Please check the examination details below before entering your candidate information

Candidate surname StemUp Model Solution | Other names

Centre Number | Candidate Number

Pearson Edexcel Level 3 GCE

Friday 14 June 2024

Afternoon (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 symbolic algebra 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.
- Values from statistical tables should be quoted in full. If a calculator is used instead of the tables the value should be given to an equivalent degree of accuracy.
- Inexact 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 ▶

P72797A

©2024 Pearson Education Ltd.
F:1/1/1/1/

P 7 2 7 9 7 A 0 1 2 4

Pearson

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

x	-1	0	1	3	5
$P(X = x)$	0.2	0.1	0.2	0.25	0.25

(a) Find Var(X)

(3)

(b) Find Var(X^2)

(3)

(a) $E(X) = -1\times0.2 + 0\times0.1 + 1\times0.2 + 3\times0.25 + 5\times0.25$

$= -0.2 + 0 + 0.2 + 0.75 + 1.25 = 2$

$E(X^2) = (-1)^2\times0.2 + 0 + 1^2\times0.2 + 3^2\times0.25 + 5^2\times0.25$

$= 0.2 + 0.2 + 2.25 + 6.25 = 8.9$

$Var(X) = E(X^2) - [E(X)]^2 = 8.9 - 2^2$

$Var(X) = 4.9$

(b) $E(X^4) = (-1)^4\times0.2 + 0 + 1^4\times0.2 + 3^4\times0.25 + 5^4\times0.25$

$= 0.2 + 0.2 + 20.25 + 156.25 = 176.9$

$Var(X^2) = E(X^4) - [E(X^2)]^2$

$= 176.9 - 8.9^2$

$= 176.9 - 79.21$

$Var(X^2) = 97.69$

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
DO NOT WRITE IN THIS AREA

Question 1 continued

(Total for Question 1 is 6 marks)

2. The number of errors made by a secretary is modelled by a Poisson distribution with a mean of 2.4 per 100 words.

A 100-word piece of work completed by the secretary is selected at random.

(a) Find the probability that

(i) there are exactly 3 errors,

(ii) there are fewer than 2 errors.

(2)

After a long holiday, a randomly selected piece of work containing 250 words completed by the secretary is examined to see if the rate of errors has changed.

(b) Stating your hypotheses clearly, and using a 5% level of significance, find the critical region for a suitable test.

(4)

(c) Find P(Type I error) for the test in part (b)

(1)

(a) X = number of errors in 100 words, $X \sim Po(2.4)$

(i) $P(X = 3) = \frac{e^{-2.4} \times 2.4^3}{3!} = 0.209014... = 0.209$

(ii) $P(X < 2) = P(X \leq 1) = e^{-2.4}(1 + 2.4) = 0.308441... = 0.308$

(b) λ = mean number of errors per 100 words

H_0: $\lambda = 2.4$ H_1: $\lambda \neq 2.4$

E = number of errors in 250 words; under H_0 $E \sim Po(6)$

Two-tailed test: 2.5% in each tail

Lower tail: $P(E \leq 1) = 0.0174 < 0.025$, $P(E \leq 2) = 0.0620 > 0.025$

Upper tail: $P(E \geq 12) = 1 - P(E \leq 11) = 1 - 0.9799 = 0.0201 < 0.025$

$P(E \geq 11) = 1 - P(E \leq 10) = 1 - 0.9574 = 0.0426 > 0.025$

Critical region: $E \leq 1$ or $E \geq 12$

(c) $P(\text{Type I error}) = P(E \leq 1) + P(E \geq 12)$

$= 0.0174 + 0.0201$

$= 0.0375$

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
DO NOT WRITE IN THIS AREA

Question 2 continued

(Total for Question 2 is 7 marks)

3. Tisam took a survey of students' favourite colours. The results are summarised in the table below.

		Colour					
		Red	**Blue**	**Green**	**Yellow**	**Black**	**Total**
Year group	**1–5**	34	15	14	22	3	88
	6–9	23	32	12	9	8	84
	10–12	5	28	19	8	8	68
	Total	62	75	45	39	19	240

Tisam carries out a suitable test to see if there is any association between favourite colour and year group.

(a) Write down the hypotheses for a suitable test.

(1)

For her table, Tisam only needs to check one cell to show that none of the expected frequencies are less than 5

(b) (i) Identify this cell, giving your reason.

