AQA

Please write clearly in block capitals.

Centre number | | | | | | Candidate number | | | | |

Surname StemUp Model Solution

Forename(s) _____

Candidate signature _____

I declare this is my own work.

A-level
FURTHER MATHEMATICS

Paper 3 Statistics

Thursday 11 June 2020 Afternoon Time allowed: 2 hours

Materials
- You must have the AQA formulae and statistical tables booklet for A-level Mathematics and A-level Further Mathematics.
- You should have a scientific calculator that meets the requirements of the specification. (You may use a graphical calculator.)
- You must ensure you have the other optional Question Paper/Answer Book for which you are entered (**either** Discrete **or** Mechanics). You will have 2 hours to complete **both** papers.

Instructions
- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- You must answer each question in the space provided for that question. If you require extra space for your answer(s), use the lined pages at the end of this book. Write the question number against your answer(s).
- Do **not** write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work you do not want to be marked.

For Examiner's Use	
Question	Mark
1	
2	
3	
4	
5	
6	
7	
8	
9	
TOTAL	

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 50.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.
- You do not necessarily need to use all the space provided.

Do not write outside the box

Answer **all** questions in the spaces provided.

1 The continuous random variable X has probability density function

$$f(x) = \begin{cases} \frac{1}{5} & 1 \leq x \leq 6 \\ 0 & \text{otherwise} \end{cases}$$

Find $P(X \geq 3)$

Circle your answer.

[1 mark]

$\frac{1}{5}$ $\frac{2}{5}$ $\frac{3}{5}$ $\frac{4}{5}$

2 Jamie is conducting a hypothesis test on a random variable which has a normal distribution with standard deviation 1

The hypotheses are

$H_0 : \mu = 5$

$H_1 : \mu > 5$

He takes a random sample of size 4

The mean of his sample is 6

He uses a 5% level of significance.

Before Jamie conducted the test, what was the probability that he would make a Type I error?

Circle your answer.

[1 mark]

0.0228 0.0456 0.0500 0.1587

Do not write outside the box

3 The mass of male giraffes is assumed to have a normal distribution.

Duncan takes a random sample of 600 male giraffes.

The mean mass of the sample is 1196 kilograms.

The standard deviation of the sample is 98 kilograms.

3 (a) Construct a 94% confidence interval for the mean mass of male giraffes, giving your values to one decimal place.

[3 marks]

$94\% \Rightarrow P(Z < z) = 0.97 \Rightarrow z = 1.8808$

$CI = \bar{x} \pm z \times \frac{s}{\sqrt{n}}$

$= 1196 \pm 1.8808 \times \frac{98}{\sqrt{600}}$

$= 1196 \pm 7.5247...$

94% confidence interval = (1188.5, 1203.5)

3 (b) Explain whether or not your answer to part **(a)** would change if a sample of size 5 was taken with the same mean and standard deviation.

[1 mark]

Yes, it would change. The sample size n is part of the calculation $(s/\sqrt{n})$, so with $n = 5$ a t distribution (t_4) would be used instead of the normal (z) distribution, and the interval would be much wider.

Turn over for the next question

Do not write outside the box

4 The discrete random variable X follows a discrete uniform distribution and takes values 1, 2, 3, …, n.

The discrete random variable Y is defined by $Y = 2X$

4 (a) Using the standard results for $\sum n$, $\sum n^2$ and $\text{Var}(aX + b)$, prove that

$$\text{Var}(Y) = \frac{n^2 - 1}{3}$$

[7 marks]

$E(X) = \Sigma\, x\, P(X = x) = \frac{1}{n}(1 + 2 + \ldots + n)$

$= \frac{1}{n} \times \frac{1}{2} n(n + 1) = \frac{n + 1}{2}$

$E(X^2) = \frac{1}{n}(1^2 + 2^2 + \ldots + n^2)$

$= \frac{1}{n} \times \frac{1}{6} n(n + 1)(2n + 1) = \frac{(n + 1)(2n + 1)}{6}$

$\text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{(n + 1)(2n + 1)}{6} - \frac{(n + 1)^2}{4}$

$= \frac{2n^2 + 3n + 1}{6} - \frac{n^2 + 2n + 1}{4}$

$= \frac{4n^2 + 6n + 2 - 3n^2 - 6n - 3}{12}$

$= \frac{n^2 - 1}{12}$

$\text{Var}(Y) = \text{Var}(2X) = 2^2\, \text{Var}(X)$ $\quad (\text{Var}(aX + b) = a^2\, \text{Var}(X))$

$= 4 \times \frac{n^2 - 1}{12}$

$= \frac{n^2 - 1}{3}$ as required

Do not write outside the box

4 (b) A spinning toy can land on one of four values: 2, 4, 6 or 8

Using a discrete uniform distribution, find the probability that the next value the toy lands on is greater than 2

[1 mark]

$P(\text{next value} > 2) = P(4) + P(6) + P(8) = 3 \times \frac{1}{4} = 0.75$

4 (c) State an assumption that is required for the discrete uniform distribution used in part **(b)** to be valid.

