MODEL ANSWERS

Please check the examination details below before entering your candidate information

Candidate surname	Other names

Pearson Edexcel Level 3 GCE

Centre Number | Candidate Number

Tuesday 18 June 2019

Morning (Time: 1 hour 30 minutes) | Paper Reference **9FM0/3B**

Further Mathematics

Advanced

Paper 3B: Further Statistics 1

You must have:
Mathematical Formulae and Statistical Tables (Green), calculator

Total Marks

Candidates may use any calculator permitted by Pearson regulations. Calculators must not have the facility for algebraic manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions

- Use **black** ink or ball-point pen.
- If pencil is used for diagrams/sketches/graphs it must be dark (HB or B).
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions and ensure that your answers to parts of questions are clearly labelled.
- Answer the questions in the spaces provided – *there may be more space than you need.*
- You should show sufficient working to make your methods clear. Answers without working may not gain full credit.
- Answers should be given to three significant figures unless otherwise stated.

Information

- A booklet 'Mathematical Formulae and Statistical Tables' is provided.
- There are 7 questions in this question paper. The total mark for this paper is 75.
- The marks for **each** question are shown in brackets – *use this as a guide as to how much time to spend on each question.*

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.

Turn over

1/1/1/1/1/

Pearson

Answer ALL questions. Write your answers in the spaces provided.

1. A chocolate manufacturer places special tokens in 2% of the bars it produces so that each bar contains at most one token. Anyone who collects 3 of these tokens can claim a prize.

Andreia buys a box of 40 bars of the chocolate.

(a) Find the probability that Andreia can claim a prize.
(2)

Barney intends to buy bars of the chocolate, one at a time, until he can claim a prize.

(b) Find the probability that Barney can claim a prize when he buys his 40th bar of chocolate.
(3)

(c) Find the expected number of bars that Barney must buy to claim a prize.
(1)

(a) Let X = no. of prizes Andreia wins out of 40 bars
$2\% = 0.02$
$\therefore X \sim B(40, 0.02)$ Binomial Distribution

To claim a prize Andreia needs 3 or more tokens $\Rightarrow X \geq 3$

$P(X \geq 3) = 1 - P(X \leq 2) = 1 - 0.95432... = 0.04567... \approx 0.0457$ (4dp)

$\boxed{\therefore P(X \geq 3) = 0.0457}$

(b.) Let Y = no. of the bar when Barney can claim a prize
Need 3 tokens to claim a prize.
$\therefore Y \sim \text{NegBin}(3, 0.02)$ Negative Binomial Distribution

Let $Z \sim B(39, 0.02)$ → Used to find probability of Barney winning 2 tokens out of 39 bars, before he can win the 3rd token at the 40th bar.

$P(Y = 40) = P(Z = 2) \times 0.02 = 0.14035... \times 0.02 = 0.002807... \approx 0.0028$ (4dp)

$\boxed{\therefore P(Y = 40) = 0.0028}$

DO NOT WRITE IN THIS AREA

Question 1 continued

(c.) Expected Value = Mean = $\frac{r}{p}$

$\boxed{\therefore E(Y) = \frac{3}{0.02} = 150}$

(Total for Question 1 is 6 marks)

P 6 1 1 8 0 A 0 3 2 4

Turn over

2. Indre works on reception in an office and deals with all the telephone calls that arrive. Calls arrive randomly and, in a 4-hour morning shift, there are on average 80 calls.

(a) Using a suitable model, find the probability of more than 4 calls arriving in a particular 20-minute period one morning.
(3)

Indre is allowed 20 minutes of break time during each 4-hour morning shift, which she can take in 5-minute periods. When she takes a break, a machine records details of any call in the office that Indre has missed.

One morning Indre took her break time in 4 periods of 5 minutes each.

(b) Find the probability that in exactly 3 of these periods there were no calls.
(2)

On another occasion Indre took 1 break of 5 minutes and 1 break of 15 minutes.

