrate data when the size of each group relative to the whole is required.
The group size is represented by the correct proportion of a circle.
When two sets of data are to be compared with pie charts the area of each chart is proportional to the size of the data set.

E.g. Mr Smith spends £500 per week as follows:

Rent/rate/insurance etc.	£300	- 1
Food	£ 50	- 2
Entertainment	£ 80	- 3
Savings for holidays etc	£ 70	- 4



References:
Chapter 2
Pages 62-69

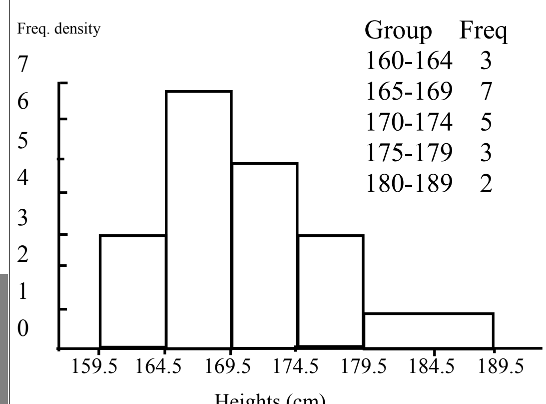
Exercise 2B
Q. 1, 3

Histograms

Histograms are used to illustrate continuous data. Each group (which must have no gaps) should abut each other. The area of the block is proportional to the frequency in order to allow for different widths of groups.
The vertical axis is therefore the frequency density and ***not*** the frequency.
Histograms are occasionally used for grouped discrete data. In this case the horizontal axis must be continuous and there must be no gaps. Hence the group 0—9, 10—19, etc would become $-\frac{1}{2}-9\frac{1}{2},\ 9\frac{1}{2}-19\frac{1}{2}$, etc,

Histogram of heights (see page 2)

Group	Freq
160-164	3
165-169	7
170-174	5
175-179	3
180-189	2



References:
Chapter 2
Pages 71-73

Example 2.1
Page 72

Exercise 2C
Q. 1

Quartiles

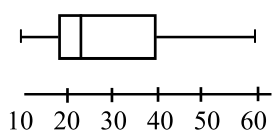
The ***Lower quartile*** is the 25% value.
The ***Median*** is the 50% value.
The ***Upper quartile*** is the 75% value.
These are denoted Q_1, Q_2 and Q_3 respectively.
The use of quartiles should be avoided for small data sets.
The ***Interquartile range (IQR)*** = $Q_3 - Q_1$

References:
Chapter 2
Page 73

Exercise 2A
Q. 6

Box and Whisker Plots

A box and whisker plot is shown here with a "box" at Q_3 and Q_1 together with a vertical line at the Median (Q_2). The ends of the horizontal line are the extreme values of the data.



References:
Chapter 2
Pages 74-77

Exercise 2C
Q. 4

Cumulative Frequency Curves.

Cumulative frequency curves show the accumulated frequency up to each upper class boundary.

On a graph, cumulative frequency is plotted against the upper class value.
It would not be appropriate to use a cumulative frequency curve for discrete values unless they are grouped.
The median (50th percentile) and the quartiles (25th and 75th percentiles) can be estimated from the graph.
The interquartile range
= upper quartile(Q_3) – lower quartile(Q_1)

N.B. For grouped values, $Q_1 = \frac{n}{2}, Q_2 = \frac{n}{2}, Q_3 = \frac{3n}{4}$.

References:
Chapter 2
Pages 40-41
Pages 73-74

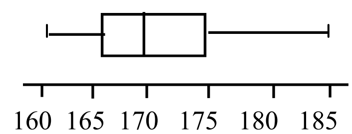
Outliers

An outlier is an extreme value. Such a value is defined quantitatively as being more than a certain "distance" from the average.
When the mean and s.d. are being used then it is defined as being more than 2 s.d. from the mean.
When the median and IQR are being used then it is defined as being more than 1.5 × the IQR above Q_3 or below Q_1.

