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

Friday 7 June 2024 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 graphical or scientific calculator that meets the requirements of the specification.
- 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 or on blank pages.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.

For Examiner's Use	
Question	Mark
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
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 random variable X has a Poisson distribution with mean 16

Find the standard deviation of X

Circle your answer.

[1 mark]

(4) 8 16 256

$\sigma = \sqrt{16} = 4$

2 The random variable T has an exponential distribution with mean 2

Find $P(T \leq 1.4)$

Circle your answer.

[1 mark]

$e^{-2.8}$ $e^{-0.7}$ $(1 - e^{-0.7})$ $1 - e^{-2.8}$

$T \sim Exp(\lambda)$ with $\lambda = 1/2 = 0.5$
$P(T \leq 1.4) = 1 - e^{-0.5 \times 1.4} = 1 - e^{-0.7}$

Do not write outside the box

3 The continuous random variable Y has **cumulative** distribution function

$$F(y) = \begin{cases} 0 & y < 2 \\ -\frac{1}{9}y^2 + \frac{10}{9}y - \frac{16}{9} & 2 \le y < 5 \\ 1 & y \ge 5 \end{cases}$$

Find the median of Y

Circle your answer.

[1 mark]

2 $\frac{10 - 3\sqrt{2}}{2}$ $\frac{7}{2}$ $\frac{10 + 3\sqrt{2}}{2}$

Turn over for the next question

$F(m) = 0.5$: $-m^2 + 10m - 16 = 4.5$

$m^2 - 10m + 20.5 = 0 \Rightarrow m = (10 \pm \sqrt{18})/2 = (10 \pm 3\sqrt{2})/2$

m must satisfy $2 \le m < 5$, so $m = (10 - 3\sqrt{2})/2 = 2.88$

Do not write outside the box

4 Research has shown that the mean number of volcanic eruptions on Earth each day is 20

Sandra records 162 volcanic eruptions during a period of **one week**.

Sandra claims that there has been an increase in the mean number of volcanic eruptions per week.

Test Sandra's claim at the 5% level of significance.

[6 marks]

Let X = number of volcanic eruptions in one week

Mean per week = $7 \times 20 = 140$

H_0: $\lambda = 140$ H_1: $\lambda > 140$ (λ = mean number per week)

Under H_0, $X \sim Po(140)$

$P(X \geq 162) = 1 - P(X \leq 161)$

$= 1 - 0.96304$

$= 0.03696... = 0.037$ (3 d.p.)

$0.037 < 0.05$, so the result is significant at the 5% level

Reject H_0

There is sufficient evidence at the 5% level to suggest that the mean number of volcanic eruptions per week has increased, supporting Sandra's claim.

Do not write outside the box

5 The continuous random variable X has probability density function

$$f(x) = \begin{cases} \frac{1}{6}e^{\frac{x}{3}} & 0 \leq x \leq \ln 27 \\ 0 & \text{otherwise} \end{cases}$$

Show that the mean of X is $\frac{3}{2}(\ln 27 - 2)$

[5 marks]

$E(X) = \int_0^{\ln 27} \frac{x}{6} e^{x/3}\, dx$

By parts: $u = x$, $\frac{dv}{dx} = \frac{1}{6}e^{x/3} \Rightarrow \frac{du}{dx} = 1$, $v = \frac{1}{2}e^{x/3}$

$E(X) = \left[\frac{x}{2}e^{x/3}\right]_0^{\ln 27} - \int_0^{\ln 27} \frac{1}{2}e^{x/3}\, dx$

$= \left[\frac{x}{2}e^{x/3} - \frac{3}{2}e^{x/3}\right]_0^{\ln 27}$

When $x = \ln 27$: $e^{(\ln 27)/3} = 27^{1/3} = 3$; when $x = 0$: $e^0 = 1$

$= \left(\frac{\ln 27}{2} \times 3 - \frac{3}{2} \times 3\right) - \left(0 - \frac{3}{2} \times 1\right)$

$= \frac{3}{2}\ln 27 - \frac{9}{2} + \frac{3}{2}$

$= \frac{3}{2}\ln 27 - 3$

$= \frac{3}{2}(\ln 27 - 2)$ as required

Turn over ▶

Do not write outside the box

6 Over time it has been accepted that the mean retirement age for professional baseball players is 29.5 years old.

Imran claims that the mean retirement age is no longer 29.5 years old.

