



Maths Questions By Topic:

Statistical Hypothesis Testing

A-Level Edexcel

 0207 060 4494

 www.expert-tuition.co.uk

 online.expert-tuition.co.uk

 enquiries@expert-tuition.co.uk

 The Foundry, 77 Fulham Palace Road, W6 8JA

Table Of Contents

New Spec

Paper 2 (AS)	Page 1
Paper 3 (A2)	Page 20

Old Spec

Statistics 2	Page 58
--------------------	---------

10. Marc took a random sample of 16 students from a school and for each student recorded
- the number of letters, x , in their last name
 - the number of letters, y , in their first name

The results from Marc’s random sample of 16 observations are given in the table below.

x	3	6	8	7	5	3	11	3	4	5	4	9	7	10	6	6
y	7	7	4	4	6	8	5	5	8	4	7	4	5	5	6	3

- (a) Use your calculator to find the product moment correlation coefficient between x and y for these data. (1)

- (b) Test whether or not there is evidence of a negative correlation between the number of letters in the last name and the number of letters in the first name.

You should

- state your hypotheses clearly
 - use a 5% level of significance
- (3)

12. A random sample of 15 days is taken from the large data set for Perth in June and July 1987. The scatter diagram in Figure 1 displays the values of two of the variables for these 15 days.

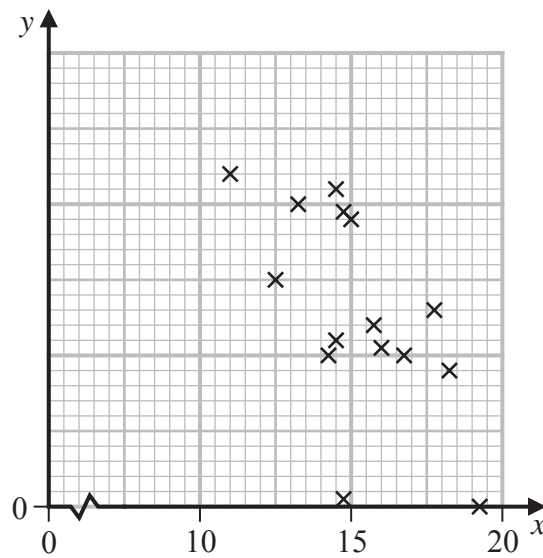


Figure 1

Stav believes that there is a correlation between Daily Total Sunshine and Daily Maximum Relative Humidity at Heathrow.

He calculates the product moment correlation coefficient between these two variables for a random sample of 30 days and obtains $r = -0.377$

Carry out a suitable test to investigate Stav's belief at a 5% level of significance. State clearly

- your hypotheses
- your critical value

(3)

16. Tessa owns a small clothes shop in a seaside town. She records the weekly sales figures, £ w , and the average weekly temperature, $t^{\circ}\text{C}$, for 8 weeks during the summer. The product moment correlation coefficient for these data is -0.915

(a) Stating your hypotheses clearly and using a 5% level of significance, test whether or not the correlation between sales figures and average weekly temperature is negative. (3)

(b) Suggest a possible reason for this correlation. (1)

Tessa suggests that a linear regression model could be used to model these data.

(c) State, giving a reason, whether or not the correlation coefficient is consistent with Tessa's suggestion. (1)

(d) State, giving a reason, which variable would be the explanatory variable. (1)

Tessa calculated the linear regression equation as $w = 10\,755 - 171t$

(e) Give an interpretation of the gradient of this regression equation. (1)

17. The lifetime, L hours, of a battery has a normal distribution with mean 18 hours and standard deviation 4 hours.
Alice's calculator requires 4 batteries and will stop working when any one battery reaches the end of its lifetime.
After her exams, Alice believed that the lifetime of the batteries was more than 18 hours. She took a random sample of 20 of these batteries and found that their mean lifetime was 19.2 hours.

Stating your hypotheses clearly and using a 5% level of significance, test Alice's belief.

(5)

18. *Kaff* coffee is sold in packets. A seller measures the masses of the contents of a random sample of 90 packets of *Kaff* coffee from her stock. The results are shown in the table below.

