6.13 STATISTICS 1, S1 (4766) AS

Objectives

To enable students to build on and extend the data handling and sampling techniques they have learnt at GCSE.

To enable students to apply theoretical knowledge to practical situations using simple probability models.

To give students insight into the ideas and techniques underlying hypothesis testing.

Assessment

Examination (72 marks)
1 hour 30 minutes
The examination paper has two sections:

Section A: 5 - 7 questions, each worth at most 8 marks.
Section Total: 36 marks

Section B: two questions, each worth about 18 marks.
Section Total: 36 marks

Assumed Knowledge

Candidates are expected to know the content for Intermediate Tier GCSE*. In addition, they need to know the binomial expansion as covered in *C1*.

*See note on page 34.

Calculators

In the MEI Structured Mathematics specification, no calculator is allowed in the examination for *C1*. For all other units, including this one, a graphical calculator is allowed.

*The use of an asterisk * in a competence statement indicates assumed knowledge. These items will not be the focus of examination questions and are included for clarity and completeness. However, they may be used within questions on more advanced statistics.*

STATISTICS 1, S1		
Specification	**Ref.**	**Competence Statements**
PROCESSES		
This section is fundamental to all the statistics units in this specification (Statistics 1-4). In this unit, the ideas may be used in examination questions but will not be their main subject.		
Statistical modelling.	S1p1	Be able to abstract from a real world situation to a statistical description (model).
	2	Be able to apply an appropriate analysis to a statistical model.
	3	Be able to interpret and communicate results.
	4	Appreciate that a model may need to be progressively refined.
Sampling.	5	* Understand the meanings of the terms population and sample.
	6	* Be aware of the concept of random sampling.
DATA PRESENTATION		
Classification and visual presentation of data.	S1D1	* Know how to classify data as categorical, discrete or continuous.
	2	* Understand the meaning of and be able to construct frequency tables for ungrouped data and grouped data.
	3	* Know how to display categorical data using a pie chart or a bar chart.
	4	Know how to display discrete data using a vertical line chart.
	5	Know how to display continuous data using a histogram for both unequal and equal class intervals.
	6	* Know how to display and interpret data on a stem and leaf diagram.
	7	* Know how to display and interpret data on a box and whisker plot.
	8	Know how to display and interpret a cumulative frequency distribution.
	9	Know how to classify frequency distributions showing skewness.

STATISTICS 1, S1			
Ref.	**Notes**	**Notation**	**Exclusions**
	PROCESSES		
	This section is fundamental to all the statistics units in this specification (Statistics 1-4). In this unit, the ideas may be used in examination questions but will not be their main subject.		
S1p1	Approximation and simplification involving appropriate distributions and probability models.		Formal definitions.
2			
3	Their implications in real-world terms.		
4	Check against reality.		
5			
6			
	DATA PRESENTATION		
S1D1			
2	Define class intervals and class boundaries.		
3			
4			
5	Area proportional to frequency. Use of the term frequency density will be expected.		
6	The term stemplot is also widely used. Stem and leaf diagrams will be expected to be sorted.		
7	The term boxplot is also widely used. The term outlier can be applied to data which are at least $1.5 \times$ IQR beyond the nearer quartile.		
8			
9	Positive and negative skewness.		Measures of skewness.

STATISTICS 1, S1		
Specification	**Ref.**	**Competence Statements**
DATA PRESENTATION (continued)		
Measures of central tendency and dispersion.	10	Know how to find median*, mean*, mode* and midrange.
	11	Know the usefulness of each of the above measures of central tendency.
	12	Know how to find range*, percentiles, quartiles* and interquartile range*.
	13	Know how to calculate and interpret mean squared deviation, root mean squared deviation, variance and standard deviation.
	14	Be able to use the statistical functions of a calculator to find mean, root mean square deviation and standard deviation.
	15	Know how the mean and standard deviation are affected by linear coding.
	16	Understand the term outlier.

STATISTICS 1, S1			
Ref.	**Notes**	**Notation**	**Exclusions**
	DATA PRESENTATION		
10	For raw data, frequency distributions, grouped frequency distributions.	Mean = $\bar{x}$	
11			
12			
13	For raw data, frequency distributions, grouped frequency distributions. The term outlier can be applied to data which are at least 2 standard deviations from the mean. $msd = \frac{S_{xx}}{n} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$, $rmsd = \sqrt{msd}$.	*msd, rmsd*	Corrections for class interval in these calculations.
	Sample variance: $s^2 = \frac{S_{xx}}{n-1} = \frac{1}{(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2$. (✝)	s^2	
	Sample standard deviation: $s = \sqrt{\text{variance}}$. (§)	s	
14			
15	$y_i = a + bx_i \Rightarrow \bar{y} = a + b\bar{x}$, $s_y{}^2 = b^2 s_x{}^2$		Proof of equivalence will not be tested.
16	The term outlier can be applied to data which are: (a) at least 2 standard deviations from the mean; (b) at least 1.5 × IQR beyond the nearer quartile.		

