Please check the examination details below before entering your candidate information

Candidate surname	Other names

Pearson Edexcel Level 3 GCE

Centre Number | Candidate Number

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.
- Good luck with your examination.

Turn over ▶

P66799A

1/1/1/1/

1. Kelly throws a tetrahedral die n times and records the number on which it lands for each throw.

She calculates the expected frequency for each number to be 43 if the die was unbiased.

The table below shows three of the frequencies Kelly records but the fourth one is missing.

Number	1	2	3	4
Frequency	47	34	36	x

(a) Show that $x = 55$

(1)

Kelly wishes to test, at the 5% level of significance, whether or not there is evidence that the tetrahedral die is unbiased.

(b) Explain why there are 3 degrees of freedom for this test.

(1)

(c) Stating your hypotheses clearly and the critical value used, carry out the test.

(5)

(a) If the die is unbiased, $\frac{n}{4} = 43$ since each of the 4 outcomes should occur the same amount of times.

$n = 172$

$x = 172 - 47 - 34 - 36$

$= 55$ hence shown B1

(b) We have 4 columns and only 1 constraint, that the totals need to agree.

$\therefore \nu = 4 - 1 = 3$ dof. B1

(c) H_0: The die is unbiased

H_1: The die is biased B1

Use $\chi^2 = \sum \frac{O_i^2}{E_i} - N$ to get our test statistic

$\chi^2 = \frac{47^2}{43} + \frac{34^2}{43} + \frac{36^2}{43} + \frac{55^2}{43} - 172$ M1

$= 6.744$ A1

Get the critical value from tables:

$\chi^2_3(5\%) = 7.815$ B1

Compare χ^2 and critical value: $6.744 < 7.815$.

Our χ^2 is smaller than the critical value, $\therefore$ does not fall in the critical region. Insufficient evidence to reject H_0, the die is unbiased. A1

Question 1 continued

(Total for Question 1 is 7 marks)

2. On a weekday, a garage receives telephone calls randomly, at a mean rate of 1.25 per 10 minutes.

(a) Show that the probability that on a weekday at least 2 calls are received by the garage in a 30-minute period is 0.888 to 3 decimal places.

(2)

(b) Calculate the probability that at least 2 calls are received by the garage in fewer than 4 out of 6 randomly selected, non-overlapping 30-minute periods on a weekday.

(2)

The manager of the garage randomly selects 150 non-overlapping 30-minute periods on weekdays.
She records the number of calls received in each of these 30-minute periods.

(c) Using a Poisson approximation show that the probability of the manager finding at least 3 of these 30-minute periods when exactly 8 calls are received by the garage is 0.664 to 3 significant figures.

(4)

(d) Explain why the Poisson approximation may be reasonable in this case.

(1)

two-tailed

The manager of the garage decides to test whether the number of calls received on a Saturday is different from the number of calls received on a weekday. She selects a Saturday at random and records the number of telephone calls received by the garage in the first 4 hours.

(e) Write down the hypotheses for this test.

(1)

The manager found that there had been 40 telephone calls received by the garage in the first 4 hours.

(f) Carry out the test using a 5% level of significance.

(4)

(a) Rate for 10min. $\Rightarrow$ 1.25 calls.

Rate for 30min. $\Rightarrow 3 \times 1.25 = 3.75$ calls.

$C \rightarrow$ # of calls in 30 minutes define variable

$C \sim Po(3.75)$ M1

$P(C \geq 2) = 0.88829... \Rightarrow$ 0.888 to 3dp hence shown A1

(b) We need to use binomial for this:

$X \rightarrow$ # of periods with at least 2 calls define variable

$X \sim B(6, 0.888)$ use the given probability of at least 2 calls M1

$P(X < 4) = P(X \leq 3)$

$= 0.02163 \Rightarrow$ 0.0216 to 3sf A1

Question 2 continued

(c) Use $C \sim Po(3.75)$ to get the probability of 8 calls in 30min.

$P(C=8) = 0.02281$ small $p \Rightarrow$ suitable for approximation B1

$E \to$ # of periods with exactly 8 calls

$E \sim B(150, 0.02281)$ M1

Since n is big and p is small, we can use poisson approximation

find λ:

$\lambda = 150 \times 0.02281 = 3.4215$ (mean of E)

$E \sim Po(3.421)$ M1

$P(E \geq 3) = 1 - P(E \leq 2)$

$= 0.664$ A1

(d) The number of periods is large and the probability of receiving 8 calls in 30min is small. B1

(e) 10min → 1.25 calls

4h = 24min → 30 calls = λ

Hypotheses

$H_0: \lambda = 30$

$H_1: \lambda \neq 30$ B1

(f) W → # of calls in 4hours on a Saturday

$W \sim Po(30)$ B1

$P(W \geq 40) = 1 - P(W \leq 39)$ M1 two-tailed, ∴ divide 5% by 2

$= 0.04625 > 0.025$ A1

So 40 does not fall in the critical region, insufficient evidence to reject H_0 and to suggest that the rate of calls on Saturday is different than on weekdays. A1

Question 2 continued

Question 2 continued

(Total for Question 2 is 14 marks)

3. A courier delivers parcels. The random variable X represents the number of parcels delivered successfully each day by the courier where $X \sim B(400, 0.64)$

A random sample $X_1, X_2, \ldots X_{100}$ is taken.

