OCR
Oxford Cambridge and RSA

AS Level Further Mathematics B (MEI)

Y416/01 Statistics b

Practice Paper – Set 2

Time allowed: 1 hour 15 minutes

You must have:
- Printed Answer Booklet
- Formulae Further Mathematics B (MEI)

You may use:
- a 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.** If additional space is required, you should use the lined page(s) at the end of the Printed Answer Booklet. The question number(s) must be clearly shown.
- Do **not** write in the barcodes.
- 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 mark 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.

Answer **all** the questions.

1 A recent survey found that hourly rates of pay for adult employees in the UK had mean £15.43 and standard deviation £6.35. The minimum legal hourly rate of pay for any age group at this time was £5.60.

(i) Explain why a Normal distribution would not be suitable for modelling the hourly rate of pay for adults. **[2]**

(ii) A random sample of 40 adult employees is selected. Find the probability that the mean hourly rate of pay of these 40 people is greater than £16. **[3]**

(iii) Explain why you can find the probability in part **(ii)** despite the fact that the distribution of hourly rates of pay is not known. **[3]**

2 The cumulative distribution function of the random variable X is given by

$$\mathrm{F}(x) = \begin{cases} 0 & x < 0, \\ k(10x - x^2) & 0 \leqslant x \leqslant 5, \\ 1 & x > 5, \end{cases}$$

where k is a constant.

(i) Show that $k = \frac{1}{25}$. **[2]**

(ii) Find $\mathrm{P}(X < 1)$. **[2]**

(iii) Find the probability density function of X. **[2]**

(iv) Sketch the graph of the probability density function of X. **[2]**

(v) Explain why the mode is zero. **[1]**

(vi) Find

- the mean of X,
- the variance of X. **[5]**

3 On the bottles of a brand of shower gel it states that the volume in the bottle is 250 ml. In fact the volume is Normally distributed with mean 251.4 ml and standard deviation 4.1 ml. Regulations require that the average content of a bottle is at least 250 ml and that no more than 5% of bottles contain less than 241 ml.

(i) Show that the requirement about the proportion of bottles containing at least 241 ml is satisfied. **[1]**

The machine that fills the bottles undergoes maintenance. Following this, a random sample of 50 filled bottles is selected and the contents measured, in order to check whether the mean volume is still 251.4 ml. Software is used to carry out a hypothesis test at the 5% significance level. The output from the software is shown in Fig. 2.

Z Test of a Mean

Null Hypothesis μ = 251.4

Alternative Hypothesis ○ < ○ > ◉ ≠

Sample

Mean 249.9

σ 5.52

N 50

Result

Z Test of a Mean

Mean	249.9
σ	5.52
SE	0.7806
N	50
Z	-1.9215
P	0.0547

Fig. 2

(ii) State the meaning of the symbol μ in the output. **[1]**

(iii) Use the output from the software to complete the test. **[2]**

(iv) Using the same data and significance level, the software could have been set to carry out a test with alternative hypothesis $\mu < 251.4$ instead of $\mu \neq 251.4$. Explain what the result of this test would have been. **[2]**

4 An advert for a new sports drink claims that it improves performance more than traditional sports drinks. A sports coach wants to check if this claim is true. He selects ten of the people he coaches. On two separate occasions he asks them to carry out 50 repetitions of a particular exercise as quickly as possible. Each person drinks a bottle of the new sports drink on one occasion and a bottle of traditional sports drink on the other occasion. They are then timed while they do the exercises.

Fig. 4.1 is a spreadsheet showing the times in seconds taken for each exercise using the two drinks, together with the differences in the times.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Person	A	B	C	D	E	F	G	H	I	J	
2	Traditional	47.0	55.4	57.6	48.3	53.2	61.4	65.8	58.2	74.1	43.8	
3	New	46.2	52.7	56.4	48.9	51.6	62.8	61.6	57.6	73.1	42.6	
4	Difference	0.8	2.7	1.2	-0.6	1.6	-1.4	4.2	0.6	1.0	1.2	
5												

Fig. 4.1

The coach wants to find a 95% confidence interval for the mean time difference.

(i) (*A*) State a distributional assumption necessary in order to construct a valid confidence interval. **[2]**

(*B*) State a property of the sample required in order to construct a valid confidence interval. **[1]**

In order to check whether the property in part **(i)**(*A*) holds, the coach carries out a Kolmogorov-Smirnov test and also uses the spreadsheet to produce a Normal probability plot. These are shown in Fig. 4.2.

