Oxford Cambridge and RSA

A Level Mathematics A

H240/02 Pure Mathematics and Statistics

Question Set 4

1. The stem-and-leaf diagram shows the heights, in centimetres, of 17 plants, measured correct to the nearest centimetre.

5	5 7 9 9
6	3 4 5 5 5 9 9
7	4 5 7 9 9
8	
9	9

Key: 5 | 6 means 56

(a) Find the median and inter-quartile range of these heights. **[3]**

median = 65
IQR = 76 − 61 = 15

(b) Calculate the mean and standard deviation of these heights. **[2]**

mean = 69, sd = 10.5

(c) State one advantage of using the median rather than the mean as a measure of average for these heights. **[1]**

Less affected by outliers

2 **(a)** The masses, in grams, of plums of a certain kind have the distribution N(55, 18).

(i) Find the probability that a plum chosen at random has a mass between 50.0 and 60.0 grams. **[1]** 0.701

(ii) The heaviest 5% of plums are classified as extra large.

Find the minimum mass of extra large plums. **[1]** 62.0

(iii) The plums are packed in bags, each containing 10 randomly selected plums.

$\bar{X} \sim N\left(55, \frac{18}{10}\right)$.

Find the probability that a bag chosen at random has a total mass of less than 530 g. **[4]**

$P = 0.0680$

(b) The masses, in grams, of apples of a certain kind have the distribution $N(67, \sigma^2)$. It is given that half of the apples have masses between 62 g and 72 g.

Determine σ. **[5]**

$\frac{72-67}{\sigma} = 0.674$. $\frac{62-67}{\sigma} = -0.674$. $\underline{\sigma = 7.42}$

3 The level, in grams per millilitre, of a pollutant at different locations in a certain river is denoted by the random variable X, where X has the distribution $N(\mu, 0.000\,040\,9)$.

In the past the value of μ has been 0.0340.

This year the mean level of the pollutant at 50 randomly chosen locations was found to be 0.0325 grams per millilitre.

Test, at the 5% significance level, whether the mean level of pollutant has changed. **[7]**

$H_0: \mu = 0.034$. $H_1: \mu \neq 0.034$ where μ = pop mean pollutant

$\bar{X} \sim N\left(0.034, \frac{0.0000409}{50}\right)$.

$P(\bar{X} < 0.0325) = 0.0486 > 0.025$ so

insufficient evidence to reject H_0

4

1. The stem-and-leaf diagram shows the heights, in centimetres, of 17 plants, measured correct to the nearest centimetre.

5	5 7 9 9
6	3 4 5 5 5 9 9
7	4 5 7 9 9
8	
9	9

Key: 5 | 6 means 56

(a) Find the median and inter-quartile range of these heights. **[3]**

median = 65
IQR = 76 − 61 = 15

(b) Calculate the mean and standard deviation of these heights. **[2]**

mean = 69   sd = 0.5

(c) State one advantage of using the median rather than the mean as a measure of average for these heights. **[1]**

Less affected by outliers

2 **(a)** The masses, in grams, of plums of a certain kind have the distribution N(55, 18).

(i) Find the probability that a plum chosen at random has a mass between 50.0 and 60.0 grams. **[1]**

0.701

(ii) The heaviest 5% of plums are classified as extra large.

Find the minimum mass of extra large plums. **[1]**

62.0

(iii) The plums are packed in bags, each containing 10 randomly selected plums.

Find the probability that a bag chosen at random has a total mass of less than 530 g. **[4]**

$\bar{X} \sim N(55, \frac{18}{10})$.
P = 0.0680

(b) The masses, in grams, of apples of a certain kind have the distribution N(67, σ^2). It is given that half of the apples have masses between 62 g and 72 g.

Determine σ. **[5]**

$\frac{72-67}{} = 0.674$, $\frac{62-67}{} = -0.674$ $\sigma = 7.4$

3 The level, in grams per millilitre, of a pollutant at different locations in a certain river is denoted by the random variable X, where X has the distribution $N(\mu, 0.000\,040\,9)$.

In the past the value of μ has been 0.0340.

This year the mean level of the pollutant at 50 randomly chosen locations was found to be 0.0325 grams per millilitre.

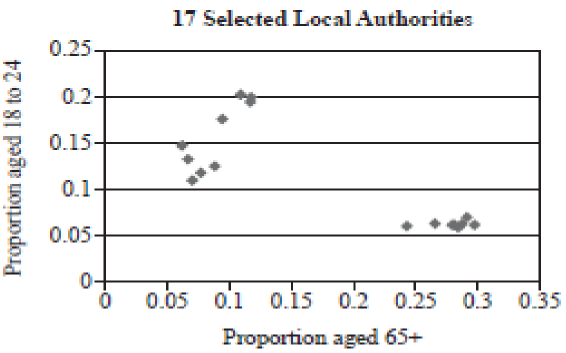
Test, at the 5% significance level, whether the mean level of pollutant has changed. [7]

$H_0: \mu = 0.034$. $H_1: \mu \neq 0.034$ where μ = pop mean pollutant

$\bar{X} \sim N\left(0.034, \frac{0.0000409}{50}\right)$

$P(\bar{X} < 0.0325) = 0.0486 > 0.025$ so insufficient evidence to reject H_0

4 A trainer was asked to give a lecture on population profiles in different Local Authorities (LAs) in the UK. Using data from the 2011 census, he created the following scatter diagram for 17 selected LAs.



