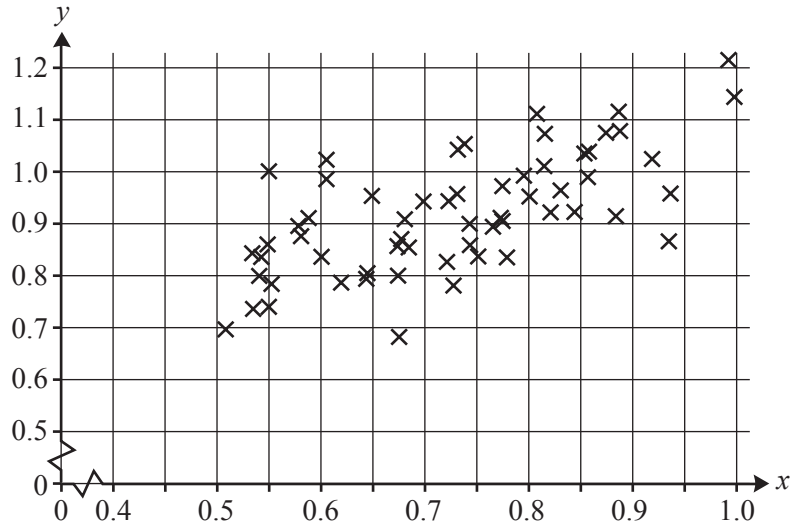


- 1 Salbutamol is a drug used to improve lung function. In a medical trial, a random sample of 60 people with impaired lung function was selected. The forced expiratory volume in one second (FEV1) was measured for each person, both before being given salbutamol and again after a two-week course of the drug. The variables x and y , measured in suitable units, represent FEV1 before and after the two-week course respectively. The data are illustrated in the scatter diagram below, together with the summary statistics for these data.



Summary statistics:

$$n = 60, \quad \Sigma x = 43.62, \quad \Sigma y = 55.15, \quad \Sigma x^2 = 32.68, \quad \Sigma y^2 = 51.44, \quad \Sigma xy = 40.66$$

- (i) Calculate the sample product moment correlation coefficient. [5]
- (ii) Carry out a hypothesis test at the 5% significance level to investigate whether there is positive correlation between FEV1 before and after the course. [6]
- (iii) State the distributional assumption which is necessary for this test to be valid. State, with a reason, whether the assumption appears to be valid. [2]
- (iv) Explain the meaning of the term 'significance level'. [2]
- (v) Calculate the values of the summary statistics if the data point $x = 0.55$, $y = 1.00$ had been incorrectly recorded as $x = 1.00$, $y = 0.55$. [3]

- 2 Suppose that 3% of the population of a large city have red hair.

- (i) A random sample of 10 people from the city is selected. Find the probability that there is at least one person with red hair in this sample. [2]

A random sample of 60 people from the city is selected. The random variable X represents the number of people in this sample who have red hair.

- (ii) Explain why the distribution of X may be approximated by a Poisson distribution. Write down the mean of this Poisson distribution. [3]

- (iii) Hence find

(A) $P(X = 2)$, [2]

(B) $P(X > 2)$. [2]

- (iv) Discuss whether or not it would be appropriate to model X using a Normal approximating distribution. [2]

A random sample of 5000 people from the city is selected.

- (v) State the exact distribution of the number of people with red hair in the sample. [2]

- (vi) Use a suitable Normal approximating distribution to find the probability that there are at least 160 people with red hair in the sample. [5]

- 3 The scores, X , in Paper 1 of an English examination have an underlying Normal distribution with mean 76 and standard deviation 12. The scores are reported as integer marks. So, for example, a score for which $75.5 \leq X < 76.5$ is reported as 76 marks.

- (i) Find the probability that a candidate's reported mark is 76. [4]

- (ii) Find the probability that a candidate's reported mark is at least 80. [3]

- (iii) Three candidates are chosen at random. Find the probability that exactly one of these three candidates' reported marks is at least 80. [2]

The proportion of candidates who receive an A* grade (the highest grade) must not be greater than 10% but should be as close as possible to 10%.

- (iv) Find the lowest reported mark that should be awarded an A* grade. [5]

The scores in Paper 2 of the examination have an underlying Normal distribution with mean μ and standard deviation 12.

- (v) Given that 20% of candidates receive a reported mark of 50 or less, find the value of μ . [4]

- 4 An art gallery is holding an exhibition. A random sample of 150 visitors to the exhibition is selected. The visitors are asked which of four artists they prefer. Their preferences, classified according to whether the visitor is female or male, are given in the table.

		Artist preferred			
		Monet	Renoir	Degas	Cézanne
Sex	Male	8	25	18	19
	Female	18	35	10	17

- (i) Carry out a test at the 10% significance level to examine whether there is any association between artist preferred and sex of visitor. Your working should include a table showing the contributions of each cell to the test statistic. [12]
- (ii) For each artist, comment briefly on how the preferences of each sex compare with what would be expected if there were no association. [6]

Copyright Information

OCR is committed to seeking permission to reproduce all third-party content that it uses in its assessment materials. OCR has attempted to identify and contact all copyright holders whose work is used in this paper. To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced in the OCR Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download from our public website (www.ocr.org.uk) after the live examination series. If OCR has unwittingly failed to correctly acknowledge or clear any third-party content in this assessment material, OCR will be happy to correct its mistake at the earliest possible opportunity.

For queries or further information please contact the Copyright Team, First Floor, 9 Hills Road, Cambridge CB2 1GE.

OCR is part of the Cambridge Assessment Group; Cambridge Assessment is the brand name of University of Cambridge Local Examinations Syndicate (UCLES), which is itself a department of the University of Cambridge.