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 13 June 2025

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 ▶

P76380A

©2025 Pearson Education Ltd.
Y:1/1/1/

P 7 6 3 8 0 A 0 1 2 4

Pearson

1. Irina is practising her serves in badminton and counts the number of her serves that are faults. She finds that 15% of her serves are faults.

Assuming that each serve is independent,

(a) find the probability that

(i) Irina's 4th fault comes on her 20th serve,

(2)

(ii) in 18 serves, Irina has 4 faults.

(2)

With practice, Irina reduces her proportion of faults, p, so that the mean number of serves until her 4th fault is at least 32

(b) Find the maximum value of p

(3)

(a)(i) X = number of serves up to and including the 4th fault

$X \sim NB(4, 0.15)$

$P(X = 20) = {}^{19}C_3 \times (0.15)^4 \times (1 - 0.15)^{16}$

$= 969 \times 0.15^4 \times 0.85^{16}$

$= 0.036424... = 0.0364$

(a)(ii) Y = number of faults in 18 serves, $Y \sim B(18, 0.15)$

$P(Y = 4) = {}^{18}C_4 \times (0.15)^4 \times (0.85)^{14}$

$= 3060 \times 0.15^4 \times 0.85^{14}$

$= 0.159203... = 0.159$

(b) Number of serves until 4th fault $\sim NB(4, p)$

Mean $= \frac{4}{p}$

$\frac{4}{p} \geq 32$

$\Rightarrow 32p \leq 4 \Rightarrow p \leq \frac{1}{8}$

Maximum value of $p = \frac{1}{8} = 0.125$

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 7 marks)

2. The discrete random variable X has probability distribution

x	-1	a	b
$P(X = x)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$

where a and b are positive constants.

(a) Find an expression for $E(X)$ in terms of a and b

(2)

The discrete random variable Y is defined as $Y = a + bX$

Given that $\text{Var}(Y) = \frac{1}{4}\text{Var}(X)$ and $E(Y) = \frac{5}{16}$

(b) find the value of $E(X)$

(7)

(a) $E(X) = -1 \times \frac{1}{2} + a \times \frac{1}{4} + b \times \frac{1}{4}$

$E(X) = \frac{1}{4}a + \frac{1}{4}b - \frac{1}{2}$

(b) $\text{Var}(Y) = \text{Var}(a + bX) = b^2\,\text{Var}(X)$

$b^2\,\text{Var}(X) = \frac{1}{4}\,\text{Var}(X) \Rightarrow b^2 = \frac{1}{4} \Rightarrow b = \frac{1}{2} \quad (b > 0)$

$E(Y) = a + bE(X)$

$a + \frac{1}{2}\left(\frac{1}{4}a + \frac{1}{4} \times \frac{1}{2} - \frac{1}{2}\right) = \frac{5}{16}$

$a + \frac{1}{8}a + \frac{1}{16} - \frac{1}{4} = \frac{5}{16}$

$\frac{9}{8}a = \frac{5}{16} + \frac{3}{16} = \frac{1}{2} \Rightarrow a = \frac{4}{9}$

$E(X) = \frac{1}{4}\left(\frac{4}{9}\right) + \frac{1}{4}\left(\frac{1}{2}\right) - \frac{1}{2}$

$= \frac{1}{9} + \frac{1}{8} - \frac{1}{2} = \frac{8 + 9 - 36}{72}$

$E(X) = -\frac{19}{72}$

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 9 marks)

3. A journalist uses a gender equality test on some randomly chosen films.
She believes that whether or not a film passes the test is associated with the period in which the film is first released. Her data is summarised in Table 1.
The journalist decides to test whether the data supports her belief.

Table 1: Observed frequencies

Test result \ Period	1990–1999	2000–2004	2005–2013	Total
Pass	45	70	65	180
Fail	75	90	55	220
Total	120	160	120	400

Some of the expected frequencies she calculates are shown in Table 2 on page 7.

(a) Complete Table 2 showing expected frequencies.

(2)

(b) Write down hypotheses for a suitable test to assess the journalist's belief.

(1)

(c) Showing your working clearly, test the journalist's belief using a 5% level of significance.
You should state your critical value and conclusion clearly.

