Please check the examination details below before entering your candidate information

Candidate surname	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

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 $\text{Var}(X)$

(3)

(b) Find $\text{Var}(X^2)$

(3)

Important Formulae

$E(X) = \sum x P(X=x)$

$E(X^2) = \sum x^2 P(X=x)$

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

(a) Let's get $E(X)$:

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

$= \cancel{-0.2} + \cancel{0.2} + 0.75 + 1.25$

$E(X) = 2$ M1

Let's get $E(X^2)$:

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

$= 0.2 + 0.2 + 9\times0.25 + 25\times0.25$

$= 0.4 + 34\times0.25$

$= 0.4 + 8.5$

$E(X^2) = 8.9$ M1

Substitute into formula:

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

$= 8.9 - (2)^2$

$= 8.9 - 4 = 4.9$

$Var(X) = 4.9$ A1

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

Question 1 continued

(b) We will use the same logic as above:

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

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

let's get $E(X^4)$:

$E(X^4) = (-1)^4(0.2) + \cancel{0^4(0.1)}^{0} + 1^4 \times (0.2) + 3^4(0.25) + 5^4(0.25)$

$= 0.2 + 0.2 + 81 \times 0.25 + 625 \times 0.25$

$E(X^4) = 176.9$ M1

We have $E(X^2)$ from (a):

$E(X^2) = 8.9$

Substitute:

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

$= 176.9 - (8.9)^2$ M1

$= 97.69 \rightarrow 97.7$

$\therefore$ $Var(X^2) = 97.7$ to 3sf A1

(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. average rate

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. two tailed

(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) $X \to$ # of errors in the 100-word piece   define variable

$X \sim Po(2.4)$ λ was given

i. $P(X=3) = 0.20901 \longrightarrow 0.209$ to 3sf   B1

= → "exactly"

ii. $P(X \leq 2) = 0.30844 \longrightarrow 0.308$ to 3sf   B1

fewer than

Question 2 continued

(b) Hypotheses

$H_0: \lambda = 2.4$ # of errors in 100-words. to get the corresponding rate for 250-words, multiply by 2.5.

$H_1: \lambda \neq 2.4$ two-tailed test.

2.4×2.5=6

$H_0: \mu = 6$

$H_1: \mu \neq 6$ B1

E → # of errors per 250-word piece define variable

$E \sim Po(6)$ M1

5% significance ⟶ 0.025 in each tail

Lower tail

$P(E \leq c_1) < 0.025$ critical value

$P(E \leq 1) = 0.0174 < 0.025 \therefore$ falls in crit region

$P(E \leq 2) = 0.0620 > 0.025 \therefore$ doesn't fall in crit. region

substitute values to get the first E value that falls in the critical region

$\therefore$ lower tail of critical region:

$E \leq 1$ c.v. is 1

Upper tail

$P(E \geq c_2) = 0.025$

$1 - P(E \leq c_2 - 1) = 0.025$ subtract 1 so that c_2 is included

$P(E \leq c_2 - 1) = 0.975$

$P(E \leq 10) = 0.957 < 0.975 \therefore$ not in crit. region

$P(E \leq 11) = 0.980 > 0.975 \therefore$ in crit. region

$\therefore c_2 - 1 = 11 \quad c_2 = 12$ critical value

$\therefore$ Upper tail of critical region:

$E \geq 12$ M1

Full critical Region:

$E \leq 1 \cup E \geq 12$ A1

(c) Type I Error is the probability of falsefully rejecting H_0. (H_0 is true but we reject it)

It is just the actual significance of the test.

actual significance $= P(E \leq 1) + P(E \geq 12)$

$= 0.0174 + (1 - 0.979908)$

$= 0.0174 + 0.0201$

$= 0.0375$

P(Type I Error) = 0.0375 B1

(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) Hypotheses

H_0: There is no association between favorite color and yeargroup

H_1: There is an association between favorite color and yeargroup B1

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 3 continued

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

That is Black, 10-12 B1

ii. For contingency tables we calculate the expected frequency like this:

			Total
	a	b	a+b
	c	d	c+d
Total	a+c	b+d	a+b+c+d

row total

column total

grand total

to get the E_i of this, we calculate

$$E_i = \frac{\text{row total} \times \text{column total}}{\text{grand total}}$$

$\therefore$ here it would be $E_i = \frac{(a+b)(a+c)}{a+b+c+d}$

Substitute: $E_i = \frac{68 \times 19}{240} = 5.3833$

5.38 to 3sf expected frequency B1

(c) Calculate degrees of freedom

★ For contingency tables the formula for DoF is:

DoF = (#of rows - 1)(#of columns - 1)

$\nu = (5-1) \times (3-1) = 8$ degrees of freedom B1

critical value from tables: $\chi^2_8 (1\%) = 20.090$ c.v. B1

his $\chi^2 = 38.449 > 20.090$ $\therefore$ falls in critical region $\therefore$ reject H_0

There is significant evidence of an association between between color chosen and yeargroup B1

(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 N 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) Geometric Distribution since we are modelling the number of trials until the first success

