SPECIMEN MATERIAL

Please write clearly, in block capitals.

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

Surname ____________________

Forename(s) ____________________

Candidate signature ____________________

AS
FURTHER MATHEMATICS

Paper 2 – Statistics **MODEL ANSWERS**

Exam Date Morning Time allowed: 1 hour 30 minutes

Materials

For this paper you must have:

- You must ensure you have the other optional question paper/answer booklet for which you are entered (**either** Mechanics **or** Discrete). You will have 1 hour 30 minutes to complete both papers.
- The AQA booklet of formulae and statistical tables.
- You may use a graphics calculator.

Instructions

- Use black ink or black ball-point pen. Pencil should be used for drawing.
- Answer **all** questions.
- You must answer each question in the space provided for that question. If you require extra space, use an AQA supplementary answer book; do **not** use the space provided for a different question.
- 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 that 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.

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

1 The random variable T has probability density defined by

$$f(t) = \begin{cases} \dfrac{t}{8} & 0 \le t \le k \\ 0 & \text{otherwise} \end{cases}$$

Find the value of k

[1 mark]

$\frac{1}{16}$ $\frac{1}{4}$ 4 16

$$\int_0^a \frac{t}{8}\,dt = \left[\frac{t^2}{16}\right]_0^a = \frac{1}{16}a^2 = 1$$

$\Rightarrow a^2 = 16$

$\Rightarrow a = 4 \quad (a > 0)$

2 The discrete random variable X has probability distribution defined by

$$P(X = x) = \begin{cases} 0.1 & x = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 \\ 0 & \text{otherwise} \end{cases}$$

Find the value of $P(4 \le X \le 7)$

Circle your answer.

[1 mark]

0.2 0.3 0.4 0.5

$P(4 \le X \le 7) = P(X=4) + P(X=5) + P(X=6) + P(X=7)$

$= 0.1 \times 4$

$= 0.4$

3 The discrete random variable R has the following probability distribution.

r	-2	0	a	4
P($R = r$)	0.3	b	c	0.1

It is known that E(R) = 0.2 and Var(R) = 3.56

Find the values of a, b and c.

[4 marks]

$0.3 + b + c + 0.1 = 1 \Rightarrow b + c = 0.6$

$E(R) = -2(0.3) + 0(b) + ac + 4(0.1) = 0.2$

$\Rightarrow -0.6 + ac + 0.4 = 0.2$

$\Rightarrow ac = 0.4$

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

$= 4(0.3) + 0(b) + a^2c + 16(0.1) - 0.2^2 = 3.56$

$= 1.2 + a^2c + 1.6 - 0.04 = 3.56$

$\Rightarrow a^2c = 0.8$

$\frac{a^2c}{ac} = \frac{0.8}{0.4}$ so $a = 2$

Therefore $b = 0.4$ and $c = 0.2$

Turn over for the next question

4 The number of printers, V, bought during one day from the *Verigood* store can be modelled by a Poisson distribution with mean 4.5

The number of printers, W, bought during one day from the *Winnerprint* store can be modelled by a Poisson distribution with mean 5.5

4 (a) Find the probability that the total number of printers bought during one day from *Verigood* and *Winnerprint* stores is greater than 10.

[2 marks]

$V \sim Po(4.5)$ and $W \sim Po(5.5)$ so $V+W \sim Po(10)$

$P(V+W > 10) = 1 - P(V+W \leq 10)$

$= 0.417$

4 (b) State the circumstance under which the distributional model you used in part **(a)** would not be valid.

[1 mark]

The sales of printers from the verigood and winnerprint stores are not independent

5 Participants in a school jumping competition gain a total score for each jump based on the length, L metres, jumped beyond a fixed point and a mark, S, for style.

L may be regarded as a continuous random variable with probability density function

$$f(l) = \begin{cases} wl & 0 \le l \le 15 \\ 0 & \text{otherwise} \end{cases}$$

where w is a constant.

S may be regarded as a discrete random variable with probability function

$$P(S = s) = \begin{cases} \frac{1}{15}s & s = 1, 2, 3, 4, 5 \\ 0 & \text{otherwise} \end{cases}$$

Assume that L and S are independent.