(6)

The journalist wants to publish her data and findings.
She decides to publish the data as **percentages** of pass and fail for each period, as shown in Table 3, rather than as **frequencies**.

Table 3: Percentages

Test result \ Period	1990–1999	2000–2004	2005–2013
Pass (%)	37.5	43.75	54.17
Fail (%)	62.5	56.25	45.83
Total (%)	100	100	100

The editor suggests carrying out the hypothesis test of the journalist's belief using the percentages in Table 3.

(d) (i) Describe the effect this would have on the test statistic, justifying your answer.

(ii) Explain whether or not the journalist should follow her editor's suggestion.

(3)

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

Table 2: Expected frequencies

Test result \ Period	1990–1999	2000–2004	2005–2013
Pass	54	72	54
Fail	66	88	66

(a) $E = \frac{\text{row total} \times \text{column total}}{\text{grand total}}$

Pass, 1990–1999: $\frac{180 \times 120}{400} = 54$ Pass, 2000–2004: $\frac{180 \times 160}{400} = 72$

Fail, 1990–1999: $\frac{220 \times 120}{400} = 66$ Fail, 2000–2004: $\frac{220 \times 160}{400} = 88$

(b) H_0: whether a film passes the test is not associated with the period in which the film is first released

H_1: whether a film passes the test is associated with the period in which the film is first released

(c) $\frac{(O-E)^2}{E}$:

Pass: $\frac{(45-54)^2}{54} = 1.5$ $\frac{(70-72)^2}{72} = 0.0556$ $\frac{(65-54)^2}{54} = 2.2407$

Fail: $\frac{(75-66)^2}{66} = 1.2273$ $\frac{(90-88)^2}{88} = 0.0455$ $\frac{(55-66)^2}{66} = 1.8333$

$\chi^2 = \Sigma\frac{(O-E)^2}{E} = 1.5 + 0.0556 + 2.2407 + 1.2273 + 0.0455 + 1.8333$

$\chi^2 = 6.9023\ldots = 6.90$

$\nu = (3-1)(2-1) = 2$

Critical value $\chi^2_2(5\%) = 5.991$

$6.90 > 5.991$, so the result is significant: reject H_0.

There is sufficient evidence to support the journalist's belief that whether a film passes the gender equality test is associated with the period in which the film is first released.

Question 3 continued

(d)(i) The observed totals for each period (120, 160, 120) are all more than 100, so converting to percentages makes every O_i and E_i smaller, and every difference $O_i - E_i$ smaller.
e.g. 1990–1999 pass: $O = 37.5$, $E = 45$ (instead of 45 and 54)
Therefore the test statistic would decrease
(recalculating with the percentages gives $\chi^2 = 5.73...$).

(d)(ii) No, the journalist should not follow her editor's suggestion.
The test must use the real observed frequencies, not percentages:
the percentages give $\chi^2 = 5.73 < 5.991$, so using them would change the result of the test (H_0 would not be rejected).

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 12 marks)

4. The discrete random variable X has probability generating function

$$G_X(t) = \frac{1}{\left(1 - \frac{2t}{5}\right)} - k$$

where k is a constant.

(a) Find the exact value of k

(2)

(b) Find the exact value of $E(X)$

(3)

(c) Find $P(X = 2)$

(3)

The random variable Y is defined as $Y = 4X - 1$

(d) Find the probability generating function of Y

(2)

(a) $G_X(1) = 1$: $\frac{1}{1 - 2/5} - k = 1 \Rightarrow \frac{5}{3} - k = 1 \Rightarrow k = \frac{2}{3}$

(b) $G_X(t) = (1 - \frac{2t}{5})^{-1} - k$

$G'_X(t) = -1 \times (-\frac{2}{5})(1 - \frac{2t}{5})^{-2} = \frac{2}{5}(1 - \frac{2t}{5})^{-2}$

$E(X) = G'_X(1) = \frac{2}{5} \times (\frac{3}{5})^{-2} = \frac{2}{5} \times \frac{25}{9} = \frac{10}{9}$

(c) $G''_X(t) = (-2) \times \frac{2}{5} \times (-\frac{2}{5})(1 - \frac{2t}{5})^{-3} = \frac{8}{25}(1 - \frac{2t}{5})^{-3}$