Mass w (g)	Midpoint y (g)	Frequency f
$240 \leq w < 245$	242.5	8
$245 \leq w < 248$	246.5	15
$248 \leq w < 252$	250.0	35
$252 \leq w < 255$	253.5	23
$255 \leq w < 260$	257.5	9

(You may use $\sum fy^2 = 5\,644\,171.75$)

A histogram is drawn and the class $245 \leq w < 248$ is represented by a rectangle of width 1.2 cm and height 10 cm.

- (a) Calculate the width and the height of the rectangle representing the class $255 \leq w < 260$. (3)
- (b) Use linear interpolation to estimate the median mass of the contents of a packet of *Kaff* coffee to 1 decimal place. (2)
- (c) Estimate the mean and the standard deviation of the mass of the contents of a packet of *Kaff* coffee to 1 decimal place. (3)

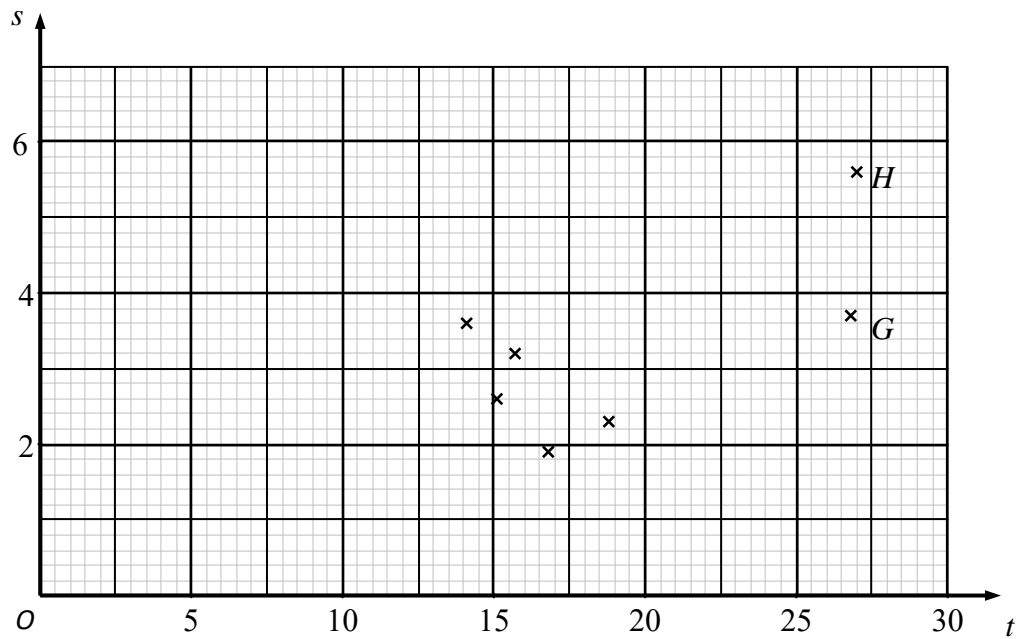
The seller claims that the mean mass of the contents of the packets is more than the stated mass. Given that the stated mass of the contents of a packet of *Kaff* coffee is 250 g and the actual standard deviation of the contents of a packet of *Kaff* coffee is 4 g,

- (d) test, using a 5% level of significance, whether or not the seller's claim is justified. State your hypotheses clearly.

(You may assume that the mass of the contents of a packet is normally distributed.)

(5)

19. A researcher believes that there is a linear relationship between daily mean temperature and daily total rainfall. The 7 places in the northern hemisphere from the large data set are used. The mean of the daily mean temperatures, t °C, and the mean of the daily total rainfall, s mm, for the month of July in 2015 are shown on the scatter diagram below.



The researcher calculated the product moment correlation coefficient for the 7 places and obtained $r = 0.658$.

Stating your hypotheses clearly, test at the 10% level of significance, whether or not the product moment correlation coefficient for the population is greater than zero.

