

# OCR

Oxford Cambridge and RSA

Accredited

## AS Level Further Mathematics B (MEI)

### Y416 Statistics b

#### Sample Question Paper

## Date – Morning/Afternoon

Time allowed: 1 hour 15 minutes

#### OCR supplied materials:

- Printed Answer Booklet
- Formulae Further Mathematics B (MEI)

#### You must have:

- Printed Answer Booklet
- Formulae Further Mathematics B (MEI)
- Scientific or graphical calculator



#### INSTRUCTIONS

- Use black ink. HB pencil may be used for graphs and diagrams only.
- Complete the boxes provided on the Printed Answer Booklet with your name, centre number and candidate number.
- Answer **all** the questions.
- **Write your answer to each question in the space provided in the Printed Answer Booklet.**
- Additional paper may be used if necessary but you must clearly show your candidate number, centre number and question number(s).
- Do **not** write in the bar codes.
- You are permitted to use a scientific or graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.

#### INFORMATION

- The total number of marks for this paper is **60**.
- The marks for each question are shown in brackets [ ].
- You are advised that an answer may receive **no marks** unless you show sufficient detail of the working to indicate that a correct method is used. You should communicate your method with correct reasoning.
- The Printed Answer Booklet consists of **12** pages. The Question Paper consists of **8** pages.

## 2

Answer **all** the questions.

- 1 Abby runs a stall at a charity event. Visitors to the stall pay to play a game in which six fair dice are rolled. If the difference between the highest and lowest scores is less than 3 then the player wins £5. Otherwise the player wins nothing.

Abby designs the spreadsheet shown in Fig. 1 to estimate the probability of a player winning, by simulating 20 goes at the game. Cell C5, highlighted, shows that the 2nd dice in simulated game 4 scores 5. Cells H5 and I5 show the highest and lowest scores, respectively, in game 4, and cell J5 gives the difference between them.

		=RANDBETWEEN(1,6)								
	A	B	C	D	E	F	G	H	I	J
1		dice 1	dice 2	dice 3	dice 4	dice 5	dice 6	High score	Low score	Difference
2	game 1	2	2	4	2	3	3	4	2	2
3	game 2	2	6	3	2	1	2	6	1	5
4	game 3	3	1	5	3	4	6	6	1	5
5	game 4	6	5	2	5	6	3	6	2	4
6	game 5	6	3	3	5	3	2	6	2	4
7	game 6	5	6	3	5	1	4	6	1	5
8	game 7	2	3	1	2	6	4	6	1	5
9	game 8	6	6	6	6	1	5	6	1	5
10	game 9	3	6	2	5	4	1	6	1	5
11	game 10	5	1	1	4	6	1	6	1	5
12	game 11	2	5	6	1	6	5	6	1	5
13	game 12	2	5	6	6	6	6	6	2	4
14	game 13	2	2	2	2	4	4	4	2	2
15	game 14	1	6	6	6	3	5	6	1	5
16	game 15	2	2	3	3	5	1	5	1	4
17	game 16	1	2	3	4	3	3	4	1	3
18	game 17	5	2	4	2	1	6	6	1	5
19	game 18	6	1	5	2	1	5	6	1	5
20	game 19	1	3	5	1	3	5	5	1	4
21	game 20	5	4	3	2	5	1			

Fig. 1

## 3

- (i) (A) Write down the numbers in columns H, I and J for game 20. [1]
- (B) Use the spreadsheet to estimate the probability of a player winning a game. [2]
- (ii) State how the estimate of probability in (i) (B) could be improved. [1]
- (iii) Give one advantage and one disadvantage of using this simulation technique compared with working out the theoretical probability. [2]

All profit made by the stall is given to charity. Abby has to decide how much to charge players to play.

- (iv) If Abby charges £1 per game, estimate the total profit when 50 players each play the game once. [3]

- 2 The cumulative distribution function of the continuous random variable,  $Y$ , is given below.

$$F(y) = \begin{cases} 0 & y < 0 \\ \frac{y^3 - y^2}{4} & 1 \leq y \leq 2 \\ 1 & y > 2 \end{cases}$$

- (i) Find  $P(Y \leq 1.5)$  [2]
- (ii) Verify that the median of  $Y$  lies between 1.6 and 1.7. [3]
- (iii) Find the probability density function of  $Y$ . [2]

- 3 At a factory, flour is packed into bags. A model for the mass in grams of flour packed into each bag is  $1500 + X$ , where  $X$  is a continuous random variable with probability density function

$$f(x) = \begin{cases} kx(6-x) & 0 \leq x \leq 6 \\ 0 & \text{elsewhere,} \end{cases}$$

where  $k$  is a constant.

- (i) Show that  $k = \frac{1}{36}$ . [2]
- (ii) Find the probability that a randomly selected bag of flour contains 1505 grams of flour or more. [3]
- (iii) Find
- the mean of  $X$ ,
  - the standard deviation of  $X$ . [6]

## 4

- 4 An online encyclopedia claims that the average mass of an adult European hedgehog is 720 g. In an investigation to check this average figure, the masses in grams of twelve randomly chosen adult European hedgehogs are measured and shown below.

705	730	720	691	718	680
731	723	745	708	724	736

(i) What assumption is required to carry out a Wilcoxon test in this situation? [1]

(ii) Given that this assumption is met, carry out a 2-tail Wilcoxon test at the 5% level to test whether the median mass is 720 g. You should state your hypotheses and complete the table of calculations in the Printed Answer Booklet. [7]

- 5 A particular alloy of bronze is specified as containing 11.5% copper on average. A researcher takes a random sample of 14 specimens of this bronze and undertakes an analysis of each of them. The percentages of copper are found to be as follows.