The total score for a participant in this competition, T, is given by $T = L^2 + \frac{1}{2}S$

Show that the expected total score for a participant is $114\frac{1}{3}$

[5 marks]

$\int_0^{15} wl \, dl = 1 \Rightarrow w\left[\frac{1}{2}l^2\right]_0^{15} = 1$

so $w\left(\frac{1}{2}(15)^2\right) = 1 \Rightarrow w = \frac{2}{15^2} = \frac{2}{225}$

$E(L^2) = \int_0^{15} l^2 (wl) \, dl = w\int_0^{15} l^3 \, dl = \frac{2}{225}\left[\frac{1}{4}l^4\right]_0^{15}$

$= \frac{2}{225}\left(\frac{1}{4}(15)^4\right) = \frac{2}{4}(15)^2 = \frac{225}{2}$

$E(T) = E\left(L^2 + \frac{1}{2}S\right) = E(L^2) + E\left(\frac{1}{2}S\right) = E(L^2) + \frac{1}{2}E(S)$

$$E(S) = \frac{1}{15}(1)(1) + \frac{1}{15}(2)(2) + \frac{1}{16}(3)(3) + \frac{1}{15}(4)(4) + \frac{1}{15}(5)(5)$$

$$= \frac{1}{15}(1^2 + 2^2 + 3^2 + 4^2 + 5^2) = \frac{1}{15}(55) = \frac{11}{3}$$

$$E(T) = E(L^2) + \frac{1}{2}E(S)$$

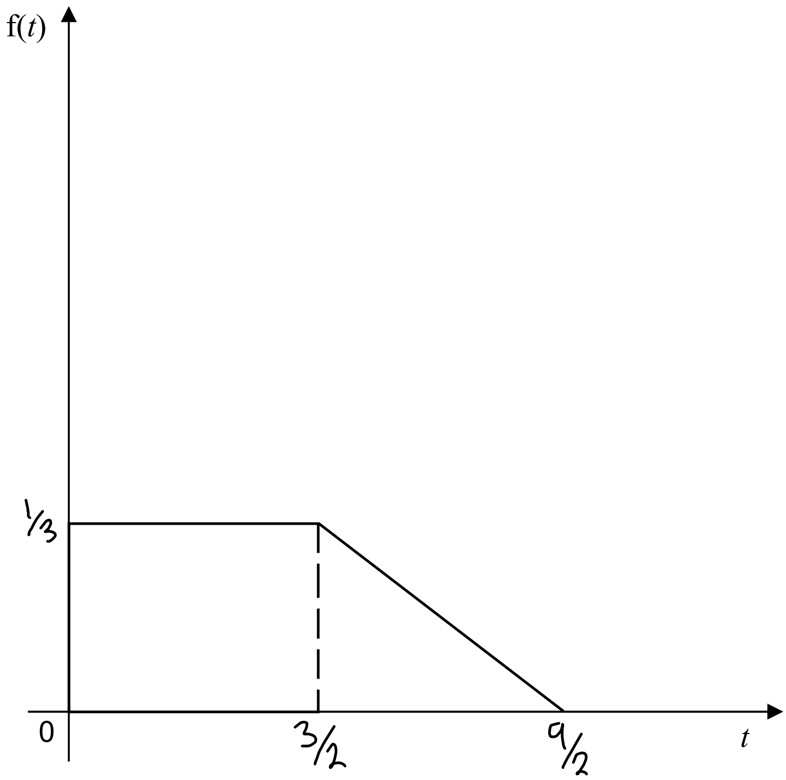
$$= \frac{225}{2} + \frac{1}{2}\left(\frac{11}{3}\right) = \frac{343}{3} = 114\frac{1}{3}$$

6 The continuous random variable T has probability density function defined by

$$f(t) = \begin{cases} \dfrac{1}{3} & 0 \le t \le \dfrac{3}{2} \\ \dfrac{9-2t}{18} & \dfrac{3}{2} \le t \le \dfrac{9}{2} \\ 0 & \text{otherwise} \end{cases}$$

6 (a) (i) Sketch this probability density function below.

[2 marks]



6 (a) (ii) State the median of T.

