Please check the examination details below before entering your candidate information

Candidate surname StemUp Model Solution | 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 ▶

P61180A

©2019 Pearson Education Ltd.

1/1/1/1/1/

P 6 1 1 8 0 A 0 1 2 4

Pearson

stemup.co.uk

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) X = number of tokens in 40 bars, $X \sim B(40, 0.02)$

$P(X \geqslant 3) = 1 - P(X \leqslant 2)$

$= 1 - 0.954329...$

$= 0.045670... \approx 0.0457$

(b) Y = number of the bar on which Barney gets his 3rd token

$Y \sim$ Negative Binomial$(3, 0.02)$

$P(Y = 40) = {}^{39}C_2 \times 0.02^2 \times 0.98^{37} \times 0.02$

$= 0.0028071... \approx 0.00281$

(c) $E(Y) = \frac{r}{p} = \frac{3}{0.02} = 150$ bars

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

stemup.co.uk

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

Question 1 continued

(Total for Question 1 is 6 marks)

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) C = number of calls in a 20-minute period

80 calls in 240 minutes $\Rightarrow \lambda = 80 \times \frac{20}{240} = \frac{20}{3}$

$C \sim Po(\frac{20}{3})$ (calls arrive randomly, singly, at a constant rate)

$P(C > 4) = 1 - P(C \leq 4)$

$= 1 - 0.205627...$

$= 0.794372... \approx 0.794$

(b) Calls in a 5-minute period $\sim Po(\frac{5}{3})$

$P(\text{no calls in a 5-minute period}) = e^{-5/3} = 0.188875...$

X = number of the 4 periods with no calls, $X \sim B(4, 0.188875...)$

$P(X = 3) = 4 \times 0.188875...^3 \times (1 - 0.188875...)$

$= 0.021861... \approx 0.0219$

(c) S = calls in the 5-minute break, $S \sim Po(\frac{5}{3})$

T = calls in the 15-minute break, $T \sim Po(5)$

$P(S = 1) = e^{-5/3} \times \frac{5}{3} = 0.314792...$

$P(T = 1) = e^{-5} \times 5 = 0.033689...$

$P(S = 1) \times P(T = 1) = 0.314792... \times 0.033689...$ (independent)

$= 0.0106052... \approx 0.0106$

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

stemup.co.uk

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

Question 2 continued

Question 2 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

stemup.co.uk

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

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 be the number the spinner lands on.

$E(X) = 1\times0.3 + 2\times0.1 + 3\times0.2 + 4\times0.1 + 5\times0.3 = 3$

$E(X^2) = 1\times0.3 + 4\times0.1 + 9\times0.2 + 16\times0.1 + 25\times0.3 = 11.6$

$Var(X) = E(X^2) - [E(X)]^2 = 11.6 - 3^2 = 2.6$

$n = 80$ is large, so by the Central Limit Theorem

$\bar{X} \approx N(3, \frac{2.6}{80})$, i.e. $\bar{X} \approx N(3, 0.0325)$

standard deviation $= \sqrt{0.0325} = 0.180277...$

$P(\bar{X} > 3.25) = P(Z > \frac{3.25 - 3}{0.180277...})$

$= P(Z > 1.38675...)$

$= 1 - 0.917241...$

$= 0.0827589... \approx 0.0828$

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

stemup.co.uk

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

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

Table 2

(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

(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

stemup.co.uk

Question 4 continued

(a) T = number of oak trees in a square

$T \sim B(6, p)$ (binomial: 6 trees planted per square, each surviving independently with constant probability p)

(b) H_0: $B(6, p)$ is a suitable model for the number of oak trees per square

H_1: $B(6, p)$ is not a suitable model

$\text{Mean} = \frac{0\times1 + 1\times4 + 2\times21 + 3\times23 + 4\times13 + 5\times11 + 6\times7}{80} = \frac{264}{80} = 3.3$

so $\hat{p} = 3.3 \div 6 = 0.55$

Expected frequency for 6 is 2.21 < 5, so pool the classes 5 and 6:

$E = 10.87 + 2.21 = 13.08$, $O = 11 + 7 = 18$

Trees	0 or 1	2	3	4	5 or 6
O	5	21	23	13	18
E	5.53	14.89	24.26	22.24	13.08
$(O - E)^2/E$	0.0508	2.5072	0.0654	3.8389	1.8506

Test statistic $X^2 = \Sigma \frac{(O - E)^2}{E} = 8.313$

$\nu = 5 - 1 - 1 = 3$ (one parameter, p, estimated)

Critical value $\chi^2_3(0.05) = 7.815$

$8.313 > 7.815$, so the result is significant: reject H_0

There is evidence that Liam's binomial model is not suitable for the number of oak trees per square.