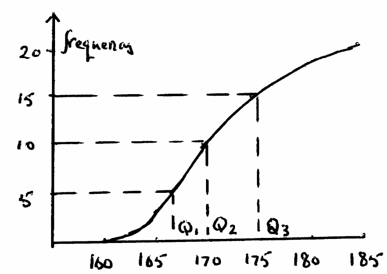
E.g. the heights of a group of students are summarised in the table.

Height(cm)	Freq.	Cum Freq.	Upper ht
		0	159.5
159.5 – 164.5	3	3	164.5
164.5 – 169.5	7	10	169.5
169.5 – 174.5	5	15	174.5
174.5 – 179.5	3	18	179.5
179.5 – 189.5	2	20	189.5

The box and whisker plot is as follows:



The cumulative frequency graph is as follows:



N.B. Sometimes the cumulative frequency diagram is drawn using straight lines between the points.

For this set of data and the graph:
Median (Q_2) ≈ 170 cm
$Q_1 \approx 166$ cm
$Q_3 \approx 175$ cm
IQR($Q_3 - Q_1$) ≈ 9 cm

An outlier would have a value
> 175 + 1.5 × 9 ≈ 188.5
or < 166 – 1.5 × 9 ≈ 152.5

152.5 cm is well below the range of data, but 188.5 cm is within the upper group, so it is possible that there is an outlier there.

This can only be determined from the original data points, which in this example we do not have.

References:
Chapter 3
Pages 86-90

Exercise 3A
Q. 1

Probability

Describes the likelihood of different possible outcomes or events occurring as a result of some trial.
For events with equally likely outcomes

Theoretical probability $P(A) = \dfrac{n(A)}{N}$

Experimental probability = $\dfrac{No.of\ successful\ outcomes}{No.of\ trials}$

(This is an estimate of the probability from an experiment.) $0 \le P(A) \le 1$.

E.g. In a school of 800 pupils, 500 are boys. If a pupil is chosen at random the probability of a boy can be estimated as

$$\frac{500}{800} = \frac{5}{8}$$

E.g. The probability of drawing an ace from a pack of cards = $\dfrac{4}{52} = \dfrac{1}{13}$

E.g. A biased coin is tossed 100 times and it comes down tails 75 times.
Experimental probability that the next toss will give heads = 0.25

References:
Chapter 3
Pages 90-91

Sample Space* or *Probability Space is an aid to consider the set of all possible outcomes.

Venn diagrams are useful tools in understanding the logic of a problem.

The complement of event A, denoted A', is the event "A does not happen". $P(A) + P(A') = 1$

The ***Expectation*** of an outcome = np

E.g. When two dice are thrown the sample space for the total score is as follows:

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$P(7) = \dfrac{6}{36} = \dfrac{1}{6}$

$P(7') = 1 - \dfrac{1}{6} = \dfrac{5}{6}$

References:
Chapter 3
Pages 92-95

Example 3.5
Page 95

Exercise 3A
Q. 6

Mutually Exclusive events are events that cannot happen simultaneously.
The Addition Rule for mutually exclusive events is

$$P(A \text{ or } B) = P(A \cup B) = P(A) + P(B)$$

Non-mutually exclusive events

$$P(A \text{ or } B) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

where
$P(A \cap B)$ means = P(A and B)

E.g. probability of drawing an ace or a king from a pack of cards when one card is drawn. (Mutually exclusive events)

$$= P(A) + P(K) = \frac{4}{52} + \frac{4}{52} = \frac{8}{52} = \frac{2}{13}$$

E.g. Probability of drawing an ace or a spade from a pack of cards when one card is drawn (non-mutually exclusive events)

$$= P(A) + P(S) - P(A \cap S) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}$$

References:
Chapter 3
Pages 98-102

Exercise 3B
Q. 6

References:
Chapter 3
Pages 107-112

Exercise 3C
Q. 6

Independent events are events where the outcome of one has no effect on the possible outcome of another.

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

If there is a dependence then they are called Dependent Events.

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

where $P(B|A)$ means the probability of event B occurring, given that A has happened. $P(B|A)$ is called a conditional probability. A conditional probability is used when your estimate of the probability of an event occuring is altered by a knowledge of whether another event has occurred.