(ii) Calculate the expected frequency for this cell.

(2)

The test statistic for Tisam's test is 38.449

(c) Using a 1% level of significance, complete the test.
You should state your critical value and conclusion clearly.

(3)

(a) H_0: there is no association between favourite colour and year group

H_1: there is an association between favourite colour and year group

(b)(i) Black and Year group 10–12: it has the lowest row total (68) and the lowest column total (19), so the smallest expected frequency.

(ii) Expected frequency $= \frac{68 \times 19}{240} = 5.38333... = 5.38$

(c) $\nu = (5 - 1)(3 - 1) = 8$

Critical value $\chi^2_8(1\%) = 20.090$

$38.449 > 20.090$, so the result is significant: reject H_0.

There is evidence of an association between favourite colour and year group (students in different year groups have different favourite colours).

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
DO NOT WRITE IN THIS AREA

Question 3 continued

(Total for Question 3 is 6 marks)

4. Every morning Geethaka repeatedly rolls a fair, six-sided die until he rolls a 3 and then he stops. The random variable X represents the number of times he rolls the die each morning.

(a) Suggest a suitable model for the random variable X

(1)

(b) Show that $P(X \leqslant 3) = \dfrac{91}{216}$

(2)

After 64 mornings Geethaka will calculate the mean number of times he rolled the die.

(c) Estimate the probability that the mean number of rolls is between 5.6 and 7.2

(5)

Nira wants to check Geethaka's die to decide whether or not the probability of rolling a 3 with his die is less than $\dfrac{1}{6}$

Nira rolls the die repeatedly until she rolls a 3
She obtains $x = 16$

(d) By carrying out a suitable test, determine what Nira's conclusion should be. You should state your hypotheses clearly and use a 5% level of significance.

(4)

(a) $X \sim Geo(\frac{1}{6})$, geometric distribution with $p = \frac{1}{6}$

(b) $P(X \leq 3) = 1 - P(X > 3) = 1 - (\frac{5}{6})^3$

$= 1 - \frac{125}{216} = \frac{91}{216}$ as required

(c) $E(X) = \frac{1}{p} = 6$ $\quad Var(X) = \frac{1-p}{p^2} = \frac{5/6}{1/36} = 30$

By the Central Limit Theorem, with $n = 64$:

$\bar{X} \approx N(6, \frac{30}{64})$, i.e. $\bar{X} \approx N(6, 0.46875)$

standard deviation $= \sqrt{0.46875} = 0.684653...$

$P(5.6 < \bar{X} < 7.2) = P(\frac{5.6-6}{0.684653...} < Z < \frac{7.2-6}{0.684653...})$

$= P(-0.58424 < Z < 1.75271)$

$= 0.960175... - 0.279531...$

$= 0.680644... = 0.681$

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 4 continued

(d) p = probability of rolling a 3 with the die

$H_0: p = \frac{1}{6}$ $\quad$ $H_1: p < \frac{1}{6}$

Under H_0: $X \sim \text{Geo}(\frac{1}{6})$

$P(X \geqslant 16) = P(X > 15) = (\frac{5}{6})^{15}$

$= 0.064905... > 0.05$

Not significant, so do not reject H_0.

There is insufficient evidence that the probability of rolling a 3 with Geethaka's die is less than $\frac{1}{6}$ (no evidence the die is biased).

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

Question 4 continued

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

DO NOT WRITE IN THIS AREA

Question 4 continued

(Total for Question 4 is 12 marks)

5. Some of the components produced by a factory are defective. The management requires that no more than 3% of the components produced are defective.
Niluki monitors the production process and takes a random sample of n components.

(a) Write down the hypotheses Niluki should use in a test to assess whether or not the proportion of defective components is greater than 0.03

(1)

Niluki defines the random variable D_n to represent the number of defective components in a sample of size n. She considers two tests **A** and **B**

In test **A**, Niluki uses $n = 100$ and if $D_{100} \geqslant 5$ she rejects H_0

(b) Find the size of test **A**

(2)

In test **B**, Niluki uses $n = 80$ and

- if $D_{80} \geqslant 5$ she rejects H_0
- if $D_{80} \leqslant 3$ she does not reject H_0
- if $D_{80} = 4$ she takes a second random sample of size 80 and if $D_{80} \geqslant 1$ in this second sample then she rejects H_0 otherwise she does not reject H_0

(c) Find the size of test **B**

(3)

Given that the actual proportion of defective components is 0.06

(d) (i) find the power of test **A**

(ii) find the expected number of components sampled using test **B**

(3)

Given also that, when the actual proportion of defective components is 0.06, the power of test **B** is 0.713

(e) suggest, giving your reasons, which test Niluki should use.