He takes a random sample of 5 recently retired professional baseball players and records their retirement ages, x. The results are

$$\sum x = 152.1 \qquad \text{and} \qquad \sum (x - \bar{x})^2 = 7.81$$

6 (a) State an assumption that you should make about the distribution of the retirement ages to investigate Imran's claim.

[1 mark]

Assume that the retirement ages are normally distributed.

6 (b) Investigate Imran's claim, using the 10% level of significance.

[8 marks]

H_0: $\mu = 29.5$ H_1: $\mu \neq 29.5$ (μ = mean retirement age, years)

$\bar{x} = \frac{152.1}{5} = 30.42$

$s^2 = \frac{7.81}{5-1} = 1.9525$ $(s = 1.3973...)$

$t = \frac{\bar{x} - \mu}{\sqrt{(s^2 \div n)}} = \frac{30.42 - 29.5}{\sqrt{(1.9525 \div 5)}} = \frac{0.92}{0.62490...}$

$= 1.4722... = 1.47$ (3 s.f.)

Under H_0, $t \sim t_4$ $(\nu = 5 - 1 = 4)$

10% two-tailed test: critical values $t = \pm 2.132$

$-2.132 < 1.47 < 2.132$, so 1.47 is not in the critical region

Do not reject H_0

There is insufficient evidence at the 10% level to suggest that the mean retirement age of professional baseball players is no longer 29.5 years old, so Imran's claim is not supported.

Do not write outside the box

Turn over for the next question

Do not write outside the box

7 The random variable X has a discrete uniform distribution and takes values 1, 2, 3, …, n

The random variable Y has probability distribution function

$$P(Y = y) = \begin{cases} \frac{1}{n} & y = 13, 16, 19 \ldots, 3n + 10 \\ 0 & \text{otherwise} \end{cases}$$

7 (a) State the variance of X

[1 mark]

$Var(X) = \frac{n^2 - 1}{12}$

7 (b) Hence find the variance of Y

Fully justify your answer.

[3 marks]

When $X = 1, 2, 3, \ldots, n$, $3X + 10 = 13, 16, 19, \ldots, 3n + 10$,

each value having probability $1/n$, so Y has the same distribution

as $3X + 10$, i.e. $Y = 3X + 10$

$Var(Y) = Var(3X + 10) = 3^2\, Var(X) = 9\, Var(X)$

$= 9 \times \frac{n^2 - 1}{12}$

$Var(Y) = \frac{3(n^2 - 1)}{4}$

Do not write outside the box

8 The discrete random variable X has probability distribution

x	7	13	17	21
$\mathbf{P}(X = x)$	0.4	0.35	0.1	0.15

Show that $P(X \leq 17 \mid X > E(X)) = 0.75$

[5 marks]

$E(X) = \Sigma\, x\, P(X = x)$

$= 7 \times 0.4 + 13 \times 0.35 + 17 \times 0.1 + 21 \times 0.15$

$= 2.8 + 4.55 + 1.7 + 3.15 \quad = 12.2$

$P(X > E(X)) = P(X > 12.2) = P(X \geq 13)$

$= 0.35 + 0.1 + 0.15 = 0.6$

$P(X \leq 17 \cap X > 12.2) = P(13 \leq X \leq 17)$

$= 0.35 + 0.1 = 0.45$

$P(X \leq 17 \mid X > E(X)) = \dfrac{P(X \leq 17 \cap X > E(X))}{P(X > E(X))}$

$= \dfrac{0.45}{0.6}$

$= 0.75$ as required

Do not write outside the box

9 A company owns three shops, A, B and C, which are based in different towns.

Each shop gives a questionnaire to 250 of their customers, and every customer completes the questionnaire.

One of the questions asks whether the customer rates the shop as good, satisfactory or poor.

For shop A, 26% of customers rate the shop as good and 38% of customers rate the shop as poor.

For shop B, 32% of customers rate the shop as good and 40% of customers rate the shop as satisfactory.