He selected the 17 LAs using the following method. The proportions of people aged 18 to 24 and aged 65+ in any Local Authority are denoted by P_{young} and P_{senior} respectively. The trainer used a spreadsheet to calculate the value of $k = \frac{P_{\text{young}}}{P_{\text{senior}}}$ for each of the 348 LAs in the UK. He then used specific ranges of values of k to select the 17 LAs.

(a) Estimate the ranges of values of k that he used to select these 17 LAs. [3]

$\frac{0.11}{0.3} \leq k \leq \frac{0.21}{0.24}$

$0.367 \leq k \leq 0.724$

(b) Using the 17 LAs the trainer carried out a hypothesis test with the following hypotheses.

H_0: There is no linear correlation in the population between P_{young} and P_{senior}
H_1: There is negative linear correlation in the population between P_{young} and P_{senior}

He found that the value of Pearson's product-moment correlation coefficient for the 17 LAs is -0.797, correct to 3 significant figures.

(i) Use the table on page 9 to show that this value is significant at the 1% level. [2]

Critical value at 1% level n=17 $\longrightarrow -0.557 > -0.797$

The trainer concluded that there is evidence of negative linear correlation between P_{young} and P_{senior} in the population.

(ii) Use the diagram to comment on the reliability of this conclusion. [2]

Not reliable as data points are clustered together in 2 places with a gap in the middle

(c) Describe one outstanding feature of the population in the areas represented by the points in the bottom right hand corner of the diagram. Large 65+ Population [1]

(d) The trainer's audience included representatives from several universities.

Suggest a reason why the diagram might be of particular interest to these people. [1]

Could be used for further research by them

Critical values of Pearson's product-moment correlation coefficient

1-tail test	5%	2.5%	1%	0.5%
2-tail test	10%	5%	2%	1%
n				
1	-	-	-	-
2	-	-	-	-
3	0.9877	0.9969	0.9995	0.9999
4	0.9000	0.9500	0.9800	0.9900
5	0.8054	0.8783	0.9343	0.9587
6	0.7293	0.8114	0.8822	0.9172
7	0.6694	0.7545	0.8329	0.8745
8	0.6215	0.7067	0.7887	0.8343
9	0.5822	0.6664	0.7498	0.7977
10	0.5494	0.6319	0.7155	0.7646
11	0.5214	0.6021	0.6851	0.7348
12	0.4973	0.5760	0.6581	0.7079
13	0.4762	0.5529	0.6339	0.6835
14	0.4575	0.5324	0.6120	0.6614
15	0.4409	0.5140	0.5923	0.6411
16	0.4259	0.4973	0.5742	0.6226
17	0.4124	0.4821	0.5577	0.6055
18	0.4000	0.4683	0.5425	0.5897
19	0.3887	0.4555	0.5285	0.5751
20	0.3783	0.4438	0.5155	0.5614
21	0.3687	0.4329	0.5034	0.5487
22	0.3598	0.4227	0.4921	0.5368
23	0.3515	0.4132	0.4815	0.5256
24	0.3438	0.4044	0.4716	0.5151
25	0.3365	0.3961	0.4622	0.5052
26	0.3297	0.3882	0.4534	0.4958
27	0.3233	0.3809	0.4451	0.4869
28	0.3172	0.3739	0.4372	0.4785
29	0.3115	0.3673	0.4297	0.4705
30	0.3061	0.3610	0.4226	0.4629

5 A random variable X has probability distribution defined as follows.

$$P(X = x) = \begin{cases} kx & x = 1, 2, 3, 4, 5, \\ 0 & \text{otherwise,} \end{cases}$$

where k is a constant.

(a) Show that $P(X = 3) = 0.2$. **[3]**

$k(1+2+3+4+5)=1 \quad k=\frac{1}{15}$

$\frac{1}{15}\times 3 = 0.2$

(b) Show in a table the values of X and their probabilities. **[2]**

x	1	2	3	4	5
$P(X=x$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{4}{15}$	$\frac{5}{15}$

(c) Two independent values of X are chosen, and their total T is found.

(i) Find $P(T = 7)$. **[3]**

$3,4 \quad 4,3 \quad 2,5 \quad 5,2 \therefore \frac{3}{15}\times\frac{4}{15}\times 2+\frac{5}{15}\times\frac{2}{15}\times 2=\frac{44}{225}$

(ii) Given that $T = 7$, determine the probability that one of the values of X is 2. **[4]**

$$\frac{\frac{5}{15}\times\frac{2}{15}\times 2}{\frac{44}{225}} = \frac{5}{11}$$

6 It is known that 26% of adults in the UK use a certain app. A researcher selects a random sample of 5000 adults in the UK. The random variable X is defined as the number of adults in the sample who use the app.

Given that $P(X < n) < 0.025$, calculate the largest possible value of n. **[5]**

$X \sim B(5000, 0.26)$

using inverse binomial $n \approx 1239$

$P(X \leq 1239) = 0.0251$

$P(X \leq 1238 = 0.0233 \quad \therefore n = 1238$

Total Marks for Question Set 4: 49 Marks