$X \sim Geo(\frac{1}{6})$ B1

(b) $P(X \leq 3) = P(X=1) + P(X=2) + P(X=3)$ OR $P(X \leq 3) = 1 - P(X > 3)$ no more than 3 fails

less than 3 tries $=$ success + (fail)(success) + (fail)(fail)(success) $= 1 - (\frac{5}{6})^3$ M1

$= \frac{1}{6} + \frac{1}{6} \times \frac{5}{6} + (\frac{5}{6})^2 \times \frac{1}{6}$ M1 $= 1 - \frac{125}{216}$

$= \frac{91}{216}$ A1 $= \frac{91}{216}$ A1

(c) From Formula Booklet for $X \sim Geo(p)$

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

Substitute:

$(\mu)\ E(X) = 1 - \frac{1}{6} = \frac{5}{6}$ $(\sigma^2)\ Var(X) = \frac{\frac{5}{6}}{(\frac{1}{6})^2} = 30$ M1A1

To use the central limit theorem let's convert to a normal distribution:

$X \sim Geo(\frac{1}{6}) \longrightarrow \bar{X} \approx \sim N(\mu, \sqrt{\frac{\sigma^2}{N}})$

Substitute: $\bar{X} \approx \sim N(6, \sqrt{\frac{30}{64}})$ M1A1

$P(5.6 < \bar{X} < 7.2) = 0.68064 \longrightarrow 0.681$ to 3sf A1

Question 4 continued

(d) Hypotheses

$H_0: p = \frac{1}{6}$

$H_1: p < \frac{1}{6}$ B1

We are given the found value 16. We can test wether this falls in the critical region.

M1A1

$P(X \geq 16) = P(X > 15) = (1 - \frac{1}{6})^{15} = (\frac{5}{6})^{15} = 0.0649 > 0.05$ $\therefore$ 16 does not fall in the critical region.

insufficient evidence to reject H_0.

insufficient evidence that $p < \frac{1}{6}$ A1

this is a geometric hypothesis test, so the value falls in the critical region when more trials are needed. The more trials it takes, the more likely it is that $p < \frac{1}{6}$.

calculate the prob. of 15 failures $\therefore$ first success in 16^{th} trial or later

Question 4 continued

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) Hypotheses

$H_0: p = 0.03$

$H_1: p > 0.03$ B1

(b) $D_{100} \sim B(100, 0.03)$ we use Binomial as we have constant probability and fixed number of trials

Size is the probability of rightfully rejecting H_0 ∴ P(falling in critical region)

$\underbrace{P(D_{100} \geqslant 5)}_{\text{given crit. region}} = 1 - P(D_{100} \leqslant 1)$ M1

$= 0.18214 \rightarrow 0.182$ to 3sf size A1

Question 5 continued

(c) Size is the probability of rightfully rejecting H_0 $\therefore$ P(falling in critical region)

The critical regions are:

$\rightarrow D_{80} \geq 5$

$\rightarrow D_{80} = 4$ and then $D_{80} \geq 1$ "and then" → multiply

$\therefore$ Size: $P(D_{80} \geq 5) + P(D_{80} = 4) \times P(D_{80} \geq 1)$ M1

$= 0.09279 + (0.12654)(0.91255)$ A1

$= 0.20826 \longrightarrow$ 0.208 to 3sf size A1

(d) i. Power is the probability of rightfully rejecting H_0 using the actual probability $\therefore$ P(in crit. region | p = actual)

For test A:

$X \sim B(100, 0.06)$

$P(X \geq 5) = 0.72322 \longrightarrow$ 0.723 to 3sf power B1

ii. For test B:

$Y \sim B(80, 0.06)$

80 components are sampled just for the test.

Every time $Y = 4$, a second component is sampled.

$\therefore$ total $= 80 + 80P(Y = 4)$ M1

$= 94.87 \longrightarrow$ 95 components A1

(e) The size and power are similar, but as test B uses fewer components I would suggest the use of test B. B1

Question 5 continued

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

Let's differentiate to get $G'_X(t)$ and $G''_X(t)$

$G_X(t) = (4-3t)^{-\frac{1}{2}}$ ✱ chain rule: multiply by the derivative of the bracket

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

derivative of bracket

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

$t=1$: $G'_X(1) = \frac{3}{2}(1)^{-\frac{3}{2}} = \frac{3}{2}$ M1A1

$G''_X(t) = -\frac{3}{2} \times \frac{3}{2}(4-3t)^{-\frac{5}{2}} \times -3$ ✱ use chain rule again

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

$G''_X(1) = \frac{27}{4}(1)^{-\frac{5}{2}} = \frac{27}{4}$ M1A1

Substitute:

$Var(X) = \frac{27}{4} + \frac{3}{2} - \left(\frac{3}{2}\right)^2$

$= 6$

$\therefore Var(X) = 6$ M1A1

Question 6 continued

(b) **Method 1** – Maclaurin

$P(X\le 2) = P(X=0) + P(X=1) + P(X=2)$

General Maclaurin Expansion

$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \dots$

$P(X=2) = \frac{G''_X(0)}{2!}x^2 = \frac{1}{2} \times \frac{27}{4} \times \frac{1}{32} = \frac{27}{256}$ M1A1