(c) Find the probability that Indre missed exactly 1 call in each of these 2 breaks.
(3)

(a.) Let C = no. of calls in a 20-min period

$\lambda = 80$ calls in 4 hours Poisson Distribution

λ : Mins

$\div 12$ (80 : 240) $\div 12$

$\frac{20}{3}$: 20

$\therefore \lambda = \frac{20}{3}$ calls in 20 mins

$\therefore C \sim Po\left(\frac{20}{3}\right)$

$P(C>4) = 1 - P(C \leq 4) = 1 - 0.20562... = 0.79437... \approx 0.7944$ (4dp)

$\boxed{\therefore P(C>4) = 0.7944}$

(b) Let X = no. of 5-min period with no calls

$\lambda = \frac{20}{3} \div \frac{20}{5} = \frac{5}{3}$ calls in 5 mins $\longrightarrow \therefore Y \sim Po\left(\frac{5}{3}\right)$

Probability of no calls $= P(Y=0) = 0.18887...$

$\therefore X \sim B(4, 0.18887...)$ $\longrightarrow$ $\boxed{\therefore P(X=3) = 0.02185... \approx 0.0219 \text{ (4dp)}}$

Total of 4 periods.

P 6 1 1 8 0 A 0 4 2 4

Question 2 continued

(c.) $\lambda = \frac{5}{3}$ calls in 5 mins $\rightarrow$ $\therefore Y \sim Po\left(\frac{5}{3}\right)$

$\lambda = \frac{5}{3} \times \frac{15}{5} = 5$ calls in 15 mins $\rightarrow$ $\therefore Z \sim Po(5)$

P(Misses exactly one call in each break) $= P(Y=1) \times P(Z=1)$

$= 0.31479... \times 0.03368...$

$= 0.01060...$

≈ 0.0106 (4dp)

$\boxed{\therefore \text{P(Misses exactly one call in each break)} = 0.0106}$

Turn over

Question 2 continued

Question 2 continued

(Total for Question 2 is 8 marks)

3. A biased spinner can land on the numbers 1, 2, 3, 4 or 5 with the following probabilities.

Number on spinner	1	2	3	4	5
Probability	0.3	0.1	0.2	0.1	0.3

The spinner will be spun 80 times and the mean of the numbers it lands on will be calculated.
Find an estimate of the probability that this mean will be greater than 3.25

(6)

Let X = the number when the spinner is spun

Mean, $E(X) = 1(0.3) + 2(0.1) + 3(0.2) + 4(0.1) + 5(0.3) = 3 = \mu$

$E(X^2) = 1^2(0.3) + 2^2(0.1) + 3^2(0.2) + 4^2(0.1) + 5^2(0.3) = 11.6$

Variance, $\sigma^2 = E(X^2) - E(X)^2 = 11.6 - 3^2 = 2.6$

Central Limit Theorem : $\bar{X} \approx\sim N\left(\mu, \frac{\sigma^2}{n}\right)$

$\bar{X} \approx\sim N\left(3, \frac{2.6}{80}\right) \Rightarrow \therefore \bar{X} \approx\sim N\left(3, \left(\frac{\sqrt{13}}{20}\right)^2\right) \Rightarrow \therefore Z \sim N\left(3, \left(\frac{\sqrt{13}}{20}\right)^2\right)$

$P(\bar{X} > 3.25) = P(Z > 3.25) = 0.08275\ldots \approx 0.0828$ (4dp)

$\boxed{\therefore P(\bar{X} > 3.25) = 0.0828}$

Question 3 continued

(Total for Question 3 is 6 marks)

4. Liam and Simone are studying the distribution of oak trees in some woodland. They divided the woodland into 80 equal squares and recorded the number of oak trees in each square. The results are summarised in Table 1 below.

Number of oak trees in a square	0	1	2	3	4	5	6	7 or more
Frequency	1	4	21	23	13	11	7	0

Table 1

Liam believes that the oak trees were deliberately planted, with 6 oak trees per square and that a constant proportion p of the oak trees survived.

(a) Suggest the model Liam should use to describe the number of oak trees per square.