DATA PRESENTATION	
Notation for sample variance and sample standard deviation	
The notations s^2 and s for sample variance and sample standard deviation, respectively, are written into both British Standards (BS3534-1, 1993) and International Standards (ISO 3534). The definitions are those given above in equations (✝) and (§). The calculations are carried out using divisor $(n-1)$. In this specification, the usage will be consistent with these definitions. Thus the meanings of 'sample variance', denoted by s^2, and 'sample standard deviation', denoted by s, are uniquely defined, as calculated with divisor $(n-1)$.	In early work in statistics it is common practice to introduce these concepts with divisor n rather than $(n-1)$. However there is no recognised notation to denote the quantities so derived. In this specification, in order to ensure unambiguity of meaning, these quantities will be referred to by the functional names of 'mean square deviation' and 'root mean square deviation'. The letters *msd* and *rmsd* will be used to denote their values. Students should be aware of the variations in notation used by manufacturers on calculators and know what the symbols on their particular models represent.

STATISTICS 1, S1		
Specification	**Ref.**	**Competence Statements**
PROBABILITY		
Probability of events in a finite sample space.	S1u1	Know how to calculate the probability of one event.
	2	Understand the concept of a complementary event and know that the probability of an event may be found by finding that of its complementary event.
Probability of two or more events which are:	3	Know how to draw sample space diagrams to help calculate probabilities.
	4	Know how to calculate the expected frequency of an event given its probability.
(i) mutually exclusive;	5	Understand the concepts of mutually exclusive events and independent events.
	6	Know to add probabilities for mutually exclusive events.
	7	Know to multiply probabilities for independent events.
	8	Know how to use tree diagrams to assist in the calculation of probabilities.
(ii) not mutually exclusive.	9	Know how to calculate probabilities for two events which are not mutually exclusive.
Conditional probability.	10	Be able to use Venn diagrams to help calculations of probabilities for up to three events.
	11	Know how to calculate conditional probabilities by formula, from tree diagrams or sample space diagrams
	12	Know that $P(B\|A) = P(B) \Leftrightarrow B$ and A are independent.
DISCRETE RANDOM VARIABLES		
Probability distributions.	S1R1	Be able to use probability functions, given algebraically or in tables.
Calculation of probability, expectation (mean) and variance.	2	Be able to calculate the numerical probabilities for a simple distribution.
	3	Be able to calculate the expectation (mean), $E(X)$, in simple cases and understand its meaning.
	4	Be able to calculate the variance, $\mathrm{Var}(X)$, in simple cases.

STATISTICS 1, S1			
Ref.	**Notes**	**Notation**	**Exclusions**
	PROBABILITY		
S1u1			
2		P(A) A' is the event 'Not A'	
3			
4		Expected frequency: n P(A)	
5			Formal notation and definitions.
6	To find P(A or B).		
7	To find P(A and B) Including the use of complementary events. e.g. finding the probability of at least one 6 in five throws of a die.		
8			
9			
10	Candidates should understand, though not necessarily in this form, the relation: $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$.		Probability of a general or infinite number of events. Formal proofs.
11	$\mathrm{P}(A \cap B) = \mathrm{P}(A).\mathrm{P}(B \mid A)$	$\mathrm{P}(B \mid A)$	
12	In this case $\mathrm{P}(A \cap B) = \mathrm{P}(A).\mathrm{P}(B)$.		
	DISCRETE RANDOM VARIABLES		
S1R1	In *S1* questions will only be set on simple finite distributions.		
2		$\mathrm{P}(X = x)$	
3		$\mathrm{E}(X) = \mu$	
4	Knowledge of $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mu^2$.	$\mathrm{Var}(X) = \mathrm{E}[(X - \mu)^2]$	

STATISTICS 1, S1		
Specification	**Ref.**	**Competence Statements**
THE BINOMIAL DISTRIBUTION AND ITS USE IN HYPOTHESIS TESTING		
Situations leading to a binomial distribution.	S1H1	Recognise situations which give rise to a binomial distribution.
	2	Be able to identify the binomial parameter p, the probability of success.
Calculations relating to binomial distribution.	3	Be able to calculate probabilities using the binomial distribution.
	4	Know that ${}^{n}C_{r}$ is the number of ways of selecting r objects from n.
	5	Know that $n!$ is the number of ways of arranging n objects in line.
Knowledge of mean.	6	Understand and apply mean = np.
Calculation of expected frequencies.	7	Be able to calculate the expected frequencies of the various possible outcomes from a series of binomial trials.
Hypothesis testing for a binomial probability p.	8	Understand the process of hypothesis testing and the associated vocabulary.
	9	Be able to identify Null and Alternative Hypotheses (H_0 and H_1) when setting up a hypothesis test on a binomial probability model.
	10	Be able to conduct hypothesis tests at various levels of significance.
	11	Be able to identify the critical and acceptance regions.
	12	Be able to draw a correct conclusion from the results of a hypothesis test on a binomial probability model.
	13	Understand when to apply 1- tail and 2- tail tests.

STATISTICS 1, S1			
Ref.	**Notes**	**Notation**	**Exclusions**
THE BINOMIAL DISTRIBUTION AND ITS USE IN HYPOTHESIS TESTING			
S1H1			
2	As a model for observed data.	B(n, p), $q = 1 - p$ ~ means 'has the distribution'.	
3	Including use of tables of cumulative binomial probabilities.		
4		${}^{n}\mathrm{C}_{r} = \binom{n}{r} = \frac{n!}{(n-r)!r!}$	
5			
6			Formal proof of variance of the binomial distribution.
7			
8	Null hypothesis, alternative hypothesis. Significance level, 1-tail test, 2-tail test. Critical value, critical region, acceptance region.		
9		H_0, H_1	
10			Normal approximation.
11			
12			
13			