AQA

Please write clearly in block capitals.

Centre number		Candidate number	

Surname ____________________

Forename(s) ____________________

Candidate signature ____________________
I declare this is my own work.

AS
FURTHER MATHEMATICS
Paper 2 Statistics

Thursday 14 May 2020 Afternoon Time allowed: 1 hour 30 minutes

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 1 hour 30 minutes 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.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 40.

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

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

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

1 The discrete random variable X has the following probability distribution function.

$$P(X = x) = \begin{cases} 0.2 & x = 1 \\ 0.3 & x = 2 \\ 0.1 & x = 3, 4 \\ 0.25 & x = 5 \\ 0.05 & x = 6 \\ 0 & \text{otherwise} \end{cases}$$

Find the mode of X.

Circle your answer.

[1 mark]

0.1   0.25   2   3

The mode is the value of the random variable where it is most dense, that is where the pdf reaches its highest point

2 A χ^2 test is carried out in a school to test for association between the class a student belongs to and the number of times they are late to school in a week.

The contingency table below gives the expected values for the test.

Number of times late

Class		0	1	2	3	4
	A	8.12	14	15.12	14	4.76
	B	8.99	15.5	16.74	15.5	5.27
	C	11.89	20.5	22.14	20.5	6.97

Find a possible value for the degrees of freedom for the test.

Circle your answer.

[1 mark]

6   8   12   15

Because there is an expected value less than 5 two columns will be merged.

Turn over for the next question

Degrees of freedom $= (r-1)(c-1)$

$(3-1)(4-1)$

$(2)(3)$

$= 6$

Do not write outside the box

3 The random variable X represents the value on the upper face of an eight-sided dice after it has been rolled. The faces are numbered 1 to 8

The random variable X is modelled by a discrete uniform distribution with $n = 8$

3 (a) Find $\mathrm{E}(X)$

[1 mark]

$X \sim U[1, 8]$

$E(x) = \frac{a+b}{2}$

$= \frac{1+8}{2} = \frac{9}{2}$

$= 4.5$

$\therefore E(x) = 4.5$

3 (b) Find $\mathrm{Var}(X)$

[1 mark]

$Var(x) = \frac{(b-a)^2}{12}$

$= \frac{8^2-1}{12}$

$= 5.25$

$Var(x) = 5.25$

3 (c) Find $\mathrm{P}(X \geq 6)$

[1 mark]

$P(x \geq 6) = P(x=6) + P(x=7) + P(x=8)$

$= \frac{1}{8} + \frac{1}{8} + \frac{1}{8}$

$= \frac{3}{8}$

Do not write outside the box

3 (d) The dice was rolled 800 times and the results below were obtained.

x	1	2	3	4	5	6	7	8
Frequency	103	63	84	110	74	41	85	240

State, with a reason, how you would refine the model for the random variable X.

[2 marks]

Because the dice is biased, the random variable X would be modelled with a discrete random variable where the probabilities are estimated using relative frequencies.

Turn over for the next question

Do not write outside the box

4 Murni is investigating the annual salary of people from a particular town.

She takes a random sample of 200 people from the town and records their annual salary.

The mean annual salary is £28 500 and the standard deviation is £5100

Calculate a 97% confidence interval for the population mean of annual salaries for the people who live in the town, giving your values to the nearest pound.

[3 marks]

At 97% confidence interval;

$$\bar{x} \pm z\sqrt{\frac{s^2}{n}}$$

The z-value corresponding to 97% confidence interval is 2.17

$$28500 \pm 2.17\sqrt{\frac{5100^2}{200}}$$

$$28500 \pm 2.17\left(\frac{5100}{\sqrt{200}}\right)$$

$$28500 \pm 783$$

$$= (27717, 29283)$$

Do not write outside the box

5 The discrete random variable X has the following probability distribution.

x	2	4	6	9
$P(X = x)$	0.2	0.6	0.1	0.1

5 (a) Find $P(X \le 6)$

[1 mark]

$p(X \le 6) = p(X=2) + p(X=4) + p(X=6)$

$= 0.2 + 0.6 + 0.1$

$= 0.9$

$\therefore p(X \le 6) = 0.9$

5 (b) Let $Y = 3X + 2$

Show that $\text{Var}(Y) = 32.49$

[5 marks]