(2)

Liam decides to test whether or not his model is suitable and calculates the expected frequencies given in Table 2.

Number of oak trees in a square	0 or 1	2	3	4	5	6
Expected frequency	5.53	14.89	24.26	22.24	10.87	2.21

$2.21 < 5$
$\therefore$ Combine colums 5 & 6.

Table 2

0.05

(b) Showing your working clearly, complete the test using a 5% level of significance. You should state your critical value and conclusion clearly.

(7)

Simone believes that a Poisson distribution could be used to model the number of oak trees per square. She calculates the expected frequencies given in Table 3.

Number of oak trees in a square	0 or 1	2	3	4	5	6 or more
Expected frequency	12.69	16.07	s	14.58	t	9.37

Table 3

(c) Find the value of s and the value of t, giving your answers to 2 decimal places.

(4)

(d) Write down hypotheses to test the suitability of Simone's model.

(1)

The test statistic for this test is 8.749

0.05

(e) Complete the test. Use a 5% level of significance and state your critical value and conclusion clearly.

(3)

(f) Using the results of these tests, explain whether the origin of this woodland is likely to be cultivated or wild.

(2)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

P 6 1 1 8 0 A 0 1 0 2 4

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

Question 4 continued

(a) Let T = no. of oak trees in a square
∴ Using binomial distribution, $T \sim B(6, p)$

(b)

No. of oak trees in a square	0 or 1	2	3	4	5 or 6
O_i, Observed Frequency	5	21	23	13	18
E_i, Expected Frequency	5.53	14.89	24.26	22.24	13.08
$\frac{(O_i - E_i)^2}{E_i}$	0.051	2.51	0.0654	3.84	1.85

Goodness of Fit Test: $\sum \frac{(O_i - E_i)^2}{E_i} = 0.051 + 2.51 + 0.0654 + 3.84 + 1.85$
$= 8.3129\ldots$

Degrees of Freedom: $\nu = 5 - \underset{\downarrow}{2} = 3$
Subtracting columns 0 or 1 & 5 or 6.

At $\nu = 3$, $\chi^2(0.05) = 7.815$ → Critical Value

$\boxed{8.3129\ldots > 7.815 \quad \therefore \text{Liam's model is not suitable.}}$

(c.) Let R = no. of oak trees in a square for Simone's model

λ for Poisson Distribution $= E(X), \mu$

$\therefore \lambda = \frac{0(1) + 1(4) + 2(21) + 3(23) + 4(13) + 5(11) + 6(7)}{80} = 3.3$

$\therefore R \sim P_o(3.3)$

$s = P(R = 3) \times 80 = 0.220911\ldots \times 80 = 17.672\ldots \approx 17.67$ (2dp)
$t = P(R = 5) \times 80 = 0.12028\ldots \times 80 = 9.622\ldots \approx 9.62$ (2dp)

$\boxed{\therefore s = 17.67,\ t = 9.62}$

Question 4 continued

(d.) H_0 : Poisson is a good fit for no. of oak trees per square.
H_1 : Poisson is not a good fit for no. of oak trees per square.

(e.) Goodness of Fit Test : $\chi^2 = 8.749$

Degrees of Freedom : $\nu = 6 - 2 = 4$
↳ Subtracting columns 0 or 1 & 6 or more.

At $\nu = 4$, $\chi^2(0.05) = 9.488$ → Critical Value

$8.749 < 9.488$ ∴ Not enough evidence to reject Ho.
∴ Accept Ho.
∴ Poisson is a suitable model.

(f.) Poisson model has a better fit, suggesting oak trees occur at random, whereas binomial suggests oak trees have been deliberately planted or cultivated.

∴ The forest is likely to be wild, not cultivated.

Question 4 continued

(Total for Question 4 is 19 marks)

5. Information was collected about accidents on the *Seapron* bypass. It was found that the number of accidents per month could be modelled by a Poisson distribution with mean 2.5

$\lambda = 2.5$

Following some work on the bypass, the numbers of accidents during a series of 3-month periods were recorded. The data were used to test whether or not there was a change in the mean number of accidents per month.