(3)

20. A meteorologist believes that there is a relationship between the daily mean windspeed, w kn, and the daily mean temperature, t °C. A random sample of 9 consecutive days is taken from past records from a town in the UK in July and the relevant data is given in the table below.

t	13.3	16.2	15.7	16.6	16.3	16.4	19.3	17.1	13.2
w	7	11	8	11	13	8	15	10	11

The meteorologist calculated the product moment correlation coefficient for the 9 days and obtained $r = 0.609$

- (a) State what is measured by the product moment correlation coefficient. (1)
- (b) Stating your hypotheses clearly test, at the 5% significance level, whether or not the product moment correlation coefficient for the population is greater than zero. (3)

22. Past records show that the proportion of customers buying organic vegetables from *Tesson* supermarket is 0.35

During a particular day, a random sample of 40 customers from *Tesson* supermarket was taken and 18 of them bought organic vegetables.

- (a) Test, at the 5% level of significance, whether or not this provides evidence that the proportion of customers who bought organic vegetables has increased. State your hypotheses clearly.

(5)

The manager of *Tesson* supermarket claims that the proportion of customers buying organic eggs is different from the proportion of those buying organic vegetables. To test this claim the manager decides to take a random sample of 50 customers.

- (b) Using a 5% level of significance, find the critical region to enable the *Tesson* supermarket manager to test her claim. The probability for each tail of the region should be as close as possible to 2.5%

(3)

During a particular day, a random sample of 50 customers from *Tesson* supermarket is taken and 8 of them bought organic eggs.

- (c) Using your answer to part (b), state whether or not this sample supports the manager's claim. Use a 5% level of significance.

(1)

- (d) State the actual significance level of this test.

(1)

The proportion of customers who buy organic fruit from *Tesson* supermarket is 0.2. During a particular day, a random sample of 200 customers from *Tesson* supermarket is taken. Using a suitable approximation, the probability that fewer than n of these customers bought organic fruit is 0.0465 correct to 4 decimal places.

- (e) Find the value of n .

(6)

23. A potter believes that 20% of pots break whilst being fired in a kiln. Pots are fired in batches of 25.

- (a) Let X denote the number of broken pots in a batch. A batch is selected at random. Using a 10% significance level, find the critical region for a two tailed test of the potter’s belief. You should state the probability in each tail of your critical region. **(4)**

The potter aims to reduce the proportion of pots which break in the kiln by increasing the size of the batch fired. He now fires pots in batches of 50. He then chooses a batch at random and discovers there are 6 pots which broke whilst being fired in the kiln.

- (b) Test, at the 5% level of significance, whether or not there is evidence that increasing the number of pots in a batch has reduced the percentage of pots that break whilst being fired in the kiln. State your hypotheses clearly. **(5)**

25. The proportion of houses in Radville which are unable to receive digital radio is 25%. In a survey of a random sample of 30 houses taken from Radville, the number, X , of houses which are unable to receive digital radio is recorded.

A radio company claims that a new transmitter set up in Radville will reduce the proportion of houses which are unable to receive digital radio. After the new transmitter has been set up, a random sample of 15 houses is taken, of which 1 house is unable to receive digital radio.

Test, at the 10% level of significance, the radio company's claim. State your hypotheses clearly.

(5)

26. Before Roger will use a tennis ball he checks it using a ‘‘bounce’’ test. The probability that a ball from Roger’s usual supplier fails the bounce test is 0.2. A new supplier claims that the probability of one of their balls failing the bounce test is less than 0.2. Roger checks a random sample of 40 balls from the new supplier and finds that 3 balls fail the bounce test.

Stating your hypotheses clearly, use a 5% level of significance to test the new supplier’s claim.

(5)

27. In a manufacturing process 25% of articles are thought to be defective. Articles are produced in batches of 20

- (a) A batch is selected at random. Using a 5% significance level, find the critical region for a two tailed test that the probability of an article chosen at random being defective is 0.25

You should state the probability in each tail which should be as close as possible to 0.025

(5)

The manufacturer changes the production process to try to reduce the number of defective articles. She then chooses a batch at random and discovers there are 3 defective articles.

- (b) Test at the 5% level of significance whether or not there is evidence that the changes to the process have reduced the percentage of defective articles. State your hypotheses clearly.

(5)

29. A test statistic has a distribution $B(25, p)$.

Given that

$$H_0 : p = 0.5 \quad H_1 : p \neq 0.5$$

- (a) find the critical region for the test statistic such that the probability in each tail is as close as possible to 2.5%. (3)

- (b) State the probability of incorrectly rejecting H_0 using this critical region. (2)

31. A shopkeeper knows, from past records, that 15% of customers buy an item from the display next to the till. After a refurbishment of the shop, he takes a random sample of 30 customers and finds that only 1 customer has bought an item from the display next to the till.