[1 mark]

Area of the rectangle $= \frac{1}{3}\left(\frac{3}{2}\right) = \frac{1}{2}$

This means that the median of t is $3/2$

6 (b) (i) Find $\mathrm{E}(T)$

[2 marks]

$$E(T) = \int_0^{3/2} \frac{1}{3}t\,dt + \int_{3/2}^{9/2} \frac{t(9-2t)}{18}\,dt$$

$$= \left[\frac{1}{6}t^2\right]_0^{3/2} + \left[\frac{1}{4}t^2 - \frac{1}{27}t^3\right]_{3/2}^{9/2}$$

$$= \frac{1}{6}(3/2)^2 + \frac{1}{4}(9/2)^2 - \frac{1}{27}(9/2)^3 - \frac{1}{4}(3/2)^2 + \frac{1}{27}(3/2)^3$$

$$= \frac{13}{8}$$

$y - y_1 = m(x - x_1)$

$y - 0 = \frac{-1/3}{6/2}(x - 9/2)$

$y = -\frac{1}{9}(x - \frac{9}{2})$

$y = -\frac{2}{18}(x - \frac{9}{2})$

$y = \frac{1}{18}(9 - 2x)$

6 (b) (ii) Given that $\mathrm{E}\left(T^2\right) = \dfrac{15}{4}$, find Var $(4T - 5)$

[3 marks]

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

$$= \frac{15}{4} - \left(\frac{13}{8}\right)^2 = \frac{71}{64}$$

$$Var(4T-5) = 16\,Var(T) = 16\left(\frac{71}{64}\right) = \frac{71}{4}$$

Turn over for the next question

7 A dairy industry researcher, Robyn, decided to investigate the milk yield, classified as low, medium or high, obtained from four different breeds of cow, A, B, C and D.

The milk yield of a sample of 105 cows was monitored and the results are summarised in contingency **Table 1**.

Table 1

		Yield			
		Low	**Medium**	**High**	**Total**
Breed	**A**	4	5	12	**21**
	B	10	6	4	**20**
	C	8	17	7	**32**
	D	5	20	7	**32**
	Total	**27**	**48**	**30**	**105**

The sample of cows may be regarded as random.

Robyn decides to carry out a χ^2-test for association between milk yield and breed using the information given in **Table 1**.

7 (a) Contingency **Table 2** gives some of the expected frequencies for this test.

Complete **Table 2** with the missing expected values.

[2 marks]

Table 2

		Yield		
		Low	**Medium**	**High**
Breed	**A**	5.4	9.6	6
	B	5.14	9.14	5.71
	C	8.23	14.63	9.14
	D	8.23	14.63	9.14

$$E_i = \frac{(\text{row total})(\text{column total})}{\text{table total}}$$

7 (b) (i) For Robyn's test, the test statistic $\sum \frac{(O-E)^2}{E}$ = 19.4 correct to three significant figures.

Use this information to carry out Robyn's test, using the 1% level of significance.

[5 marks]

H_0: Milk yield is independent of breed

H_1: Milk yield is not independent of breed

$\nu = (4-1)(3-1) = (3)(2) = 6$ degrees of freedom

Critical value: 16.81

$19.4 > 16.81$

Therefore, there is evidence that we should reject H_0 and that milk yield is not independent of breed

7 (b) (ii) By considering the observed frequencies given in **Table 1** with the expected frequencies in **Table 2**, interpret, in context, the association, if any, between milk yield and breed.

[2 marks]

The largest values of $|O-E|$ belong to Breed A/High and Breed B/Low so these are the largest sources of association. This means that far more than expected Breed A cows had a high milk yield

8 In a small town, the number of properties sold during a week in spring by a local estate agent, Keith, can be regarded as occurring independently and with constant mean μ. Data from several years have shown the value of μ to be 3.5.

A new housing development was built on the outskirts of the town and the properties on this development were offered for sale by the builder of the development, not by the local estate agents.

During the first **four** weeks in spring, when properties on the new development were offered for sale by the builder, Keith sold a total of 8 properties.

Keith claims that the sale of new properties by the builder reduced his mean number of properties sold during a week in spring.

8 (a) Investigate Keith's claim, using the 5% level of significance.

[6 marks]

$H_0: \lambda = 14 \quad H_1: \lambda < 14$

$T \sim Po(14)$

$P(T \leq 8) = 0.062 > 0.05$

Therefore there is insufficient evidence to reject H_0 and suggest that Keith's mean number of properties sold per week in spring has reduced

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

[1 mark]

Type II Error = P(Accept H_0 | H_0 is false)

In this context, a type II Error is concluding that there is no significant evidence to suggest that Keith's mean number of properties sold per week in spring has reduced when it actually has

8 (c) State **one** advantage and **one** disadvantage of using a 1% significance level rather than a 5% level of significance in a hypothesis test.

[2 marks]

+ Less likely to reject Ho when actually it is true

− Less likely to accept H_1 when Ho is false

END OF QUESTIONS