Kolmogorov-Smirnov Test	
p-value	0.455
n	10

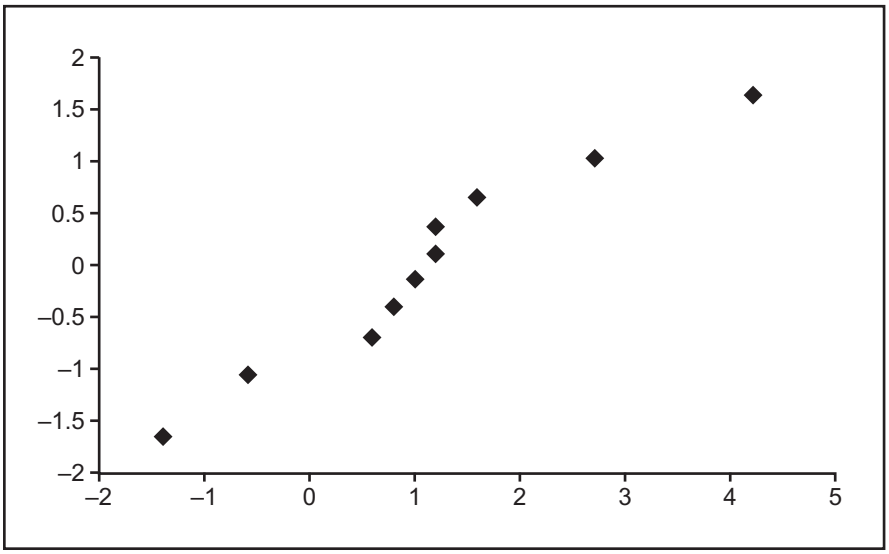

Fig. 4.2

(ii) Explain how the Kolmogorov-Smirnov test and the Normal probability plot suggest that the assumption in part **(i)**(*A*) is justified. **[2]**

You should now assume that the conditions required in order to construct a valid confidence interval are satisfied. You are given that the sample mean and sample standard deviation of the differences are 1.13 seconds and 1.56 seconds respectively.

(iii) Showing your working, find a 95% confidence interval for the mean time difference. **[5]**

(iv) Explain why the confidence interval suggests that the claim in the advert may be true. **[2]**

5 In an Olympic distance triathlon, competitors swim 1.5 km, cycle 40 km and run 10 km. The times in minutes taken by a randomly chosen competitor for the swimming, cycling and running stages are denoted by S, C and R respectively. These times are modelled by independent Normal distributions with means and variances as shown in the table.

	Mean	Variance
Swimming time, S minutes	33.36	39.12
Cycling time, C minutes	84.35	154.06
Running time, R minutes	53.88	81.42

(i) Find the probability that a randomly chosen competitor's total time for all three stages exceeds 3 hours. **[3]**

(ii) Find $\mathrm{P}(C > R + 30)$. **[2]**

(iii) Find $\mathrm{P}(C < 3S)$. **[2]**

(iv) Comment briefly on the assumption that S, C and R are independent. **[1]**

6 A student is investigating continuous uniform distributions, using three independent random variables X, Y and Z. Both X and Y have a continuous uniform distribution over [0, 10] and Z has a continuous uniform distribution over [0, 20].

(i) The value of $Z - X - Y$ always lies between a and b. Write down the values of a and b. **[2]**

The student wants to know whether there is any difference between $P(Z - X - Y > 5)$ and $P(Z - 2X > 5)$. She designs the spreadsheet shown in Fig. 6 to simulate 30 values of X, Y and Z together with the corresponding values of $Z - X - Y$ and $Z - 2X$. She uses the spreadsheet to calculate the mean and standard deviation of $Z - X - Y$ and $Z - 2X$ for the 30 simulated values. Some rows of the spreadsheet have been deliberately omitted.

	A	B	C	D	E
1	***X***	***Y***	***Z***	***Z−X−Y***	***Z−2X***
2	6.06	1.76	8.65	0.82	-3.47
3	0.11	9.45	2.06	-7.51	1.83
4	2.63	1.21	12.88	9.04	7.63
29	2.00	7.46	14.59	5.14	10.59
30	3.69	7.68	6.60	-4.77	-0.78
31	8.35	6.67	5.84	-9.19	-10.86
32					
33		Number > 5		5	8
34					
35			Mean	-2.64	-2.21
36			SD	7.07	8.69
37					

Fig. 6

(ii) Use the values in the spreadsheet to estimate each of the following.

- $P(Z - X - Y > 5)$
- $P(Z - 2X > 5)$

[2]

(iii) Show that $E(Z - X - Y)$ and $E(Z - 2X)$ are both 0. **[3]**

(iv) The student repeats the simulation several times and notices that the standard deviation of $Z - X - Y$ is usually less than that of $Z - 2X$. Show why this is the case. **[5]**

END OF QUESTION PAPER

BLANK PAGE