- (a) Stating your hypotheses clearly, and using a 5% level of significance, test whether or not there has been a change in the proportion of customers buying an item from the display next to the till.

(6)

During the refurbishment a new sandwich display was installed. Before the refurbishment 20% of customers bought sandwiches. The shopkeeper claims that the proportion of customers buying sandwiches has now increased. He selects a random sample of 120 customers and finds that 31 of them have bought sandwiches.

- (b) Using a suitable approximation and stating your hypotheses clearly, test the shopkeeper's claim. Use a 10% level of significance.

(8)

32. A student takes a multiple choice test. The test is made up of 10 questions each with 5 possible answers. The student gets 4 questions correct. Her teacher claims she was guessing the answers. Using a one tailed test, at the 5% level of significance, test whether or not there is evidence to reject the teacher's claim.
State your hypotheses clearly.

(6)

(Total 6 marks)

33. A company claims that a quarter of the bolts sent to them are faulty. To test this claim the number of faulty bolts in a random sample of 50 is recorded.

(a) Give two reasons why a binomial distribution may be a suitable model for the number of faulty bolts in the sample. (2)

(b) Using a 5% significance level, find the critical region for a two-tailed test of the hypothesis that the probability of a bolt being faulty is $\frac{1}{4}$. The probability of rejection in either tail should be as close as possible to 0.025 (3)

(c) Find the actual significance level of this test. (2)

In the sample of 50 the actual number of faulty bolts was 8.

(d) Comment on the company's claim in the light of this value. Justify your answer. (2)

The machine making the bolts was reset and another sample of 50 bolts was taken. Only 5 were found to be faulty.

(e) Test at the 1% level of significance whether or not the probability of a faulty bolt has decreased. State your hypotheses clearly. (6)

34. (a) Define the critical region of a test statistic. (2)

A discrete random variable X has a Binomial distribution $B(30, p)$. A single observation is used to test $H_0 : p = 0.3$ against $H_1 : p \neq 0.3$

(b) Using a 1% level of significance find the critical region of this test. You should state the probability of rejection in each tail which should be as close as possible to 0.005 (5)

(c) Write down the actual significance level of the test. (1)

The value of the observation was found to be 15.

(d) Comment on this finding in light of your critical region. (2)

Question 34 continued

Lined area for writing the answer to Question 34.

(Total 10 marks)

35. Past records suggest that 30% of customers who buy baked beans from a large supermarket buy them in single tins. A new manager questions whether or not there has been a change in the proportion of customers who buy baked beans in single tins. A random sample of 20 customers who had bought baked beans was taken.

(a) Using a 10% level of significance, find the critical region for a two-tailed test to answer the manager’s question. You should state the probability of rejection in each tail which should be less than 0.05.

(5)

(b) Write down the actual significance level of a test based on your critical region from part (a).

(1)

The manager found that 11 customers from the sample of 20 had bought baked beans in single tins.

(c) Comment on this finding in the light of your critical region found in part (a).

(2)

36. A single observation x is to be taken from a Binomial distribution $B(20, p)$.

This observation is used to test $H_0 : p = 0.3$ against $H_1 : p \neq 0.3$

(a) Using a 5% level of significance, find the critical region for this test.
The probability of rejecting either tail should be as close as possible to 2.5%. **(3)**

(b) State the actual significance level of this test. **(2)**

The actual value of x obtained is 3.

(c) State a conclusion that can be drawn based on this value giving a reason for your answer. **(2)**

(Total 7 marks)

37. Sue throws a fair coin 15 times and records the number of times it shows a head.

(a) State the distribution to model the number of times the coin shows a head. **(2)**

Find the probability that Sue records

(b) exactly 8 heads, **(2)**

(c) at least 4 heads. **(2)**

Sue has a different coin which she believes is biased in favour of heads. She throws the coin 15 times and obtains 13 heads.

(d) Test Sue's belief at the 1% level of significance. State your hypotheses clearly. **(6)**