(1)

(a) $H_0: p = 0.03$ $H_1: p > 0.03$

(b) $D_{100} \sim B(100, 0.03)$

Size $= P(D_{100} \geq 5) = 1 - P(D_{100} \leq 4)$

$= 1 - 0.817855... = 0.182145... = 0.182$

(c) $D_{80} \sim B(80, 0.03)$

Size $= P(D_{80} \geq 5) + P(D_{80} = 4) \times P(D_{80} \geq 1)$

$= 0.092793... + 0.126540... \times 0.912554...$

$= 0.208267... = 0.208$

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
DO NOT WRITE IN THIS AREA

Question 5 continued

(d)(i) Test A with $p = 0.06$: $X \sim B(100, 0.06)$

Power $= P(X \geqslant 5) = 1 - P(X \leqslant 4) = 0.723224... = 0.723$

(ii) $Y \sim B(80, 0.06)$

A second sample is taken only when $Y = 4$

$P(Y = 4) = {}^{80}C_4 \times 0.06^4 \times 0.94^{76} = 0.185957...$

Expected number $= 80 + 80 \times P(Y = 4)$

$= 80 + 80 \times 0.185957...$

$= 94.876... \approx 95$ components

(e) Test A: size 0.182, power 0.723. Test B: size 0.208, power 0.713.

The tests have similar size and similar power, but test B samples fewer components on average (about 95 rather than 100), so Niluki should use test B.

Question 5 continued

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
DO NOT WRITE IN THIS AREA

Question 5 continued

(Total for Question 5 is 10 marks)

6. The random variable X has probability generating function $G_X(t)$ where

$$G_X(t) = \frac{1}{\sqrt{4-3t}}$$

(a) Use calculus to find Var(X)
Show your working clearly.

(6)

(b) Find the exact value of $P(X \leqslant 2)$

(4)

The independent random variables X_1 and X_2 each have the same distribution as X

The random variable $Y = X_1 + X_2 + 1$

(c) By finding the probability generating function of Y, state the name of the distribution of Y

(4)

(d) Hence, or otherwise, find $P(X_1 + X_2 > 5)$

(2)

(a) $G_X(t) = (4-3t)^{-1/2}$

$G'_X(t) = -\frac{1}{2}(4-3t)^{-3/2} \times (-3) = \frac{3}{2}(4-3t)^{-3/2}$

$E(X) = G'_X(1) = \frac{3}{2}(1)^{-3/2} = \frac{3}{2}$

$G''_X(t) = \frac{3}{2} \times (-\frac{3}{2})(4-3t)^{-5/2} \times (-3) = \frac{27}{4}(4-3t)^{-5/2}$

$G''_X(1) = \frac{27}{4}$

$Var(X) = G''_X(1) + G'_X(1) - [G'_X(1)]^2$

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

$Var(X) = 6$

(b) $P(X = 0) = G_X(0) = 4^{-1/2} = \frac{1}{2}$

$P(X = 1) = G'_X(0) = \frac{3}{2} \times 4^{-3/2} = \frac{3}{2} \times \frac{1}{8} = \frac{3}{16}$

$P(X = 2) = \frac{G''_X(0)}{2!} = \frac{1}{2} \times \frac{27}{4} \times 4^{-5/2} = \frac{1}{2} \times \frac{27}{4} \times \frac{1}{32} = \frac{27}{256}$

$P(X \leqslant 2) = \frac{1}{2} + \frac{3}{16} + \frac{27}{256} = \frac{128+48+27}{256}$

$P(X \leqslant 2) = \frac{203}{256}$

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
DO NOT WRITE IN THIS AREA

Question 6 continued

(c) $G_Y(t) = G_{X1}(t) \times G_{X2}(t) \times t$

$$= \frac{1}{\sqrt{(4-3t)}} \times \frac{1}{\sqrt{(4-3t)}} \times t = \frac{t}{4-3t}$$

Dividing numerator and denominator by 4:

$G_Y(t) = \frac{t/4}{1-(3/4)t} = \frac{pt}{1-(1-p)t}$ with $p = \frac{1}{4}$

This is the pgf of a geometric distribution, so $Y \sim Geo(\frac{1}{4})$

(d) $X_1 + X_2 > 5 \Leftrightarrow Y > 6$

$P(X_1 + X_2 > 5) = P(Y > 6) = (1 - \frac{1}{4})^6 = (\frac{3}{4})^6$

$$= \frac{729}{4096} = 0.177978\ldots = 0.178$$

Question 6 continued

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

DO NOT WRITE IN THIS AREA

Question 6 continued

(Total for Question 6 is 16 marks)

7. The probability of winning a prize when playing a single game of *Pento* is $\frac{1}{5}$

When more than one game is played the games are independent.
Sam plays 20 games.

(a) Find the probability that Sam wins 4 or more prizes.

(2)

Tessa plays a series of games.

(b) Find the probability that Tessa wins her 4th prize on her 20th game.

(2)

Rama invites Sam and Tessa to play some new games of *Pento*.
They must pay Rama £1 for each game they play but Rama will pay them £2 for the first time they win a prize, £4 for the second time and £($2w$) when they win their wth prize ($w > 2$)

Sam decides to play n games of *Pento* with Rama.

(c) Show that Sam's expected profit is £$\frac{1}{25}\left(n^2 - 16n\right)$

(6)

Given that Sam chose $n = 15$

(d) find the probability that Sam does not make a loss.

(4)

Tessa agrees to play *Pento* with Rama. She will play games until she wins r prizes and then she will stop.

(e) Find, in terms of r, Tessa's expected profit.

(4)

(a) X = number of prizes Sam wins, $X \sim B(20, 0.2)$

$P(X \geq 4) = 1 - P(X \leq 3) = 1 - 0.411449...$

$= 0.588551... = 0.589$

(b) Y = number of the game on which Tessa wins her 4th prize

$Y \sim$ Negative Binomial(4, 0.2)

$P(Y = 20) = {}^{19}C_3 \times 0.2^3 \times 0.8^{16} \times 0.2$

$= 969 \times 0.2^4 \times 0.8^{16}$

$= 0.043639... = 0.0436$

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
DO NOT WRITE IN THIS AREA

Question 7 continued

(c) S = number of prizes Sam wins in n games, $S \sim B(n, 0.2)$

Profit $= (2 + 4 + 6 + \ldots + 2S) - n$

$= \frac{S}{2}[2 \times 2 + 2(S - 1)] - n$ (arithmetic series, $a = d = 2$)

$= S(S + 1) - n = S^2 + S - n$

$E(S) = 0.2n$ $\quad$ $Var(S) = n \times 0.2 \times 0.8 = 0.16n$

$E(S^2) = Var(S) + [E(S)]^2 = 0.16n + 0.04n^2$

$E(\text{Profit}) = E(S^2) + E(S) - n$

$= 0.04n^2 + 0.16n + 0.2n - n$

$= 0.04n^2 - 0.64n$

$= \frac{1}{25}(n^2 - 16n)$

So Sam's expected profit is $£\frac{1}{25}(n^2 - 16n)$ as required

(d) $n = 15$, $S \sim B(15, 0.2)$. Sam does not make a loss when

$S^2 + S - 15 \geqslant 0$

$S^2 + S - 15 = 0 \Rightarrow S = \frac{-1 \pm \sqrt{61}}{2} \Rightarrow S = 3.405\ldots$ (or $-4.405\ldots$)

So $S \geqslant 3.405\ldots$, i.e. $S \geqslant 4$ ($S = 3$: $12 - 15 < 0$; $S = 4$: $20 - 15 > 0$)

$P(\text{no loss}) = P(S \geqslant 4) = 1 - P(S \leqslant 3)$

$= 1 - 0.648162\ldots$

$= 0.351838\ldots = 0.352$

(e) T = number of the game on which Tessa wins her rth prize

$T \sim$ Negative Binomial$(r, 0.2)$

Profit $= (2 + 4 + \ldots + 2r) - T = r(r + 1) - T$

$E(T) = \frac{r}{p} = \frac{r}{0.2} = 5r$

Expected profit $= r(r + 1) - E(T) = r^2 + r - 5r$

Tessa's expected profit $= £(r^2 - 4r)$

Question 7 continued

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
DO NOT WRITE IN THIS AREA

Question 7 continued

Question 7 continued

(Total for Question 7 is 18 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