$E(X) = \sum x \cdot p(X=x)$

$= (2 \times 0.2) + (4 \times 0.6) + (6 \times 0.1) + (9 \times 0.1)$

$= 0.4 + 2.4 + 0.6 + 0.9$

$= 4.3$

$E(X^2) = \sum x^2 \cdot p(X=x)$

$= (2^2 \times 0.2) + (4^2 \times 0.6) + (6^2 \times 0.1) + (9^2 \times 0.1)$

$= 0.8 + 9.6 + 3.6 + 8.1$

$= 22.1$

$Var(X) = E(X^2) - (E(X))^2$

$22.1 - (4.3)^2$

$= 3.61$

$Y = 3X + 2$

$Var(Y) = Var(3X + 2)$

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

$= 3^2 (3.61)$

$= 9 \times 3.61$

$= 32.49$

$\therefore Var(Y) = 32.49$

Do not write outside the box

5 (c) The continuous random variable T is independent of Y.

Given that $\text{Var}(T) = 5$, find $\text{Var}(T + Y)$

[1 mark]

$\text{Var}(T+Y) = \text{var}(T) + \text{var}(Y)$

$= 5 + 32.49$

$= 37.49$

$\therefore \text{Var}(T+Y) = 37.49$

Turn over for the next question

Do not write outside the box

6 The continuous random variable X has probability density function

$$f(x) = \begin{cases} \dfrac{4}{45}(x^3 - 10x^2 + 29x - 20) & 1 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

6 (a) Find $P(X < 2)$

[2 marks]

$P(X<2) = \int_1^2 f(x) = \int_1^2 \frac{4}{45}(x^3 - 10x^2 + 29x - 20)\,dx$

$= \frac{4}{45}\int_1^2 x^3 - 10x^2 + 29x - 20 \; dx$

$\frac{4}{45}\left[\frac{x^4}{4} - \frac{10x^3}{3} + \frac{29x^2}{2} - 20x \Big|_1^2\right]$

$\frac{4}{45}\left[\left(\frac{2^4}{4} - \frac{10(2^3)}{3} + \frac{29}{2}(2^2) - 20(2)\right) - \left(\frac{1}{4} - \frac{10}{3} + \frac{29}{2} - 20\right)\right]$

$\frac{4}{45}\left(-\frac{14}{3} - \left(\frac{-103}{12}\right)\right) = \frac{4}{45}\left(\frac{47}{12}\right)$

$= \frac{47}{135}$

6 (b) Verify that the median of X is 2.3, correct to two significant figures.

[4 marks]

Median for $f(x) = 0.5$, Let the median for $X = m$

$\int_1^m \frac{4}{45}(x^3 - 10x^2 + 29x - 20)\,dx = 0.5$

$= \frac{4}{45}\int_1^m (x^3 - 10x^2 + 29x - 20)\,dx = 0.5$

$\frac{4}{45}\left[\frac{x^4}{4} - \frac{10}{3}x^3 + \frac{29}{2}x^2 - 20x \Big|_1^m\right] = 0.5$

$\frac{4}{45}\left(\frac{m^4}{4} - \frac{10m^3}{3} + \frac{29m^2}{2} - 20m + \frac{103}{12}\right) = 0.5$

$\frac{m^4}{45} - \frac{8m^3}{27} + \frac{58m^2}{45} - \frac{16m}{9} + \frac{71}{270} = 0$

the value of m reduces to $2.2828 \simeq 2.3$

$\Rightarrow$ The median of X is 2.3 correct to two significant figures.

Do not write outside the box

6 (c) Find the mean of X.

[2 marks]

$E(x) = \int x \cdot f(x)\, dx$

$= \frac{4}{45}\int_1^4 x\left(x^3 - 10x^2 + 29x - 20\right) dx$

$= \frac{4}{45}\int_1^4 x^4 - 10x^3 + 29x^2 - 20x\, dx$

$= \frac{4}{45}\left[\frac{x^5}{5} - \frac{10}{4}x^4 + \frac{29}{3}x^3 - 10x^2\Big|_1^4\right]$

Turn over for the next question

$= \frac{4}{45}\left[\left(\frac{4^5}{5} - \frac{10}{4}(4^4) + \frac{29}{3}(4^3)\right) - 10(4^2) - \left(\frac{1}{5} - \frac{10}{4} + \frac{29}{3} - 10\right)\right]$

$= \frac{4}{45}\left(\frac{352}{15} + \frac{79}{30}\right)$

$= \frac{4}{45}\left(\frac{261}{10}\right)$

$= 2.32$

Do not write outside the box

7 A restaurant has asked Sylvia to conduct a χ^2 test for association between meal ordered and age of customer.

7 (a) State the hypotheses that Sylvia should use for her test.