$P(X = 2) = \frac{G''_X(0)}{2!} = \frac{1}{2} \times \frac{8}{25}$

$P(X = 2) = \frac{4}{25}$

(d) $G_Y(t) = E(t^{4X-1}) = t^{-1} E((t^4)^X) = \frac{1}{t} G_X(t^4)$

$G_Y(t) = \frac{1}{t}\left(\frac{1}{1 - 2t^4/5} - \frac{2}{3}\right)$

$= \frac{5}{t(5 - 2t^4)} - \frac{2}{3t}$

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

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 10 marks)

5. A football team scores goals at an average rate of 1.8 goals per match.

(a) Give two assumptions that would be necessary to use a Poisson distribution to model the number of goals scored in a match by the team.

(2)

Given that in their next match the team scores exactly 3 goals,

(b) find the exact probability that 2 of these goals were scored in the first half of the match.

(4)

A hockey team plays 2 games each week during a 40-week season.
Each game consists of 2 periods.
The probability that the team concedes no goals in a period is 0.55

The random variable X represents the number of **periods** in a week in which the team concedes no goals.

(c) (i) Write down a suitable distribution for X

(ii) For this 40-week season, use the Central Limit Theorem to estimate $P(\overline{X} > 2)$

(4)

(a) Goals are scored independently of each other.

Goals are scored at a constant rate (average 1.8 per match).

(b) A_1 = goals in a match ~ Po(1.8); goals in each half ~ Po(0.9)

$P(2 \text{ in 1st half} \mid 3 \text{ in match}) = \frac{P(2 \text{ in 1st half}) \times P(1 \text{ in 2nd half})}{P(3 \text{ in match})}$

$= \frac{0.9^2e^{-0.9}/2! \times 0.9e^{-0.9}/1!}{1.8^3e^{-1.8}/3!} = \frac{0.164660... \times 0.365912...}{0.160670...}$

$= \frac{0.9^3/2}{1.8^3/16} = 3 \times (\frac{1}{2})^3 = \frac{3}{8}$

(c)(i) $X \sim B(4, 0.55)$

(c)(ii) $E(X) = 4 \times 0.55 = 2.2$ $\quad Var(X) = 4 \times 0.55 \times 0.45 = 0.99$

40 weeks, so by the Central Limit Theorem

$\overline{X} \approx N(2.2, \frac{0.99}{40})$, i.e. $\overline{X} \approx N(2.2, 0.02475)$

$P(\overline{X} > 2) = P(Z > \frac{2 - 2.2}{\sqrt{0.02475}}) = P(Z > -1.27128...)$

$= 0.898186... = 0.898$

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

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. A biased coin and a fair coin are thrown repeatedly.

The probability that the biased coin shows heads is $\frac{1}{3}$

Let B represent the number of throws of the biased coin until it shows heads for the first time.
Let F represent the number of throws of the fair coin until it shows heads for the first time.

(a) Find

(i) $P(B = 3)$

(2)

(ii) $P(2 \leqslant F \leqslant 8)$

(3)

(iii) $P(\{B = 3\} \cup \{F = 3\})$

(3)

Each day Chris invites his friend Shivani to play a game with these coins.
One player's score will be X and the other player's score will be Y, where

$$X = 4F \quad \text{and} \quad Y = \frac{B^2}{2}$$

Chris and Shivani will play a large number of games and the winner will be the player with the higher **total** score.

Before their first game, Chris allows Shivani to choose whether her score for every game will be X or her score for every game will be Y

(b) State, explaining your reasoning and showing any calculations you have made, which choice Shivani should make.