Estimate the probability that the mean number of parcels delivered each day by the courier is greater than 257

(4)

"mean" → Central Limit Theorem

$X \sim B(400, 0.064)$

For Binomial, the formula for expected value and variance is:

$E(X) = \underbrace{n}_{\text{sample size}} \times p$ — probability

$Var(X) = \underbrace{n}_{\text{sample size}} p(1-p)$ probability

$E(X) = 400 \times 0.64 = 256 = \mu$

$Var(X) = 25.6(1-0.64) = 0.9216 = \sigma^2$

→ Plug the found μ and σ^2 into CLT:

$\therefore \bar{X} \approx\sim N(256, 0.9216)$ in your calculator enter $\sigma = \sqrt{0.9216}$ M1A1

$P(\bar{X} > 257) = 0.1492$ dM1

↳ 0.149 to 3sf probability that the mean # of parcels is larger than 257

A1

Question 3 continued

(Total for Question 3 is 4 marks)

4. Members of a photographic group may enter a maximum of 5 photographs into a members only competition.
Past experience has shown that the number of photographs, N, entered by a member follows the probability distribution shown below.

n	0	1	2	3	4	5
$P(N = n)$	a	0.2	0.05	0.25	b	c

Given that $E(4N + 2) = 14.8$ and $P(N = 5 \mid N > 2) = \frac{1}{2}$

(a) show that $\text{Var}(N) = 2.76$

(6)

The group decided to charge a 50p entry fee for the first photograph entered and then 20p for each extra photograph entered into the competition up to a maximum of £1 per person. Thus a member who enters 3 photographs pays 90p and a member who enters 4 or 5 photographs just pays £1

Assuming that the probability distribution for the number of photographs entered by a member is unchanged,

(b) calculate the expected entry fee per member.

(3)

Bai suggests that, as the mean and variance are close, a Poisson distribution could be used to model the number of photographs entered by a member next year.

(c) State a limitation of the Poisson distribution in this case.

(1)

Question 4 continued

(a) Use $E(4N+2) = 14.8$ Formula

$4E(N) + 2 = 14.8$ M1 $E(aX+b) = aE(X)+b$

$4E(N) = 12.8$

$E(N) = 3.2$

Formulae for Mean and Variance

$E(X) = \sum x P(X=x)$ mean, μ

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

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

$E(N) = 3.2 = 0 \times a + 1 \times 0.2 + 2 \times 0.05 + 3 \times 0.25 + 4 \times b + 5 \times c$

$3.2 = 0.2 + 0.1 + 0.75 + 4b + 5c$

$3.2 = 1.05 + 4b + 5c$

$2.15 = 4b + 5c$ Eq.1 M1

Now use $P(N=5 \mid N>2) = \frac{1}{2}$

$P(N=5 \mid N>2) = \dfrac{P(N=5)}{P(N>2)} = \dfrac{c}{0.25+b+c} = 0.5$ M1

$\Rightarrow 2c = 0.25 + b + c$

$c - b = 0.25$ Eq.2

Solve Simultaneously Eq1 & Eq2:

$2.15 = 4b + 5c$ elimination method

$c - b = 0.25 \quad |\times 4| \quad 1 = -4b + 4c$ +

$3.15 = 9c$

$c = 0.35,\ b = 0.1$ values of b and c

We need $E(N^2)$ to get $Var(N)$:

$E(N^2) = 1^2 \times 0.2 + 2^2 \times 0.05 + 3^2 \times 0.25 + 4^2 \times 0.1 + 5^2 \times 0.35$

$= 0.2 + 4 \times 0.05 + 9 \times 0.25 + 16 \times 0.1 + 25 \times 0.35$ M1

$= 13$

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

$= 13 - 3.2^2$ dM1

$= 2.76$ hence shown A1

(b)

fee	0	50	70	90	100	100
P(N=n)	a	0.2	0.05	0.25	0.2	0.35

M1

$E(\text{fee}) = 5 \times 0.2 + 70 \times 0.05 + 90 \times 0.25 + 100 \times 0.2 + 100 \times 0.35$

$= 81p$ expected entry fee per person A1

Question 4 continued

(c) The distribution will asign a large probability to values $N > 5$ B1

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. Asha, Davinda and Jerry each have a bag containing a large number of counters, some of which are white and the rest are red.
Each person draws counters from their bag one at a time, notes the colour of the counter and returns it to their bag.

