6.14 STATISTICS 2, S2 (4767) A2

Objectives

To extend students' ability to represent data in bivariate situations, with an emphasis on linear and rank order modelling, and associated hypothesis testing.

To introduce continuous probability distributions through the Normal distribution.

Assessment

Examination (72 marks)
1 hour 30 minutes
There will be four questions each worth about 18 marks.

Assumed Knowledge

Candidates are expected to know the content of *C1* and *S1*. Candidates also need to know the series expansion of e^x.

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.

STATISTICS 2, S2		
Specification	**Ref.**	**Competence Statements**

BIVARIATE DATA		
Scatter diagram.	S2b1	Know how to draw a scatter diagram.
	2	Know the difference between dependent and independent variables.
Pearson's product moment correlation coefficient (pmcc).	3	Know how to calculate the pmcc from raw data or summary statistics.
	4	Know how to carry out hypothesis tests using the pmcc and tables of critical values.
Spearman's Rank correlation coefficient.	5	Know how to calculate Spearman's rank correlation coefficient from raw data or summary statistics.
	6	Know how to carry out hypothesis tests using Spearman's rank correlation coefficient and tables of critical values.
Regression line for a random variable on a non-random variable.	7	Know how to calculate the equation of the least squares regression line using raw data or summary statistics.
	8	Know the meaning of the term residual and be able to calculate and interpret residuals.

POISSON DISTRIBUTION		
Situations leading to a Poisson distribution.	S2P1	Know the situations under which the Poisson distribution is likely to be an appropriate model.
Calculations of probability and of expected frequencies.	2	Be able to calculate the probabilities within a Poisson distribution.
	3	Be able to use the Poisson distribution as an approximation to the binomial distribution, and know when to do so.
The mean and variance of the Poisson distribution.	4	Know the mean and variance of a Poisson distribution.
The sum of independent Poisson distributions.	5	Know that the sum of two or more independent Poisson distributions is also a Poisson distribution.

STATISTICS 2, S2			
Ref.	**Notes**	**Notation**	**Exclusions**

BIVARIATE DATA

Ref.	Notes	Notation	Exclusions
S2b1			
2			
3		Sample value *r*.	
4	Only ‘H_0: No correlation’ will be tested. Hypothesis tests using Pearson’s product moment correlation coefficient require a modelling assumption that the data are drawn from a bivariate Normal distribution. This may be recognised on a scatter diagram by an approximately elliptical distribution of points. Candidates will not be required to know the formal meaning of bivariate Normality but will be expected to know that both variables must be random and that where one or both of the distributions is skewed, bimodal, etc., the procedure is likely to be inaccurate. They will also be expected to recognise (from a scatter diagram) cases of non-linear association and, where appropriate, to apply a test based on Spearman’s correlation coefficient.		
5		Sample value r_s .	
6	Only ‘H_0: No association’ will be tested. Hypothesis tests using Spearman’s rank correlation coefficient require no modelling assumption about the underlying distribution.		
7	The goodness of fit of a regression line may be judged by looking at the scatter diagram. Examination questions will be confined to cases in which a random variable, *Y*, and a non-random variable, *x*, are modelled by a relationship in which the expected value of *Y* is a linear function of *x*.		Derivation of the least squares regression line.
8	Informal checking of a model by looking at residuals.		

POISSON DISTRIBUTION

Ref.	Notes	Notation	Exclusions
S2P1		$X \sim \text{Poisson}(\lambda)$	
2	Including use of tables of cumulative Poisson probabilities.		
3			
4			Formal proof.
5	$X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\mu)$ $\Rightarrow X + Y \sim \text{Poisson}(\lambda + \mu)$ when *X* and *Y* are independent.		Formal proofs.

STATISTICS 2, S2		
Specification	**Ref.**	**Competence Statements**

CONTINGENCY TABLES

The χ^2 test.	S2H1	Be able to apply the χ^2 test to a contingency table.
	2	Be able to interpret the results of a χ^2 test.

NORMAL DISTRIBUTION

The use of the Normal distribution.	S2N1	Be able to use the Normal distribution as a model.
	2	Be able to standardise a Normal variable and use the Normal distribution tables.
The use of the Normal distribution as an approximation to the binomial and Poisson distributions.	3	Be able to use the Normal distribution as an approximation to the binomial distribution and know when it is appropriate to do so.
	4	Be able to use the Normal distribution as an approximation to the Poisson distribution and know when it is appropriate to do so.
	5	Know when to use a continuity correction and be able to do so.
Hypothesis test for a single mean.	6	Be able to carry out a hypothesis test for a single mean using the Normal distribution and know when it is appropriate to do so.

STATISTICS 2, S2			
Ref.	**Notes**	**Notation**	**Exclusions**

CONTINGENCY TABLES

S2H1	The use of Yates' continuity correction for 2×2 contingency tables will not be required in examination questions. Candidates who use it appropriately will, however, be eligible for full marks.		
2	This may involve considering the individual cells in the contingency table.		

NORMAL DISTRIBUTION

S2N1			
2		$X \sim \mathrm{N}(\mu, \sigma^2)$	Proof.
3	Variance of B(n, p) is npq B(n, p) ≈ N(np, npq)		Proof.
4	Poisson (λ) ≈ N(λ, λ)		
5			
6	In situations where either (a) the population variance is known or (b) the population variance is unknown but the sample size is large $\mathrm{E}(\bar{X}) = \mu,\ \mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}$.		