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

E.g. Probability of drawing an ace then a king if two cards are drawn from a pack and the first is replaced before the second is drawn.
(Events are independent)

$$= P(A) \times P(K) = \frac{4}{52} \times \frac{4}{52} = \frac{1}{169}$$

E.g. probability of drawing a king and then an ace if the first card is not replaced.
(Dependent events)

$$= P(A) \times P(K|A) = \frac{4}{52} \times \frac{4}{51} = \frac{4}{663}$$

Competence statements u1, u2, u3, u4, u5, u6, u7, u8, u9, u10, u11, u12

References:
Chapter 5
Pages138-140

Exercise 5A
Q. 1, 3, 4

Factorials

"6 factorial" is written 6! and its value is

$$6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1$$

It represents the number of ways of placing 6 different items in a line.

E.g. $7! = 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 5040$
The numbers get very large and different calculators revert to standard form at different values.
i.e. $14! = 8.71782912 \times 10^{10}$

References:
Chapter 5
Pages 142-146

Example 5.5
Page 144

Exercise 5B
Q. 1, 2, 3

Permutations and Combinations

A ***permutation*** is a selection of things in which the order matters.

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

A ***combination*** is a selection of things in which the order does not matter.
The number of ways of choosing r items from n distinct items is given by

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

Symmetry: $^{n}C_{r} = \binom{n}{r} = \dfrac{n!}{(n-r)!r!} = \binom{n}{n-r} = {}^{n}C_{n-r}$

E.g. the number of ways of choosing 2 items from 10.

$$^{10}C_2 = \binom{10}{2} = \frac{10!}{8!2!} = \frac{10.9}{1.2} = 45$$

E.g. the number of ways of choosing 8 items from 10.

$$^{10}C_8 = \binom{10}{8} = \frac{10!}{8!2!} = \frac{10.9}{1.2} = 45$$

i.e. $^{10}C_8 = {}^{10}C_2$

References:
Chapter 5
Page 145

Exercise 5B
Q. 8

Pascal's Triangle

The row starting 1 n… gives the $(n+1)$ coefficients in the expansion of $(a+b)^n$. They represent the number of ways of choosing 0, 1, 2,… n items from n.

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1

E.g. the number of ways of choosing 3 items from 5.
Extend Pascal's Triangle by one more line.

1 3 3 1
1 4 6 4 1
1 5 10 10 5 1

The number required is the 3rd entry in the 5th line. i.e. 10

References:
Chapter 5
Page 147-149

Exercise 5B
Q. 7

Probabilities in less simple cases

Combinations are sometimes combined by multiplication.
i.e. The number of ways of choosing 3 males from 5 and also 4 females from 6

$$= {}^{5}C_3 \times {}^{6}C_4 = 10 \times 15 = 150$$

E.g. A committee of three is to be selected from 4 boys and 3 girls.
(i) How many different selections are possible?
(ii) How many of these contain 2 boys and 1 girl?
(iii) What is the probability that the committee consists of 2 boys and 1 girl?

(i) There are 7 people so the total number of committees is $^{7}C_3 = \dfrac{7!}{3!4!} = 35$.

(ii) The number of ways of choosing 2 boys from 4 and 1 girl from 3 is $^{4}C_2 \times {}^{3}C_1 = 6 \times 3 = 18$

(iii) Prob $= \dfrac{18}{35}$

References:
Chapter 6
Pages 153-156

Exercise 6A
Q. 4

References:
Chapter 6
Pages 158-159

The Binomial Distribution

Requires

- Two outcomes where the constant probability of success is p and failure is $q = 1 - p$
- n Trials
- Outcome of any trial is independent of all other trials.

If a random variable X is binomially distributed, we write $X \sim B(n,p)$ and

$$P(X = r) = \binom{n}{r} p^r q^{n-r}$$

Mean (Expected Value) = np

E.g. A fair die is thrown 10 times. Find the probability of throwing (i) 4 sixes, (ii) at least two sixes.