(a) Stating your hypotheses clearly and using a 5% level of significance, find the critical region for this test. You should state the probability in each tail.

(5)

(b) State P(Type I error) using this test.

(1)

Data from the series of 3-month periods are recorded for 2 years.

(c) Find the probability that at least 2 of these 3-month periods give a significant result.

(3)

Given that the number of accidents per month on the bypass, after the work is completed, is actually 2.1 per month,

(d) find P(Type II error) for the test in part (a)

(3)

(a.) $H_0: \lambda = 2.5$

$H_1: \lambda \neq 2.5$ → 2-tailed test, ∴ level of significance $= \frac{5}{2} = 0.025$ at each tail

Let X = no. of accidents in a 3-month period

∴ $\lambda = 2.5 \times 3 = 7.5$

∴ $X \sim Po(7.5)$

$P(X \leq c_1) < 0.05$ & $P(X \geq c_1) < 0.05$, where c_1 & c_2 are critical values.

$P(X \leq 2) = 0.02025... < 0.025$

$P(X \leq 3) = 0.05914... > 0.025$

∴ $c_1 = 2$

$P(X \geq 13) = 1 - P(X \leq 12) = 1 - 0.95734... = 0.04266... > 0.025$

$P(X \geq 14) = 1 - P(X \leq 13) = 1 - 0.97843... = 0.02156... < 0.025$

∴ $c_2 = 14$

$\boxed{\therefore X \leq 2 \text{ or } X \geq 14}$

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

Question 5 continued

(b) $P(\text{Type I Error}) = 0.0203 + 0.0216 = \boxed{0.0419}$ (4 dp)

(c.) Let M = no. of 3-month periods with a significant result
$n = 2(12) \div 3 = 8$
$p = 0.0419$

$\therefore M \sim B(8, 0.0419)$ Binomial Distribution

$P(M \geq 2) = 1 - P(M \leq 1) = 1 - 0.95846... = 0.04153... \approx 0.0415$ (4dp)

$\boxed{\therefore P(M \geq 2) = 0.0415}$

(d.) $\lambda = 2.1 \times 3 = 6.3$ for a 3-month period $\Rightarrow \therefore Y \sim Po(6.3)$

$P(\text{Type II Error}) = P(3 \leq Y \leq 13)$
$= P(Y \leq 13) - P(Y \leq 2)$
$= 0.99451... - 0.04984....$
$= 0.94467...$
≈ 0.9447 (4dp)

$\boxed{\therefore P(\text{Type II Error}) = 0.9447}$

Question 5 continued

Question 5 continued

(Total for Question 5 is 12 marks)

6. The discrete random variable X has probability generating function

$$G_X(t) = k\ln\left(\frac{2}{2-t}\right)$$

where k is a constant.

(a) Find the exact value of k **(1)**

(b) Find the exact value of Var(X) **(7)**

(c) Find P(X = 3) **(4)**

(a) $G(1) = 1$

$k\ln\left(\frac{2}{2-1}\right) = 1$

$k\ln 2 = 1 \Rightarrow \boxed{\therefore\ k = \frac{1}{\ln 2}}$

(b) $E(X) = G_X'(1)$

$\text{Var}(X) = G_X''(1) + G_X'(1) - [G_X'(1)]^2$

$G(t) = \frac{1}{\ln 2}(\ln 2 - \ln(2-t))$

$G'(t) = \frac{1}{\ln 2}\left(\frac{1}{2-t}\right) = \frac{(2-t)^{-1}}{\ln 2} \rightarrow G'(1) = \frac{(2-1)^{-1}}{\ln 2} = \frac{1}{\ln 2}$

$G''(t) = \frac{1}{\ln 2}\left(\frac{1}{(2-t)^2}\right) = \frac{(2-t)^{-2}}{\ln 2} \rightarrow G''(1) = \frac{(2-1)^{-1}}{\ln 2} = \frac{1}{\ln 2}$