$P(X=1) = \frac{3}{2} \times \frac{1}{8} = \frac{3}{16}$

$P(X=0) = \frac{1}{2}$ M1

$\therefore P(X \le 2) = \frac{1}{2} + \frac{3}{16} + \frac{27}{256}$

$= \frac{203}{256}$ A1

Method 2 – Binomial

let's expand to get the first 3 terms:

$G_X(t) = \frac{1}{2}\left(1 - \frac{3}{4}t\right)^{-\frac{1}{2}}$ get it into the required $a(1-bx)$ form

$= \frac{1}{2}\left(1 + \left(-\frac{1}{2}\right)\left(-\frac{3}{4}\right)t + \frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)}{2}\left(-\frac{3}{4}t\right)^2 + \dots\right)$ M1

$= \frac{1}{2}\left(1 + \frac{3}{8}t + \frac{3}{8} \times \frac{9}{16}t^2\right)$ M1

$= \frac{1}{2} + \frac{3}{16}t + \frac{27}{256}t^2$ A1

substitute $t = 1$

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

$P(X \le 2) = \frac{203}{256}$ A1

(c) When $W = X + Y$, $G_W(t) = G_X(t) \times G_Y(t)$

$\therefore$ in this case:

$Y = X_1 + X_2 + 1 \longrightarrow G_Y(t) = G_{X_1}(t) \times G_{X_2}(t) \times t$

$= \frac{1}{\sqrt{4-3t}} \times \frac{1}{\sqrt{4-3t}} \times t$ M1

$G_Y(t) = \frac{t}{4-3t}$ A1

look at the form:

$G_Y(t) = \frac{\frac{1}{4}t}{1 - \frac{3}{4}t}$ div. up and down by 4 M1

the form looks like Geometric

$\therefore$ distribution is $Y \sim \text{Geo}\left(\frac{1}{4}\right)$ A1

(d) $P(X_1 + X_2 > 5)$

$Y = X_1 + X_2 + 1$

$Y - 1 = X_1 + X_2$

we want six failures

$P(Y - 1 > 5) \longrightarrow P(Y > 6) = \left(1 - \frac{1}{4}\right)^6 = \frac{729}{4096}$

$= 0.177978$

$= 0.178$ M1A1

Question 6 continued

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) $X \to$ # of Prizes Sam wins   define variable

$X \sim B(20, 0.2)$ binomial as fixed # of trials + constant probability

$P(X \geq 4) = 1 - P(X \leq 3)$ M1

$= 0.588555 \longrightarrow 0.589$ to 3sf   A1

(b) $Y \to$ # of game when Tessa wins 4^{th} prize   define variable

$Y \sim negB(4, 0.2)$ M1   Formula for $X \sim NegB(n,p)$

$P(Y=20) = \binom{19}{3} \times 0.2^3 \times 0.8^{16}$ $P(X=x) = \binom{x-1}{n-1}(p)^{n-1}(1-p)^{x-n}$

$= 0.043639$

$= 0.0436$ to 3sf   A1

Question 7 continued

(c) $S \rightarrow$ # of Prizes Sam wins in n games. define variable

$S \sim B(n, 0.2)$ M1

Profit would be an arithmetic sequence.

Profit $= (2+4+6 \ldots + 2S) - n$ M1

$\therefore$ it's sum of an arithmetic sequence $- n$

$\frac{S}{2}[4+2(S-1)] - n$

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

profit $= S^2 + S - n$ A1

Formulae for exp. value and variance

$E(X) = np \quad Var(X) = np(1-p)$

$E(S) = 0.2n$, $Var(S) = 0.2 \times 0.8n = 0.16n$ M1

we know that $Var(x) = E(x^2) - [E(x)]^2$

$\therefore Var(x) + [E(x)]^2 = E(x^2)$

$E(S^2) = 0.16n + (0.2n)^2$

$E(S^2) = 0.16n + 0.04n^2$ A1

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

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

$= 0.04n^2 + 0.36n - n$

$= \frac{1}{25}n^2 + \frac{9}{25}n - n$

$= \frac{1}{25}(n^2 + 9n - 25n)$

$= \frac{1}{25}(n^2 - 16n)$ hence shown A1

(d) We want profit ≥ 0 with $n = 15$:

$P(S^2 + S - 15 \geq 0)$ M1

Solve to get first game w/ profit.

use quadratic formula

$\frac{-1 \pm \sqrt{1^2 - 4(1)(-15)}}{2} = 3.40, \cancel{-4.41}$ Reject M1

$\therefore$ from 4 games onwards $\rightarrow$ profit

$S \sim B(15, 0.2)$

$P(S \geq 4) = 0.35183 \rightarrow 0.352$ to 3sf M1A1

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

Question 7 continued

(e) T → game on which Tessa wins r^{th} prize define variable

$T \sim NegB(r, 0.2)$ M1

Profit $= (2+4+6+\ldots+2r) - T$

$= r(r+1) - T$ A1

$E(X) = \frac{n}{p}$

Her expected profit $= r(r+1) - E(T)$ M1

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

$= r^2 + r - 5r$

$= r^2 - 4r$ A1

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