Let X = No. of sixes thrown. $p = \frac{1}{6},\ q = \frac{5}{6}$

$X \sim B(10, {}^1/_6)$

(i) $P(X = 4) = \binom{10}{4}\left(\frac{1}{6}\right)^4\left(\frac{5}{6}\right)^6 = 0.0543$

(ii) $P(X = 0 \text{ or } 1) = \left(\frac{5}{6}\right)^{10} + 10.\left(\frac{5}{6}\right)^9\left(\frac{1}{6}\right) = 0.1615 + 0.3230$

$= 0.4845$

$\Rightarrow P(\text{at least 2 sixes}) = 1 - 0.4845 = 0.5155$

Example 6.3
Page 160

Exercise 6B
Q. 7

Cumulative Binomial Probability Tables.

Binomial probabilities can be found using the formula for each term as above or by using the Cumulative Binomial Probability tables on pages 34 – 39 of the Students' Handbook.

e.g. the top left entry on page 39 is for $n = 20$ and $p = 0.05$. The entry represents the term $P(X = 0) = 0.95^{20} = 0.3585$.

The entry below it represents $P(X \le 1) = 0.7358$.
Calculating each term gives
$P(X \le 1) = P(X = 0) + P(X = 1)$
$= (0.95)^{20} + 20(0.95)^{19}(0.05) = 0.3585 + 0.3774$
$= 0.7358$

E.g. A fair die is thrown 20 times. Find the probability of getting: (i) five or fewer 6's, (ii) more than seven 6's, (iii) four 6's.

X=No of sixes thrown; $p = \frac{1}{6}, q = \frac{5}{6}$

$X \sim B(20, \frac{1}{6})$

(i) $P(X \le 5) = 0.8982$ (see page E17)

(ii) $P(X > 7) = 1 - P(X \le 7) = 1 - 0.9887 = 0.0113$

(iii) $P(X=4) = P(X \le 4) - P(X \le 3) = 0.7687 - 0.5665 = 0.2022$

References:
Chapter 7
Pages 167-175
Pages 177-179

Exercise 7A
Q. 2, 7

References:
Chapter 7
Pages 182-184

Example 7.2
Page 178

Exercise 7B
Q. 2

Exercise 7C
Q. 6

Hypothesis Testing

A null hypothesis is a hypothesis which is to be tested.
The alternative hypothesis represents the departures from the null hypothesis. In setting up the alternative hypothesis a decision needs to be made regarding whether it is to be 1-sided or 2-sided.
The significance level needs to be decided.

Hypothesis testing process

- Establish null and alternative hypotheses
- Decide on the significance level
- Collect data
- Conduct test
- Interpret result in terms of the original claim

1-tail or 2-tail tests

- 1-tail tests are applied if departures in a particular direction are emphasised. There is only one critical region.
- 2-tail tests are applied if no particular direction is emphasised. There are two critical regions.
- A 10% significance level in a 2-tail test gives a critical region of 5% in each tail.

E.g. A man claims to be able to tell the throw of a die by mind reading. He is asked to guess the throw of a die 20 times and is correct 9 times. Find the probability that he is correct 9 or more times. Comment on the test result. Find how many times he would have had to be correct to verify his claim at the 5% significance level.

X = No of correct claims; $p = \frac{1}{6}, q = \frac{5}{6}$

$X \sim B(20, \frac{1}{6})$

$H_0: p = \frac{1}{6}$ (He is correct only by chance)

$H_1: p > \frac{1}{6}$ (He is mind reading)

$P(X \ge 9) = 1 - P(X \le 8) = 1 - 0.9972 = 0.0018\ (< 0.05)$

So reject H_0 and conclude that he is mind reading.
Establish critical values at 5% :
$P(X \ge 7) = 1 - P(X \le 6) = 1 - 0.9629 = 0.031$ (3%)
$P(X \ge 6) = 1 - P(X \le 5) = 1 - 0.8982 = 0.1018$ (10%)
7 is the smallest value which lies inside the 5% region.
(He would have to be correct 7 or more times)

E.g. To test a claim a coin is biased towards heads a 1-tail test is appropriate.
If the claim is the the coin is biased (either way) a 2-tail test should be used.
$H_0: p = 0.5;\ H_1: p \ne 0.5$

References:
Chapter 4
Pages 118-124

Example 4.1
Page 121

Exercise 4A
Q 5 ,8

A Discrete Random Variable is a random variable which is a set of discrete values such as {0,1,2,3, ...n} and is used to model discrete data.