The probability of Asha getting a red counter on any one draw is 0.07

(a) Find the probability that Asha will draw at least 3 white counters before a red counter is drawn.

3 "failures"

(2)

(b) Find the probability that Asha gets a red counter for the second time on her 9th draw.

(2)

The probability of Davinda getting a red counter on any one draw is p.
Davinda draws counters until she gets n red counters. The random variable D is the number of counters Davinda draws.

Given that the mean and the standard deviation of D are 4400 and 660 respectively,

(c) find the value of p.

(4)

Jerry believes that his bag contains a smaller proportion of red counters than Asha's bag. To test his belief, Jerry draws counters from his bag until he gets a red counter. Jerry defines the random variable J to be the number of counters drawn up to and including the first red counter.

(d) Stating your hypotheses clearly and using a 10% level of significance, find the critical region for this test.

(5)

Jerry gets a red counter for the first time on his 34th draw.

(e) Giving a reason for your answer, state whether or not there is evidence that Jerry's bag contains a smaller proportion of red counters than Asha's bag.

(2)

Given that the probability of Jerry getting a red counter on any one draw is 0.011

(f) show that the power of the test is 0.702 to 3 significant figures.

(3)

(a) Geometric

$P(\text{at least 3 white}) = (1-0.07)^3$ M1

$= 0.8043 \rightarrow 0.804$ to 3sf A1

(b) Negative Binomial

$P(2^{nd}$ red on 9^{th} draw$) = \binom{8}{1} 0.93^7 \times 0.07^2$ M1

$= 0.02358 \rightarrow 0.0236$ to 3sf A1

DO NOT WRITE IN THIS AREA

Question 5 continued

(c) $D \sim NB(n, p)$

For Neg. Binomial from Formula Booklet:

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

Substitute:

$\frac{n}{p} = 4400$ $\quad \frac{n(1-p)}{p^2} = 660^2$ ← we are given σ, but $Var(X) = \sigma^2$ M1A1

$n = 4400p$ $\quad \frac{4400p(1-p)}{p^2} = 435600$

Solve for p:

$\frac{\cancel{p}(1-p)}{p^{\cancel{2}}} = \frac{4356\cancel{00}}{44\cancel{00}}$

$\frac{1-p}{p} = 99$

$1 - p = 99p$ M1

$1 = 100p$, $p = 0.01$ A1

(d) Hypotheses

$H_0: p = 0.07$

$H_1: p < 0.07$ B1

$J \rightarrow$ # of draws when Jerry gets a red counter

$J \sim Geo(0.07)$ M1

$P(J \geq c) < 0.1$ M1

$(1 - 0.07)^{c-1} < 0.1$

$0.93^{c-1} < 0.1$

$(c-1)\log 0.93 < \log 0.1$

$c - 1 > \frac{\log 0.1}{\log 0.93} \Rightarrow c > \frac{\log 0.1}{\log 0.93} + 1$ $\quad c > 32.72...$ M1

flip inequality sign as $\log 0.93 < 1$ $\quad$ round up to next integer as geometric is discrete

∴ crit. region → $J \geq 33$ A1

(e) $34 > 33$ ∴ falls in critical region. There is sufficient evidence to reject H_0 and to suggest that Jerry's bag has a smaller proportion of red counters than Asha's. M1A1

(f)

Power is the probability of rightfully rejecting H_0 using the actual probability ∴ P(in crit. region | p = actual)

$P(J \geq 33 \mid p = 0.011) = (1 - 0.011)^{32}$ M1M1

given $\quad = 0.7019 \rightarrow 0.702$ to 3sf A1

Turn over ▶

P 6 6 7 9 9 A 0 1 5 2 4

Question 5 continued

Question 5 continued

(Total for Question 5 is 18 marks)

Turn over

6. The probability generating function of the random variable X is

$$G_X(t) = k(1+2t)^5$$

where k is a constant.

(a) Show that $k = \dfrac{1}{243}$ **(2)**

(b) Find $P(X = 2)$ **(2)**

(c) Find the probability generating function of $W = 2X + 3$ **(2)**

The probability generating function of the random variable Y is

$$G_Y(t) = \frac{t(1+2t)^2}{9}$$

Given that X and Y are independent,

(d) find the probability generating function of $U = X + Y$ in its simplest form. **(2)**

(e) Use calculus to find the value of $\mathrm{Var}(U)$ **(6)**

(a) We know that $G_X(1) = 1$ M1

$1 = k(1+2(1))^5$

$1 = k(3)^5$

$k = \frac{1}{243}$ hence shown A1

(b) $P(X=2)$ is the coefficient of t^2 in the expansion of $G_X(t)$. M1

Use Binomial Expansion:

$G_X(t) = \frac{1}{243}(\ldots + \binom{5}{2}(2t)^2 + \ldots)$

coeff: $\frac{1}{243} \times 10 \times 4 = \frac{40}{243}$

$\therefore P(X=2) = \frac{40}{243}$ A1

(c) For $X = aY + b$, $G_Y(t) = t^b G_X(t^a)$

$W = 2X + 3$

$G_W(t) = t^3 \cdot G_X(t^2)$ substitute the pgf given and the k we found M1

$= \frac{t^3}{243}(1+2t^2)^5$

A1

Question 6 continued

(d) For $U = X + Y \rightarrow G_U(t) = G_X(t) \times G_Y(t)$

Substitute:

$G_U(t) = \frac{1}{243}(1+2t)^5 \times \frac{t(1+2t)^2}{9}$ M1

$= \frac{t(1+2t)^7}{2187}$ A1

(e) For PGFs:

$E(X) = G_X'(1)$

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

$G_U(t) = \frac{t(1+2t)^7}{2187}$

chain rule

$G_U'(t) = \frac{7t(1+2t)^6 \times 2}{2187} + \frac{(1+2t)^7}{2187}$

$= \frac{14t(1+2t)^6}{2187} + \frac{(1+2t)^7}{2187}$ M1

differentiate and use chain rule and product rule

★Product Rule:

$\frac{d}{dx}(uv) = uv' + u'v$

$G_U'(1) = \frac{14(3)^6}{2187} + \frac{(3)^7}{2187} = \frac{17}{3}$ A1

$G_U''(t) = \frac{6 \times 14t(1+2t)^5 \times 2}{2187} + \frac{14(1+2t)^6}{2187} + \frac{(1+2t)^7}{2187}$ M1 again: differentiate and use chain rule and product rule

$G_U''(1) = \frac{168(1)(3)^5}{2187} + \frac{14(3)^6}{2187} + \frac{(3)^7}{2187} = 28$ A1

Substitute into $Var(U) = G_U''(1) + G_U'(1) - [G_U'(1)]^2$

$Var(U) = 28 + \frac{17}{3} + \left(\frac{17}{3}\right)^2$ M1

$= \frac{14}{9}$ A1

Question 6 continued

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 14 marks)

7. A manufacturer has a machine that produces lollipop sticks.
The length of a lollipop stick produced by the machine is normally distributed with unknown mean μ and standard deviation 0.2

Farhan believes that the machine is not working properly and the mean length $\bar{L}$ of the lollipop sticks has decreased.
He takes a random sample of size n to test, at the 1% level of significance, the hypotheses

$$H_0: \mu = 15 \qquad H_1: \mu < 15$$

(a) Write down the size of this test.

(1)

Given that the actual value of μ is 14.9

(b) (i) calculate the minimum value of n such that the probability of a Type II error is less than 0.05
Show your working clearly.

(6)

(ii) Farhan uses the same sample size, n, but now carries out the test at a 5% level of significance. Without doing any further calculations, state how this would affect the probability of a Type II error.

(1)

(a) Size is the probability of rightfully rejecting H_0 ∴ P(falling in critical region)
Since Normal Distribution is continuous, that's equal to the significance level
∴ Size = 0.01 B1

(b)i. Type II Error is the probability of falsefully accepting H_0. (H_0 is false but we accept it)
Let's say the critical region is $\bar{L} < k$.
$\bar{L} \sim N(15, \frac{0.2}{\sqrt{n}}^2)$
Convert to Standard Normal
$\frac{k-15}{\frac{0.2}{\sqrt{n}}} = \text{InvN}(0.01)$ with $\mu=0, \sigma=1 \longrightarrow \frac{k-15}{\frac{0.2}{\sqrt{n}}} = -2.3263$ M1

Make k the subject: $k = 15 - \frac{0.46526}{\sqrt{n}}$ A1 convert to standard Normal:
$P(\text{Type II}) = P(\underbrace{\bar{L} > k}_{\text{falling outside the CR}} \mid \mu = 14.9) > 0.95 \longrightarrow \frac{k-14.9}{\frac{0.2}{\sqrt{n}}} > \text{InvN}(0.95)$

Substitute our k from above:
$\frac{(15 - \frac{0.46526}{\sqrt{n}}) - 14.9}{\frac{0.2}{\sqrt{n}}} > 1.6449$ M1A1

Make n the sub
$0.1 - \frac{0.46526}{\sqrt{n}} > \frac{0.32898}{\sqrt{n}}$ dM1
$0.1\sqrt{n} > 0.79424$
$\sqrt{n} > 7.9424 \Rightarrow n = 64$ A1

Question 7 continued

ii. Probability of Type II error will increase. B1

Question 7 continued

(Total for Question 7 is 8 marks)

TOTAL FOR PAPER IS 75 MARKS