[1 mark]

H_0 : There is no association between meal ordered and age of customer

H_1 : There is an association between meal ordered and age of customer

7 (b) Sylvia correctly calculates her value of the test statistic to be 44.1

She uses a 5% level of significance and the degrees of freedom for the test is 30

Sylvia accepts the null hypothesis.

Explain whether or not Sylvia was correct to accept the null hypothesis.

[4 marks]

$\chi^2 = 44.1$

Critical value $\chi^2_{(30,\ 0.05)} = 43.773$

from the chi-square table.

Critical region is $\chi^2 \geq 43.773$

$44.1 > 43.773$ and lies in the critical region hence we reject H_0.

$\Rightarrow$ Sylvia was not correct to accept the null hypothesis.

Do not write outside the box

7 (c) State in context the correct conclusion to Sylvia's test.

[1 mark]

There is significant evidence to suggest that there is an association between meal ordered and age of customer.

Turn over for the next question

Do not write outside the box

8 There are two hospitals in a city.

Over a period of time, the first hospital recorded an average of 20 births a day.

Over the same period of time, the second hospital recorded an average of 5 births a day.

Stuart claims that birth rates in the hospitals have changed over time.

On a randomly chosen day, he records a total of 16 births from the two hospitals.

8 (a) Investigate Stuart's claim, using a suitable test at the 5% level of significance.

[6 marks]

Let X = Number of births from both hospitals

$X \sim Po(25)$

$H_0 : \lambda = 25$

$H_1 : \lambda \neq 25$

$P(X \leq 16 | \lambda = 25) = P(X=1) + P(X=2) + P(X=3) + P(X=4) + \dots + P(X=16)$

$= e^{-25}\left[\frac{25^0}{0!} + \frac{25^1}{1!} + \frac{25^2}{2!} + \frac{25^3}{3!} + \frac{25^4}{4!} + \frac{25^5}{5!} + \frac{25^6}{6!} + \frac{25^7}{7!} + \frac{25^8}{8!} + \frac{25^9}{9!} + \frac{25^{10}}{10!} + \frac{25^{11}}{11!} + \frac{25^{12}}{12!} + \frac{25^{13}}{13!} + \frac{25^{14}}{14!} + \frac{25^{15}}{15!} + \frac{25^{16}}{16!}\right]$

$= 0.0377$

Because its a two-tailed test, the level of significance is $\frac{0.05}{2} = 0.025$

The critical region is $X \leq 15$ or $X \geq 36$

$\Rightarrow$ $0.0377 > 0.025$ is not in the critical region

hence we fail to reject H_0 (Accept H_0)

$\Rightarrow$ There is no significant evidence to suggest that the total birth rate in the two hospitals has changed.

Do not write outside the box

8 (b) For a test of the type carried out in part **(a)**, find the probability of making a Type I error, giving your answer to two significant figures.

[3 marks]

Type I error

$P(\text{Type I error}) = P(X \le 15) + P(X \geq 36)$

$P(X \le 15) = e^{-25}\left[\frac{25^0}{0!} + \frac{25^1}{1!} + \frac{25^2}{2!} + \frac{25^3}{3!} + \frac{25^4}{4!} + \frac{25^5}{5!} + \frac{25^6}{6!} + \frac{25^7}{7!} + \frac{25^8}{8!} + \frac{25^9}{9!} + \frac{25^{10}}{10!} + \frac{25^{11}}{11!} + \frac{25^{12}}{12!} + \frac{25^{13}}{13!} + \frac{25^{14}}{14!} + \frac{25^{15}}{15!}\right]$

$= 0.02229$

$P(X \geq 36) = 1 - P(X \le 35)$

$P(X \le 35) = e^{-25}\left[\frac{25^0}{0!} + \frac{25^1}{1!} + \frac{25^2}{2!} + \frac{25^3}{3!} + \frac{25^4}{4!} + \frac{25^5}{5!} + \frac{25^6}{6!} + \frac{25^7}{7!} + \frac{25^8}{8!} + \frac{25^9}{9!} + \frac{25^{10}}{10!} + \dots + \frac{25^{35}}{35!}\right]$

$= 0.97754$

$P(X \geq 36) = 1 - 0.97754$

$= 0.02246$

$P(\text{Type I error}) = 0.02229 + 0.02246$

$= 0.04475 \simeq 0.045$

END OF QUESTIONS

$P(\text{Type I error}) = 0.045$ (correct to 3 sf)