Please check the examination details below before entering your candidate information

Candidate surname: MODEL SOLUTIONS

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

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) Recall that $\text{Var}(X) = E[X^2] - (E[X])^2$

$E[X] = 0.2(-1) + 0.1(0) + 0.2(1) + 0.25(3) + 0.25(5) = 2$ ①

$E[X^2] = 0.2(-1)^2 + 0.1(0)^2 + 0.2(1)^2 + 0.25(3)^2 + 0.25(5)^2 = 8.9$ ①

$\Rightarrow \text{Var}(X) = 8.9 - 2^2 = 4.9$ ①

b) As $\text{Var}(X) = E[X^2] - (E[X])^2$, $\text{Var}(X^2) = E[X^4] - (E[X^2])^2$

$E[X^4] = 0.2(-1)^4 + 0.1(0)^4 + 0.2(1)^4 + 0.25(3)^4 + 0.25(5)^4 = 176.9$ ①

$\Rightarrow \text{Var}(X^2) = 176.9 - 8.9^2 = 97.7$ ① ①

DO NOT WRITE IN THIS AREA

P 7 2 7 9 7 A 0 2 2 4

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) $X \sim Po(2.4)$ (errors per 100 words)

i) $Pr(X=3) = 0.2040$ 1

ii) $Pr(X<2) = 0.3084$ 1 Using a calculator.

b) $H_0: \lambda = 2.4$ $\quad 2.4 \times \frac{250}{100} = 6$

$H_1: \lambda \neq 2.4$ 1

Let $E \sim Po(6)$ (errors per 250 words)

$Pr(E \leqslant 1) = 0.0174 < 0.025$ as two-tailed.

$Pr(E \leqslant 2) = 0.0620 > 0.025$ 1

$Pr(E \geqslant 11) = 0.0426 > 0.025$

$Pr(E \geqslant 12) = 0.0201 < 0.025$ 1

So $E \leqslant 1$ or $E \geqslant 12$ is the critical region 1

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

Question 2 continued

c) Recall that a Type I error is rejecting H_0 when H_0 is true.

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

$$= 0.0174 + 0.0201 = 0.0375 \quad (1)$$

(Total for Question 2 is 7 marks)

Turn over

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 the colour chosen and year group

H_1: There is an association between colour and year group. 1

b)i) Choose the cell with the lowest row total and column total.

'Black', '10–12'. 1

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

P 7 2 7 9 7 A 0 6 2 4

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

Question 3 continued

ii) Recall that the expected frequency is

$$\frac{\text{row total} \times \text{column total}}{\text{total}}$$

$$E = \frac{68 \times 19}{240} = 5.3833 \quad (1)$$

c) Recall that for a table with r rows and c columns, we compare our test statistic with th χ^2 distribution with $(r-1)(c-1)$ degrees of freedom.

$v = (5-1)(3-1) = 8 \quad (1)$

From the table, we have a critical value of

$20.090 \quad (1)$

$38.449 > 20.090.$

Hence, we reject H_0 as there is evidence of an association between the colour chosen and the year group. (1)

(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 $P(X \leqslant 3) = \frac{91}{216}$

(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 $\frac{1}{6}$

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) Geethaka repeats an event until sucessful.
This is represented by a geometric distribution.

$X \sim \text{Geom}(1/6)$ ①

b) Recall that $\Pr(X=x) = (1-p)^{x-1}p$

$\Pr(X \leq 3) = \Pr(X=1) + \Pr(X=2) + \Pr(X=3)$

$= \left(\frac{5}{6}\right)^0 \times \left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)^1 \times \left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)^2 \times \left(\frac{1}{6}\right)$ ①

$= \frac{1}{6} + \frac{5}{36} + \frac{25}{216}$

$= \frac{36}{216} + \frac{30}{216} + \frac{25}{216} = \frac{91}{216}$ ①

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

c) Recall that for a Geometric distribution,

$E[X] = \frac{1}{p}$ and $Var(X) = \frac{1-p}{p^2}$

$E[X] = \frac{1}{1/6} = 6$ 1

$Var(X) = \frac{1-\frac{1}{6}}{(1/6)^2} = 30$ 1

$n = 64$

So by the Central Limit Theorem,

$\bar{X} \sim N(6, 30/64)$ 1 1

$Pr(5.6 \leq \bar{X} \leq 7.2) = 0.6806$ 1

d) $H_0 : p = 1/6$

$H_1 : p < 1/6$ 1

Using a Calculator

$Pr(X \geq 16)$ 1 $= 0.0649 > 0.5$ 1

Hence, we do not reject H_0 as there is insufficient evidence to conclude that the die is biased.

1

P 7 2 7 9 7 A 0 9 2 4

Turn over

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) $H_0: p = 0.03$

$H_1: p > 0.03$ 1

b) Recall that the size of a test is the probability of making a Type I error.

$D_{100} \sim B(100, 0.03)$
1
$\Pr(D_{100} \geqslant 5) = 0.1821$ 1 using a calculator

P 7 2 7 9 7 A 0 1 2 2 4

DO NOT WRITE IN THIS AREA

Question 5 continued

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

$Pr(A \cap B) = Pr(A)\,Pr(B)$ because the events are independent

$Pr(D_{80} \geq 5) + \left(Pr(D_{80} = 4)\right)\left(Pr(D_{80} \geq 1)\right)$ ①

$= 0.0928 + 0.12654 \times 0.9126$ ① Using a calculator

$= 0.208$ ①

d) i) Recall that the power of a test is the probability of not making a Type II error.