(c) R = number of oak trees in a square, $R \sim Po(3.3)$ (λ = mean = 3.3)

$s = 80 \times P(R = 3) = 80 \times \frac{e^{-3.3} \times 3.3^3}{3!} = 17.6729...$

$s = 17.67$

$t = 80 \times P(R = 5) = 80 \times \frac{e^{-3.3} \times 3.3^5}{5!} = 9.62291...$

$t = 9.62$

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

Question 4 continued

(d) H_0: a Poisson distribution is a suitable model for the number of oak trees per square

H_1: a Poisson distribution is not a suitable model

(e) All expected frequencies are at least 5, so no pooling: 6 classes

$\nu = 6 - 1 - 1 = 4$ (one parameter, λ, estimated)

Critical value $\chi^2_4(0.05) = 9.488$

$8.749 < 9.488$, so the result is not significant: do not reject H_0

There is insufficient evidence to reject Simone's model: the Poisson model is suitable for the number of oak trees per square.

(f) The binomial model (deliberately planted) is rejected but the Poisson model is a good fit, so the oak trees occur at random in the squares.

Therefore the woodland is likely to be wild, not cultivated.

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

stemup.co.uk

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

Question 4 continued

(Total for Question 4 is 19 marks)

stemup.co.uk

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

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) λ = mean number of accidents per month

H_0: $\lambda = 2.5$ H_1: $\lambda \neq 2.5$

X = number of accidents in a 3-month period; under H_0 $X \sim Po(7.5)$

Lower tail: $P(X \leqslant 2) = 0.0203$, $P(X \leqslant 3) = 0.0591 > 0.025$

Upper tail: $P(X \geqslant 14) = 1 - P(X \leqslant 13) = 1 - 0.9784 = 0.0216$

$P(X \geqslant 13) = 1 - P(X \leqslant 12) = 0.0427 > 0.025$

Critical region: $X \leqslant 2$ or $X \geqslant 14$

Probability in lower tail = 0.0203, probability in upper tail = 0.0216

(b) P(Type I error) = 0.0203 + 0.0216 = 0.0419

(c) 2 years = 8 three-month periods

M = number of periods giving a significant result, $M \sim B(8, 0.0419)$

$P(M \geqslant 2) = 1 - P(M \leqslant 1)$

$= 1 - (0.9581^8 + 8 \times 0.0419 \times 0.9581^7)$

$= 1 - 0.958462...$

$= 0.041537... \approx 0.0415$

(d) Actual $\lambda = 2.1$, so in a 3-month period $Y \sim Po(6.3)$

P(Type II error) = $P(3 \leqslant Y \leqslant 13) = P(Y \leqslant 13) - P(Y \leqslant 2)$

$= 0.994514... - 0.049846... = 0.94466... \approx 0.945$

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

stemup.co.uk

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

Question 5 continued

Question 5 continued

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

stemup.co.uk

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

Question 5 continued

(Total for Question 5 is 12 marks)

stemup.co.uk

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

$k = \frac{1}{\ln 2}$

(b) $G_X(t) = \frac{1}{\ln 2}\left[\ln 2 - \ln(2-t)\right]$

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

$E(X) = G'(1) = \frac{1}{\ln 2}$

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

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

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

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

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

$\text{Var}(X) = \frac{1}{\ln 2}\left(2 - \frac{1}{\ln 2}\right)$

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

stemup.co.uk

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

Question 6 continued

(c) By Maclaurin, $P(X = 3)$ = coefficient of t^3 in $G_X(t) = \frac{G'''(0)}{3!}$

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

$G'''(0) = \frac{2}{8 \ln 2} = \frac{1}{4 \ln 2}$

$P(X = 3) = \frac{1}{4 \ln 2} \div 6 = \frac{1}{24 \ln 2}$

$= 0.0601122\ldots \approx 0.0601$



Question 6 continued

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

stemup.co.uk

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

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)(i) $B \sim Geo(\frac{1}{3})$

$P(B = 4) = (\frac{2}{3})^3 \times \frac{1}{3} = \frac{8}{81}$

(ii) $P(B \leq 5) = 1 - P(B > 5) = 1 - (\frac{2}{3})^5 = 1 - \frac{32}{243} = \frac{211}{243}$

(b) $E(B) = \frac{1}{p} = 3$, $Var(B) = \frac{1-p}{p^2} = \frac{2/3}{1/9} = 6$

$E(B^2) = Var(B) + [E(B)]^2 = 6 + 3^2 = 15$

(c) Blue: $X = B$, so the expected score is $E(B^2) = 15$

Red: $X = R$, the spin on which red first appears, $R \sim Geo(\frac{2}{3})$

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

stemup.co.uk

Question 7 continued

Expected score $E(e^R) = \sum_{x=1}^{\infty} e^x \times \left(\frac{1}{3}\right)^{x-1} \times \frac{2}{3}$

$= \frac{2e}{3} \sum_{x=1}^{\infty} \left(\frac{e}{3}\right)^{x-1}$ (geometric series with $r = \frac{e}{3} < 1$)

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

$= 19.2978...$

$19.2978... > 15$, so the expected score is greater when choosing red.

I recommend that Tamara chooses red, since her expected score is about 19.3 compared with 15 for blue.

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

Question 7 continued

(Total for Question 7 is 12 marks)

TOTAL FOR PAPER IS 75 MARKS

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

stemup.co.uk