Altogether, there are 210 good ratings and 261 satisfactory ratings.

9 (a) Complete the following table with the observed frequencies.

[3 marks]

		Rating		
		Good	**Satisfactory**	**Poor**
Shop	**A**	65	90	95
	B	80	100	70
	C	65	71	114

9 (b) Carry out a test for association between shop and rating, using the 1% level of significance.

[8 marks]

H_0: there is no association between shop and rating

H_1: there is an association between shop and rating

Totals: each shop 250; Good 210, Satisfactory 261, Poor 279; total 750

Expected frequency E = (row total × column total) ÷ 750

e.g. $E(A, Good) = 250 \times 210 \div 750 = 70$

Do not write outside the box

Expected	Good	Satisfactory	Poor
A	70	87	93
B	70	87	93
C	70	87	93

All expected frequencies ≥ 5, so no cells need combining

$$\chi^2 = \Sigma \frac{(O - E)^2}{E}$$

$$= \frac{(65 - 70)^2 + (80 - 70)^2 + (65 - 70)^2}{70} + \frac{(90 - 87)^2 + (100 - 87)^2 + (71 - 87)^2}{87}$$

$$+ \frac{(95 - 93)^2 + (70 - 93)^2 + (114 - 93)^2}{93}$$

$$= 0.3571 + 1.4286 + 0.3571 + 0.1034 + 1.9425 + 2.9425$$

$$+ 0.0430 + 5.6882 + 4.7419$$

$$= 17.604... = 17.60 \text{ (2 d.p.)}$$

$\nu = (3 - 1)(3 - 1) = 4$

Critical value at 1% level: $\chi^2_4 = 13.277$

$17.60 > 13.277$, so the test statistic is in the critical region

Reject H_0

There is sufficient evidence at the 1% level to suggest that there is an association between shop and rating.

Turn over for the next question

Do not write outside the box

10 The length, X cm, of teeth for an extinct breed of shark is known to have a normal distribution.

A random sample of 250 teeth has the following summary statistics.

$\bar{x}$	s
10.2	1.3

10 (a) Construct a 91% confidence interval for the population mean length of the shark teeth.

Give your limits to two decimal places.

[3 marks]

$n = 250$ is large, so $\bar{x}$ is approximately normal (σ estimated by s)

$P(Z < z) = 0.955 \Rightarrow z = 1.6954$

$CI = \bar{x} \pm z \frac{s}{\sqrt{n}} = 10.2 \pm 1.6954 \times \frac{1.3}{\sqrt{250}}$

$= 10.2 \pm 0.13939...$

91% confidence interval = (10.06, 10.34)

Do not write outside the box

10 (b) The random sample is used to carry out a hypothesis test using the 9% level of significance.

The hypotheses are

$H_0: \mu = 10.36$

$H_1: \mu \neq 10.36$

10 (b) (i) State, with a reason, whether the null hypothesis will be rejected.

[1 mark]

H_0 is rejected, because 10.36 lies outside the 91% confidence interval (10.06, 10.34) (a 9% two-tailed test matches a 91% CI).

10 (b) (ii) The hypothesis test is repeated with a second random sample of 250 teeth which also has $s = 1.3$ cm

Find the power of this test assuming that the population mean length of the shark teeth is 10.32 cm

Give your answer to two decimal places.

[3 marks]

Acceptance region: $10.36 \pm 1.6954 \times \frac{1.3}{\sqrt{250}} = 10.36 \pm 0.13939...$

Reject H_0 if $\bar{x} < 10.2206$ or $\bar{x} > 10.4994$

If $\mu = 10.32$: $\bar{X} \sim N(10.32, \frac{1.3^2}{250})$

$P(\bar{X} < 10.2206) = P(Z < \frac{10.2206 - 10.32}{1.3 \div \sqrt{250}}) = P(Z < -1.2089) = 0.11335$

$P(\bar{X} > 10.4994) = P(Z > \frac{10.4994 - 10.32}{1.3 \div \sqrt{250}}) = P(Z > 2.1819) = 0.01456$

Power = 0.11335 + 0.01456 = 0.12791...

Power = 0.13 (2 d.p.)

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.**

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 © 2024 AQA and its licensors. All rights reserved.