$\text{Var}(X) = \frac{1}{\ln 2} + \frac{1}{\ln 2} - \left(\frac{1}{\ln 2}\right)^2 \rightarrow \boxed{\therefore\ \text{Var}(X) = \frac{1}{\ln 2}\left(2 - \frac{1}{\ln 2}\right)}$

Question 6 continued

(c) $P(X=3)$ needs coefficient of t^3 in the Maclaurin series

$$G'''(t) = \frac{1}{\ln 2} \times \frac{2}{(2-t)^3} = \frac{2(2-t)^{-3}}{\ln 2}$$

$$P(X=3) = \frac{G'''(0)}{3!} = \frac{2(2-0)^{-3}}{\ln 2} \div 6 = \frac{1}{4\ln 2} \times \frac{1}{6} = \frac{1}{24\ln 2}$$

$$\therefore P(X=3) = \frac{1}{24\ln 2} \approx 0.0601 \text{ (4dp)}$$

Question 6 continued

Question 6 continued

(Total for Question 6 is 12 marks)

7. A spinner can land on red or blue. When the spinner is spun, there is a probability of $\frac{1}{3}$ that it lands on blue. The spinner is spun repeatedly.

The random variable B represents the number of the spin when the spinner first lands on blue.

(a) Find (i) $P(B = 4)$

(ii) $P(B \leqslant 5)$

(4)

(b) Find $E(B^2)$

(3)

Steve invites Tamara to play a game with this spinner.

Tamara must choose a colour, either red or blue.

Steve will spin the spinner repeatedly until the spinner first lands on the colour Tamara has chosen. The random variable X represents the number of the spin when this occurs.

If Tamara chooses red, her score is e^X

If Tamara chooses blue, her score is X^2

(c) State, giving your reasons and showing any calculations you have made, which colour you would recommend that Tamara chooses.

(5)

(a.) $B \sim \text{Geo}\left(\frac{1}{3}\right)$ Geometric Distribution

(i) $P(B=4) = \left(\frac{2}{3}\right)^3 \times \frac{1}{3} = \boxed{\frac{8}{81} \approx 0.0988}$ (4dp)

(ii) $P(B \leq 5) = 1 - P(B > 5) = 1 - \left(\frac{2}{3}\right)^5 = \boxed{\frac{211}{243} \approx 0.8683}$ (4dp)

(b.) $\text{Var}(B) = E(B^2) - E(B)^2 \Rightarrow \therefore E(B^2) = \text{Var}(B) + E(B)^2$

$\text{Var}(B) = \frac{1-p}{p^2} = \frac{1-\frac{1}{3}}{\left(\frac{1}{3}\right)^2} = 6$

$E(B) = \frac{1}{p} = \frac{1}{\left(\frac{1}{3}\right)} = 3$

$\boxed{\therefore E(B^2) = 6 + 3^2 = 15}$

Question 7 continued

(c.) Let R = no. of spins when it first lands on red

$\therefore p = \frac{2}{3}$ $\therefore R \sim Geo\left(\frac{2}{3}\right)$ $\therefore X = R \sim Geo\left(\frac{2}{3}\right)$

Expected value of red $= E(e^X) = \sum_{x=1}^{\infty} e^x \cdot p(1-p)^{x-1}$

$= \sum_{x=1}^{\infty} e^x \left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^{x-1}$

$= \frac{2e}{3} \sum_{x=1}^{\infty} \left(\frac{e}{3}\right)^{x-1}$

For geometric series: $S_\infty = \frac{a}{1-r}$

$= \frac{2e}{3} \times \frac{1}{1-\frac{e}{3}}$

$= \frac{2e}{3-e}$

$\therefore E(e^X) = \frac{2e}{3-e} = 19.297... > 15 \rightarrow E(B^2) = 15$

$\boxed{\therefore \text{ Tamara should choose red since it has a greater expected value.}}$

Question 7 continued

(Total for Question 7 is 12 marks)

TOTAL FOR PAPER IS 75 MARKS