$X \sim Bin(100, 0.06)$

$Pr(X) \geq 5 = 0.7232$ ①

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

$Pr(Y = 4) = 0.186 \leftarrow$ Using a calculator

So the expected number is

$80 + 80(0.186) = 94.85$ ①

So $n = 95$ ①

e) The tests have similar size and power but Test B involves sampling fewer components so using Test B would be advised. ①

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

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

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

$\Rightarrow G''_X(t) = \frac{27}{4}(4-3t)^{-5/2}$ By Chain Rule.

$G'_X(1) = 3/2$, $[G'_X(1)]^2 = 9/4$, $G''_X(1) = \frac{27}{4}$

$\Rightarrow \text{Var}(X) = \frac{27}{4} - \frac{3}{2} - \frac{9}{4} = 6$

b) We can use a binomial expansion to find the up to the t^2 coefficient.

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

$= \frac{1}{2}(1 - 3/4\,t)^{-1/2}$ 1 as $4^{-1/2} = \frac{1}{2}$

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

P 7 2 7 9 7 A 0 1 6 2 4

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

Question 6 continued

$$= \frac{1}{2}\left(1 + (-\tfrac{1}{2})(-\tfrac{3}{4}t) + \frac{(-\tfrac{1}{2})(-\tfrac{3}{2})}{2!}(-\tfrac{3}{4}t)^2 + \ldots\right)$$

$$= \frac{1}{2}\left(1 + \frac{3}{8}t + \frac{27}{128}t^2 + \ldots\right)$$ ① ①

$$= \frac{1}{2} + \frac{3}{16}t + \frac{27}{256}t^2 + \frac{1}{2}(\ldots)$$

$$Pr(x \leq 2) = Pr(x=0) + Pr(x=1) + Pr(x=2)$$

$$= \frac{1}{2} + \frac{3}{16} + \frac{27}{256}$$

$$= \frac{203}{256}$$ ①

c) As X_1 and X_2 are independent,

$$G_Y(t) = G_X(t) \times G_X(t) \times t$$

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

Look at the P.G.F section in the formula booklet. This is in the form

$$\frac{pt}{1-(1-p)t} \sim \text{Geom}(p)$$

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

So $Y \sim \text{Geom}(1/4)$ ①

Question 6 continued

d) $\Pr(X_1 + X_2 > 5)$ Note that $Y - 1 = X_1 + Y_2$

$= \Pr(Y - 1 > 5)$ 1

$= \Pr(Y > 6) = 0.178$ 1 using a calculator

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) $X \sim Bin(20, 0.2)$ ① X number of prizes Sam wins

$Pr(X \geq 4) = 0.589$ ① using a calculator.

b) $Y \sim NegBin(4, 0.2)$ ① Y number of games when Tessa wins her 4th prize.

$Pr(Y=20) = \binom{19}{3} 0.2^3 \times 0.8^{16} \times 0.2 = 0.0436$ ①

c) $S \sim Bin(n, 0.2)$ ① S is the number of prizes Sam wins in n games.

Question 7 continued

The total cost is £n as he pays £1 per game.

The profit is $(2+4+\dots+2S)-n$ ①

$= 2(1+2+\dots+S) - n$

$= 2\left(\frac{S(S+1)}{2}\right) - n$ as $\sum_{k=r}^{S} r = \frac{S(S+1)}{2}$ in the Formula Booklet

$= S^2 + S - n$ ①

From the Formula Booklet, we know that if $X \sim Bin(n,p)$,

$E[X]=np$, $Var(X) = np(1-p)$

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

$E[S] = 0.2n$, $Var(S) = 0.2 \times 0.8n = 0.16n$ ①

$\Rightarrow 0.16n = E[S^2] - (0.2n)^2$

$\Rightarrow E[S^2] = 0.04n^2 + 0.16n$ ①

$\Rightarrow E[S^2+S-n] = 0.04n^2 + 0.16n + 0.2n - n$

$= £\frac{1}{25}(n^2 - 16n)$ ①

Question 7 continued

d) If Sam does not make a loss,

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

$\Pr(S^2 + S - n \geqslant 0)$

$= \Pr(S^2 + S - 15 \geqslant 0)$ ①

$S^2 + S - 15 = 0$

has solutions $3.4051, -4.405$ ① using a calculator

As S is a positive integer, we have

$S \geqslant 4$ ①

$\Pr(S \geqslant 4) = 0.3518$ ①

e) $T \sim \text{NegBin}(r, 0.2)$ ① T is the game where Tessa wins her r^{th} prize

Profit $= (2 + 4 + \dots + 2r) - T = r(r+1) - T$ ① – similar to b

$E[r(r+1) - T] = r(r+1) - E[T]$ ①

Recall that the expectation of a negative binomial is $\frac{r}{p}$

$= r(r+1) - \frac{r}{0.2} = r^2 - 4r$ ①

P 7 2 7 9 7 A 0 2 2 2 4