(5)

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

$P(B = 3) = (\frac{2}{3})^2 \times \frac{1}{3} = \frac{4}{27}$ $(= 0.148)$

(a)(ii) $F \sim Geo(\frac{1}{2})$, $P(F \leq n) = 1 - (1 - p)^n$

$P(2 \leq F \leq 8) = P(F \leq 8) - P(F \leq 1)$
$= [1 - (\frac{1}{2})^8] - [1 - \frac{1}{2}]$

$= \frac{1}{2} - \frac{1}{256} = \frac{127}{256}$ $(= 0.496)$

(a)(iii) $P(F = 3) = (\frac{1}{2})^2 \times \frac{1}{2} = \frac{1}{8}$; B and F independent

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

$P(\{B = 3\} \cup \{F = 3\}) = P(B = 3) + P(F = 3) - P(B = 3) \times P(F = 3)$

$= \frac{4}{27} + \frac{1}{8} - \frac{4}{27} \times \frac{1}{8}$

$= \frac{32 + 27 - 4}{216} = \frac{55}{216} \quad (= 0.255)$

(b) Compare the expected scores $E(X) = E(4F)$ and $E(Y) = E(\frac{B^2}{2})$

$E(F) = \frac{1}{1/2} = 2 \quad \Rightarrow \quad E(X) = 4E(F) = 4 \times 2 = 8$

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

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

$E(Y) = \frac{1}{2} E(B^2) = \frac{1}{2} \times 15 = 7.5$

Over a large number n of games the total scores are about 8n and 7.5n.
Since 8 > 7.5, Shivani should choose X, as it has the higher expected score (so the higher expected total).

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 13 marks)

7. A software company develops computer programs.
Its *Lovelace* program is found to glitch 7 times on average when it is run.

The company makes some alterations to improve the program.
It then tests to see if it has been successful in reducing the average number of glitches per run.
The glitches can be assumed to occur independently.

(a) (i) State suitable hypotheses for the test.

(ii) Find the critical region for the test at a 10% level of significance.

(iii) State the size of the test.

(4)

The altered program is actually found to glitch 5 times on average when run.

(b) Find the probability of a Type II error for the company's test.

(3)

Due to the high value of the probability of a Type II error in part (b), the company makes an entirely new version of the program.
The new version of the program is designed to have an average of fewer than k glitches per run.

The random variable T represents the number of **runs** of the program until a run with no glitches occurs.

(c) Write down the name of a suitable distribution for T

(1)

If the new version has an average of λ glitches per run, this new version of the program will be tested using the hypotheses

$$H_0: \quad \lambda = k \qquad\qquad H_1: \quad \lambda < k$$

The company will reject H_0 if the first run without any glitches of the new version occurs within n runs.

(d) Show that the power function of the test is given by

$$1 - (1 - e^{-\lambda})^n$$

(2)

The company requires P(Type II error) < 0.2
Assuming that the new version of the program has an average of 2.75 glitches per run,

(e) find the minimum value of n

(4)

(a)(i) λ = mean number of glitches per run

H_0: $\lambda = 7$ $\quad$ H_1: $\lambda < 7$

(a)(ii) Under H_0, $X \sim Po(7)$

$P(X \leq 3) = 0.0818 < 0.10$ $\quad$ $P(X \leq 4) = 0.1730 > 0.10$

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

Critical region: $X \leqslant 3$

(a)(iii) Size of test = $P(X \leqslant 3 \mid \lambda = 7) = 0.0818$

(b) $Y \sim Po(5)$

$P(\text{Type II error}) = P(Y \geqslant 4 \mid \lambda = 5) = 1 - P(Y \leqslant 3)$

$= 1 - 0.2650 = 0.7350 = 0.735$

(c) Geometric distribution

(d) $P(\text{a run has no glitches}) = P(G = 0) = e^{-\lambda}$, so $T \sim Geo(e^{-\lambda})$

Power = $P(\text{reject } H_0) = P(T \leqslant n)$

$= 1 - P(T > n) = 1 - P(\text{first } n \text{ runs all have glitches})$

$= 1 - (1 - e^{-\lambda})^n$ as required

(e) $P(\text{Type II error}) = 1 - \text{power} = (1 - e^{-2.75})^n$

$(1 - e^{-2.75})^n < 0.2$

$n \log(1 - e^{-2.75}) < \log 0.2$

$\log(1 - e^{-2.75}) < 0$, so dividing reverses the inequality:

$n > \dfrac{\log 0.2}{\log(1 - e^{-2.75})} = 24.3622\ldots$

Minimum value of $n = 25$

Question 7 continued

(Total for Question 7 is 14 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