A Level Further Mathematics A
Y542/01 Statistics

Practice Paper – Set 3
Time allowed: 1 hour 30 minutes

You must have:
• Printed Answer Booklet
• Formulae A Level Further Mathematics A

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.
• Give non-exact numerical answers correct to 3 significant figures unless a different degree of accuracy is specified in the question.
• The acceleration due to gravity is denoted by $g\,\text{m}\,\text{s}^{-2}$. Unless otherwise instructed, when a numerical value is needed, use $g = 9.8$.

INFORMATION
• The total mark for this paper is **75**.
• The marks for each question are shown in brackets **[ ]**.
• **You are reminded of the need for clear presentation in your answers.**
• The Printed Answer Booklet consists of **12** pages. The Question Paper consists of **8** pages.

Answer **all** the questions.

1 The performance of a piece of music is being recorded. The piece consists of three sections, *A*, *B* and *C*. The times, in seconds, taken to perform the three sections are normally distributed random variables with the following means and standard deviations.

Section	Mean	Standard deviation
A	264	13
B	173	9
C	264	13

(a) Assume first that the times for the three sections are independent. Find the probability that the total length of the performance is greater than 720.0 seconds. **[3]**

(b) In fact sections *A* and *C* are musically identical, and the recording is made by using a single performance of section *A* twice, together with a performance of section *B*. In this case find the probability that the total length of the performance is greater than 720.0 seconds. **[2]**

2 In a fairground game a competitor scores 0, 1, 2 or 3 with probabilities given in the following table, where *a* and *b* are constants.

Score	0	1	2	3
Probability	*a*	*b*	*b*	*b*

The competitor's expected score is 0.9.

(a) Show that $b = 0.15$. **[2]**

(b) Find the variance of the score. **[3]**

(c) The competitor has to pay £2.50 to take part, and wins a prize of £2*X*, where *X* is the score achieved. Find the expectation of the competitor's loss. **[2]**

3 (a) Alex places 20 black counters and 8 white counters into a bag. She removes 8 counters at random without replacement. Find the probability that the bag now contains exactly 5 white counters. **[3]**

(b) Bill arranges 8 blue counters and 4 green counters in a random order in a straight line. Find the probability that exactly three of the green counters are next to one another. **[4]**

4 Leyla investigates the number of shoppers who visit a shop between 10.30 am and 11 am on Saturday mornings. She makes the following assumptions.

- Shoppers visit the shop independently of one another.
- The average rate at which shoppers visit the shop between these times is constant.

(a) State an appropriate distribution with which Leyla could model the number of shoppers who visit the shop between these times. **[1]**

Leyla uses this distribution, with mean 14, as her model.

(b) Calculate the probability that, between 10.35 am and 10.50 am on a randomly chosen Saturday, at least 10 shoppers visit the shop. **[3]**

Leyla chooses 25 Saturdays at random.

(c) Find the expected number of Saturdays, out of 25, on which there are no visitors to the shop between 10.35 am and 10.50 am. **[3]**

(d) In fact on 5 of these Saturdays there were no visitors to the shop between 10.35 am and 10.50 am. Use this fact to comment briefly on the validity of the model that Leyla has used. **[1]**

5 The birth rate, x per thousand members of the population, and the life expectancy at birth, y years, in 14 randomly selected African countries are given in the table.

Country	x	y	Country	x	y
Benin	4.8	59.2	Mozambique	5.4	54.63
Cameroon	4.7	54.87	Nigeria	5.7	52.29
Congo	4.9	61.42	Senegal	5.1	65.81
Gambia	5.7	59.83	Somalia	6.5	54.88
Liberia	4.7	60.25	Sudan	4.4	63.08
Malawi	5.1	60.97	Uganda	5.8	57.25
Mauretania	4.6	62.77	Zambia	5.4	58.75

$n = 14, \sum x = 72.8, \sum y = 826, \sum x^2 = 392.96, \sum y^2 = 48\,924.54, \sum xy = 4279.16$

(a) Calculate Pearson's product-moment correlation coefficient r for the data. **[2]**

(b) State what would be the effect on the value of r if the birth rate were given per hundred and not per thousand. **[1]**

(c) Explain what the sign of r tells you about the relationship between life expectancy and birth rate for these countries. **[1]**

(d) Test at the 5% significance level whether there is correlation between birth rate and life expectancy at birth in African countries. **[5]**

(e) A researcher wants to estimate the life expectancy at birth in Zimbabwe, where the birth rate is 3.9 per thousand. Explain whether a reliable estimate could be obtained using the regression line of y on x for the given data. **[1]**

6 The reaction times, in milliseconds, of all adult males in a standard experiment have a symmetrical distribution with mean and median both equal to 700 and standard deviation 125.

The reaction times of a random sample of 6 international athletes are measured and the results are as follows:

702   631   540   714   575   480

It is required to test whether international athletes have a mean reaction time which is less than 700.

(a) Assume first that the reaction times of international athletes have the distribution $N(\mu, 125^2)$. Test at the 5% significance level whether $\mu < 700$. **[7]**

(b) Now assume only that the distribution of the data is symmetrical, but **not** necessarily normal.

  (i) State with a reason why a Wilcoxon test is preferable to a sign test. **[1]**

  (ii) Use an appropriate Wilcoxon test at the 5% significance level to test whether the median reaction time of international athletes is less than 700. **[6]**

(c) Explain why the significance tests in part **(a)** and part **(b)(ii)** could produce different results. **[1]**

7 Sasha tends to forget his passwords. He investigates whether the number of attempts he needs to log on to a system with a password can be modelled by a geometric distribution. On 60 occasions he records the number of attempts he needs to log on, and the results are shown in the table.

Number of attempts	1	2	3	4 or more
Frequency	20	19	13	3

(a) Test at the 1% significance level whether the results are consistent with the distribution Geo(0.4). **[9]**

(b) Suggest which two probabilities should be changed, and in what way, to produce an improved model. (Numerical values are not required.) You should give a reason for your suggestion. **[3]**

8 A continuous random variable X has probability density function given by the following function, where a is a constant.

$$f(x) = \begin{cases} \dfrac{2x}{a^2} & 0 \leqslant x \leqslant a, \\ 0 & \text{otherwise.} \end{cases}$$

The expected value of X is 4.

(a) Show that $a = 6$. **[3]**

Five independent observations of X are obtained, and the largest of them is denoted by M.

(b) Find the cumulative distribution function of M. **[8]**

END OF QUESTION PAPER

7

BLANK PAGE