[1 mark]

The spinning toy is unbiased, i.e. it is equally likely to land on each of the four values (probability 0.25 each).

Turn over for the next question

Do not write outside the box

5 Emily claims that the average number of runners per minute passing a shop during a long distance run is 8

Emily conducts a hypothesis test to investigate her claim.

5 (a) State the hypotheses for Emily's test.

[1 mark]

H_0: $\lambda = 8$

H_1: $\lambda \neq 8$ (λ = mean number of runners passing per minute)

5 (b) Emily counts the number of runners, X, passing the shop in a randomly chosen minute.

The critical region for Emily's test is $X \leq 2$ or $X \geq 14$

During a randomly chosen minute, Emily counts 3 runners passing the shop.

Determine the outcome of Emily's hypothesis test.

[3 marks]

$X = 3$ is not in the critical region ($X \leq 2$ or $X \geq 14$)

Do not reject H_0

There is no significant evidence to suggest that the average number of runners per minute passing the shop is not 8, so there is not enough evidence to reject Emily's claim.

Do not write outside the box

5 (c) The actual average number of runners per minute passing the shop is 7

Find the power of Emily's hypothesis test, giving your answer to three significant figures.

[3 marks]

Power = P(reject H_0 when $\lambda = 7$), $X \sim Po(7)$

$P(X \leq 2) = 0.029636...$

$P(X \geq 14) = 1 - P(X \leq 13) = 1 - 0.98719 = 0.012811...$

Power = $P(X \leq 2) + P(X \geq 14) = 0.029636 + 0.012811$

$= 0.042447...$

$= 0.0424$ (3 s.f.)

Turn over for the next question

Do not write outside the box

6 The distance, X metres, between successive breaks in a water pipe is modelled by an exponential distribution. The mean of X is 25

The distance between two successive breaks is measured. A water pipe is given a 'Red' rating if the distance is less than d metres.

The government has introduced a new law changing d to 2

Before the government introduced the new law, the probability that a water pipe is given a 'Red' rating was 0.05

6 (a) Explain whether or not the probability that a water pipe is given a 'Red' rating has increased as a result of the new law.

[4 marks]

$E(X) = \frac{1}{\lambda} = 25 \Rightarrow \lambda = \frac{1}{25} = 0.04$

Before the law: $P(X < d) = 1 - e^{-0.04d} = 0.05$

$e^{-0.04d} = 0.95 \Rightarrow d = -25 \ln 0.95 = 1.282... \approx 1.3$ metres

New law, $d = 2$: $P(X < 2) = 1 - e^{-0.04 \times 2}$

$= 1 - e^{-0.08} = 0.076883... \approx 0.077$

$0.077 > 0.05$ (equivalently 2 metres > 1.3 metres)

So yes, the probability that a water pipe is given a 'Red' rating has increased as a result of the new law.

Do not write outside the box

6 (b) Find the probability density function of the random variable X.

[2 marks]

$$f(x) = \begin{cases} 0.04e^{-0.04x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

6 (c) After investigation, the distances between successive breaks in water pipes are found to have a standard deviation of 5 metres.

Explain whether or not the use of an exponential model in parts **(a)** and **(b)** is appropriate.

[2 marks]

For the exponential model, s.d. $= \frac{1}{\lambda} = \frac{1}{0.04} = 25$ metres

The actual standard deviation of 5 metres is very different from the 25 metres of the model, so the exponential model is not appropriate.

Turn over for the next question

Do not write outside the box

7 The rainfall per day in February in a particular town has been recorded as having a mean of 1.8 inches.

Sienna claims that rainfall in February has increased in the town. She records the rainfall in a random sample of 12 days.

Her sample mean is 2 inches and her sample standard deviation is 0.4 inches.

It is assumed that rainfall per day has a normal distribution.

7 (a) Investigate Sienna's claim using the 5% level of significance.

[6 marks]

H_0: $\mu = 1.8$ H_1: $\mu > 1.8$

(μ = population mean rainfall per day in February, in inches)

$$t = \frac{\bar{x} - \mu}{s \div \sqrt{n}} = \frac{2 - 1.8}{0.4 \div \sqrt{12}}$$

$= 1.7320...$ $= 1.73$

Under H_0, $t \sim t_{11}$ $(\nu = 12 - 1 = 11)$

5% one-tailed critical value $t_{11} = 1.796$

$1.73 < 1.796$, so the test statistic is not in the critical region

Do not reject H_0

There is no significant evidence at the 5% level to suggest that the mean rainfall per day in February in the town has increased, so Sienna's claim is not supported.

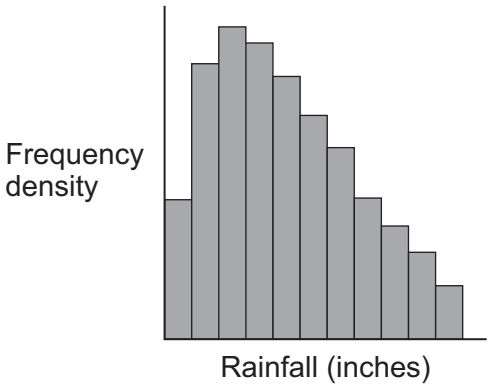
Do not write outside the box

7 (b) For the test carried out in part **(a)**, state in context the meaning of a Type II error.