11.12	11.29	11.42	11.43	11.20	11.25	11.65
11.33	11.56	11.34	11.44	11.24	11.60	11.52

The researcher uses software to draw a Normal probability plot for these data and to conduct a Kolmogorov-Smirnov test for Normality. The output is shown in Fig 5.1.

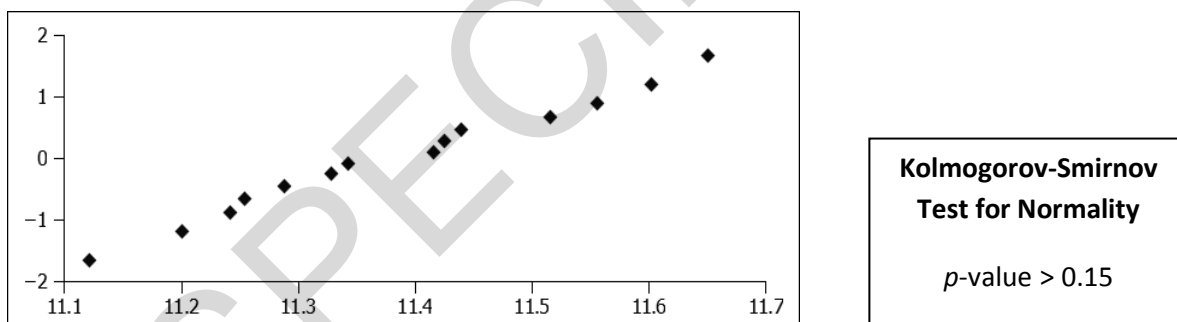
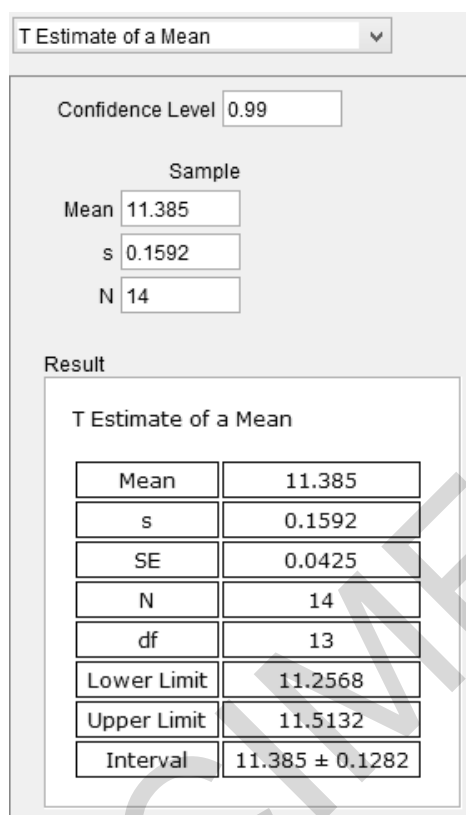


Fig 5.1

- (i) Comment on what the Normal probability plot and the  $p$ -value of the test suggest about the data. [3]

## 5

The researcher uses software to produce a 99% confidence interval for the mean percentage of copper in the alloy, based on the  $t$  distribution. The output from the software is shown in Fig 5.2.



**Fig 5.2**

- (ii) State the confidence interval which the software gives, in the form  $a < \mu < b$ . [1]
- (iii) (A) State an assumption necessary for the use of the  $t$  distribution in the construction of this confidence interval. [1]
- (B) State whether the assumption in part (iii) (A) seems reasonable. [1]
- (iv) Does the confidence interval suggest that the copper content is different from 11.5%, on average? Explain your answer. [2]
- (v) In the output from the software shown in Fig 5.2, SE stands for ‘standard error’.
- (A) Explain what a standard error is. [1]
- (B) Show how the standard error was calculated in this case. [1]
- (vi) Suggest a way in which the researcher could produce a narrower confidence interval. [1]

## 6

- 6 The table below shows the mean and variance of the test scores of a random samples of 70 girls who are starting an A level Mathematics course.

Sample mean	Sample variance
118.86	86.57

- (i) Showing your working, find a 95% confidence interval for the population mean. [4]
- (ii) Explain why you can construct the interval in part (i) despite no information about the distribution of the parent population being given. [2]
- (iii) The same random sample of girls repeats the test. The mean improvement in score is 0.9. The 95% confidence interval for the improvement is  $[-1.5, 3.3]$ . What is the sample variance for the improvement in score? [2]

- 7 Two flatmates work at the same location. One of them takes the bus to work and the other one cycles. Journey times, measured in minutes, are distributed as follows.

- By bus: Normally distributed with mean 23 and standard deviation 6
- By bicycle: Normally distributed with mean 21 and standard deviation 2

You should assume that all journey times are independent.

- (i) One morning the two flatmates set out at the same time. Find the probability that the person who takes the bus arrives before the cyclist. [3]
- (ii) Find the probability that the total time taken for 5 bus journeys is less than 2 hours. [2]
- (iii) Comment on the assumption that all journey times are independent. [1]

**END OF QUESTION PAPER**

SPECIMEN

SPECIMEN

---

Copyright Information:

OCR is committed to seeking permission to reproduce all third-party content that it uses in the 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](http://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.