The particular values taken are usually denoted by lower case letters and we usually use suffices.
e.g. $x_1, x_2, x_3, x_4, \ldots, x_n$

This means that, given the set of numbers
$x_1, x_2, x_3, x_4, \ldots, x_n$, which occur with probabilities
$p_1, p_2, p_3, p_4, \ldots, p_n$,
then $p_1 + p_2 + p_3 + p_4 + \ldots + p_n = 1$.

A random variable is usually denoted by an upper case letter, e.g. X.

$P(X = x_i) = p_i$ means the probability that the random variable X takes the particular value x_i is p_i.

For a small set of values we can often conveniently list the associated probabilities.

The rule which assigns probabilities to the various outcomes is known as the Probability Function of X.

Sometimes it is possible to write the probability function as a formula.

Example 3
The Random Variable X is given by the number of heads obtained when three fair coins are tossed together.
The probability distribution is as follows:

x	0	1	2	3
p	$^1/_8$	$^3/_8$	$^3/_8$	$^1/_8$

e.g. $P(X= 2) = {}^3/_8$

When three fair coins are tossed together the outcomes are theoretically as follows:

HHH HHT HTH THH TTH THT HTT TTT,

from which it can be seen that exactly two heads appear in three of the eight possible outcomes.

Example 4
The Random Variable Y is given by the formula
$P(Y = r) = kr$ for $r = 1, 2, 3, 4$. Find the value of k and the probability distribution.

The distribution is as follows:

y	1	2	3	4
p	k	$2k$	$3k$	$4k$

Since $\Sigma p = 1$, $k + 2k + 3k + 4k = 1$ giving $k = {}^1/_{10}$

The probability distribution is therefore as follows:

y	1	2	3	4
p	$^1/_{10}$	$^2/_{10}$	$^3/_{10}$	$^4/_{10}$

References:
Chapter 4
Pages 126-130

Example 4.4
Page 130

Expectation
If a discrete random variable, X, takes possible values
$x_1, x_2, x_3, x_4, \ldots, x_n$
with associated probabilities $p_1, p_2, p_3, p_4, \ldots, p_n$
then the Expectation $E(X)$ of X is given by

$$E(X) = \Sigma x_i p_i$$

The Expectation is just another name for the mean. It is therefore sometimes denoted by μ, the symbol used for the mean of a population.

Expectation of a function of X.
If g[X] is a function of the discrete random variable, X, then E(g[X]) is given by $E(g[X]) = \Sigma\, g[x_i]p_i$

e.g. $E(X^2) = \Sigma x_i^2 p_i$
N.B. $E(X^2)$ is not the same as $[E(X)]^2$.

E.g. For example 3 we have:

$$E[X] = \sum xp = 0\times\frac{1}{8} + 1\times\frac{3}{8} + 2\times\frac{3}{8} + 3\times\frac{1}{8}$$
$$= \frac{12}{8} = 1.5$$
$$E[X^2] = \sum x^2 p = 0^2\times\frac{1}{8} + 1^2\times\frac{3}{8} + 2^2\times\frac{3}{8} + 3^2\times\frac{1}{8}$$
$$= \frac{24}{8} = 3$$

Exercise 4B
Q. 6, 7

Exercise 4C
Q. 1, 2

Variance
The variance of a discrete random variable X, $Var(X)$, is given by the formula $Var(X) = E[(X - \mu)^2]$.
Since the variance of a distribution can also be given by the formula $s^2 = \Sigma x_i^2 p_i - (\Sigma x_i p_i)^2$,
we obtain the alternative, and very useful formula

$$Var(X) = s^2 = E(X^2) - [E(X)]^2 = E(X^2) - \mu^2 .$$

For example 3 we have:

$E[X^2] = 3$, $E[X] = 1.5$
$Var(X) = 3 - 1.5^2 = 0.75$