[1 mark]

Concluding that the mean rainfall per day in February has not increased (not rejecting H_0) when in fact it has increased.

7 (c) The distribution of rainfall per day in February in the town over 10 years is shown in the histogram.



Explain whether or not the assumption that rainfall per day in February has a normal distribution is appropriate.

[1 mark]

Not appropriate: the histogram is not symmetrical (it is positively skewed), whereas a normal distribution is symmetrical.

Turn over for the next question

Do not write outside the box

8 Ray is conducting a hypothesis test with the hypotheses

H_0 : There is no association between time of day and number of snacks eaten

H_1 : There is an association between time of day and number of snacks eaten

He calculates **expected** frequencies correct to two decimal places, which are given in the following table.

		Number of snacks eaten		
		0	1	2 or more
Time of Day	Day	23.68	21.05	5.26
	Night	21.32	18.95	4.74

Ray calculates his test statistic using $\sum \frac{(O-E)^2}{E}$

8 (a) State, with a reason, the error Ray has made and describe any changes Ray will need to make to his test.

[3 marks]

Error: one expected frequency is less than 5 (4.74 < 5), so Ray cannot carry out the test on this table as it is.

Changes: merge the '1' and '2 or more' columns into '1 or more', which gives a 2 × 2 table with $\nu = (2-1)(2-1) = 1$ degree of freedom.

With 1 degree of freedom Yates' correction must be used, so the test statistic will be

$\Sigma \frac{(|O-E|-0.5)^2}{E}$ *rather than* $\Sigma \frac{(O-E)^2}{E}$

Do not write outside the box

8 (b) Having made the necessary corrections as described in part **(a)**, the correct value of the test statistic is 8.74

Complete Ray's hypothesis test using a 1% level of significance.

[3 marks]

$\nu = 1$: χ^2 critical value at the 1% level = 6.635

8.74 > 6.635, so the test statistic is in the critical region

Reject H_0

There is significant evidence at the 1% level to suggest that there is an association between time of day and number of snacks eaten.

Turn over for the next question

Do not write outside the box

9 The continuous random variable X has the cumulative distribution function shown below.

$$\mathrm{F}(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{62}(4x^3 + 6x^2 + 3x) & 0 \leq x \leq 2 \\ 1 & x > 2 \end{cases}$$

The discrete random variable Y has the probability distribution shown below.

y	2	7	13	19
$\mathrm{P}(Y = y)$	0.5	0.1	0.1	0.3

The random variables X and Y are independent.

Find the exact value of $\mathrm{E}(X^3 + Y)$.

[6 marks]

$f(x) = F'(x) = \frac{1}{62}(12x^2 + 12x + 3) = \frac{3}{62}(4x^2 + 4x + 1), \quad 0 \leq x \leq 2$

$E(X^3) = \int_0^2 x^3 f(x)\,dx = \frac{3}{62}\int_0^2 (4x^5 + 4x^4 + x^3)\,dx$

$= \frac{3}{62}\left[\frac{2}{3}x^6 + \frac{4}{5}x^5 + \frac{1}{4}x^4\right]_0^2$

$= \frac{3}{62}\left(\frac{128}{3} + \frac{128}{5} + 4\right) = \frac{3}{62} \times \frac{1084}{15}$

$= \frac{1084}{310} = \frac{542}{155}$

$E(Y) = 2 \times 0.5 + 7 \times 0.1 + 13 \times 0.1 + 19 \times 0.3$

$= 1 + 0.7 + 1.3 + 5.7 = 8.7 = \frac{87}{10}$

$E(X^3 + Y) = E(X^3) + E(Y) = \frac{542}{155} + \frac{87}{10}$

$= \frac{1084}{310} + \frac{2697}{310} = \frac{3781}{310}$

Do not write outside the box

END OF QUESTIONS

Do not write outside the box

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED

Do not write outside the box

Question number	**Additional page, if required.** **Write the question numbers in the left-hand margin.**

Do not write outside the box

Question number	**Additional page, if required.** **Write the question numbers in the left-hand margin.**

Do not write outside the box

Question number	**Additional page, if required.** **Write the question numbers in the left-hand margin.**

Do not write outside the box

Question number	**Additional page, if required.** **Write the question numbers in the left-hand margin.**

Copyright information

For confidentiality purposes, all acknowledgements of third-party copyright material are published in a separate booklet. This booklet is published after each live examination series and is available for free download from www.aqa.org.uk.

Permission to reproduce all copyright material has been applied for. In some cases, efforts to contact copyright-holders may have been unsuccessful and AQA will be happy to rectify any omissions of acknowledgements. If you have any queries please contact the Copyright Team.

Copyright © 2020 AQA and its licensors. All rights reserved.