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

©2021 Pearson Education Ltd.

Pearson

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) $4 \times 43 = 47 + 34 + 36 + x$

$x = 172 - 117$

$x = 55$ ①

(b) $V = 4 - 1 = 3$ because the only constraint is that the totals agree. ①

(c) H_0: the die is unbiased

H_1: the die is biased ①

Perform a Chi-Squared test to compare observed and expected frequencies

Find $X^2 = \sum_{1}^{N} \frac{(O_i - E_i)^2}{E}$ as the test statistic

Test stat. $= \frac{(47-43)^2}{43} + \frac{(34-43)^2}{43} + \frac{(36-43)^2}{43} + \frac{(55-43)^2}{43}$ 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

Question 1 continued

Test stat. = 6.744 ①

Using 3 degrees of freedom, 0.05 significance level:

$\chi^2_{(3,0.05)} = 7.815$ ① ← from statistical tables in the formula booklet.

Not in the critical region since 7.815 > 6.744

therefore insufficient evidence to reject H_0. ①

Test is inconclusive, suggesting the die is unbiased.

(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)

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) $C \sim \text{Poisson}(3.75)$ ① ← 1.25 per 10 min, 3.75 per 30 min, ×3

$P(C \geq 2) = 0.8883$ ① ← using stats tables or calculator

(b) number of tests = 6 (days)

probability of desired outcome = 0.8883 (at least 2 calls)

$D \sim B(6, 0.8883)$ ① so $P(D \leq 3) = 0.0216$ 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

Question 2 continued

(c) $P(C = 8) = 0.02281$ ① ← prob of exactly 8 calls

If $X \sim B(n, p)$ where n is large, p is small, then X can be approximated by $X \sim \text{Poisson}(np)$.

let E = day with exactly 8 calls, then

$E \sim B(150, 0.02281)$

mean $= n \times p = 150 \times 0.02281 = 3.4215$ ①

$E \sim \text{Poisson}(3.4215)$

$P(E \geq 3) = 1 - P(E \leq 2)$ ①

$= 1 - 0.336$

$= 0.664$ ①

(d) The number of periods is large and the probability of receiving 8 calls in 30 minutes is small. ①

(e) $H_0: \lambda = 30$ $\quad H_1: \lambda \neq 30$ ①

↑

1.25 calls per 10 mins, 4 hrs = 24 × 10 mins

24 × 1.25 = 30 calls

Question 2 continued

(f) Test whether 40 calls is significantly different from the expected value (30).

$X \sim \text{Poisson}(30)$ (1)

$P(X \geq 40) = 1 - P(X \leq 39)$ (1)

$= 0.04625$ (1)

5% significance level given a two-tailed test,

sig. level $= 0.05 \div 2 = 0.025$ as we use $\neq$ in H_1

$0.04625 > 0.025$ $\therefore$ insufficient evidence to reject H_0. There is insufficient evidence at the 5% significance level to suggest that the number of calls received on a Saturday is different to the number of calls received on a weekday. (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 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)

Central Limit Theorem: $\bar{X} \approx N\left(\mu, \frac{\sigma^2}{n}\right)$

Given $X \sim B(400, 0.64)$.

$n = 400$

$p = 0.64$

$\mu = E(X) = np = 400 \times 0.64 = 256$ ①

$\sigma^2 = Var(X) = np(1-p) = 256(1-0.64) = 92.16$

$\therefore \bar{X} \sim N(256, 0.9216)$ ① $\leftarrow \frac{\sigma^2}{n} = \frac{92.16}{100} = 0.9216$

$P(\bar{X} > 257) \leftarrow$ can be standardised to $Z \sim N(0,1)$

$z = \frac{x-\mu}{\sigma} = \frac{257-256}{\sqrt{0.9216}} = 1.0417$ ①

$P(Z > 1.0417) = 0.1492$ ① $\leftarrow$ calculator or stats tables

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 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 $Var(N) = 2.76$

(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)

(a) $E(4N + 2) = 14\ 8$

$4E(N) + 2 = 14\ 8$

$4E(N) = 12.8$

$E(N) = 3\ 2$ 1

$E(N) = \sum n \times P(N = n)$

$3.2 = 1 \times 0\ 2 + 2 \times 0\ 05 + 3 \times 0\ 25 + 4b + 5c$ 1

$2\ 15 = 4b + 5c$

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

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

$\frac{c}{0.25+b+c} = \frac{1}{2}$ (1) ← N = 5 ← N > 2

$2c = 0.25 + b + c$

$c = 0.25 + b$ sub into $2.15 = 4b + 5c$

$2.15 = 4b + 5(0.25 + b)$

$2.15 - 1.25 = 9b$

$b = 0.1$

$c = 0.1 + 0.25 = 0.35$ (1)

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

$+ 5^2 \times 0.35$ (1)

$E(N^2) = 13$

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

$= 13 - 3.2^2$ (1)

$= 2.76$ (1)

Question 4 continued

(b)

fee	0	50	70	90	100	100
$P(N=n)$	a	0 2	0·05	0·25	b	c

1

$E(N) = 50 \times 0{\cdot}2 + 70 \times 0{\cdot}05 + 90 \times 0{\cdot}25 + 100 \times 0{\cdot}1$

$+ 100 \times 0\ 35$ 1

$E(N) = 81p$ 1

(c) Poisson distribution assigns substantial probability to $N > 5$. 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

(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.
(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)

1 - P(red counter)

(a) P(at least 3 white counters) $= (1 - 0.07)^3$ 1

$= 0.8043$ 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

Question 5 continued

(b) Using negative binomial model, probability that r^{th} success occurs on x^{th} trial is $\binom{x-1}{r-1} p^r (1-p)^{(x-1)-(r-1)}$

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

$= 0.02358$ ①

(c) $\text{mean} = \frac{n}{p} = 4400$

$\text{variance} = \frac{n(1-p)}{p^2} = 660^2$ ①

$n = 4400p$

$\frac{4400p(1-p)}{p^2} = 660^2$ $\times p$

$4400(1-p) = 660^2 p$ $\div 4400$

$1-p = 99p$ ①

$1 = 100p$

$\therefore p = 0.01$ ①

(d) $H_0: p = 0.07$ $H_1: p < 0.07$

testing whether the probability is overstated

$J \sim Geo(0.07)$

$P(J \geq c) < 0.1$ ← using a 10% = 0.1 significance level

Question 5 continued

$(1-0.07)^{c-1} < 0.1$ ← c is the number of counters drawn

$c-1 > \dfrac{\log 0.1}{\log 0.93}$

$\therefore\ c > 32.72$. so critical region is $J \geq 33$.

(e) $34 > 33$ so is in the critical region ①

There is evidence to suggest that Jerry's bag contains a smaller proportion of red counters than Asha's bag ①

(f) Power of test $= P(J \geq 33 \mid p = 0.011)$ ①

probability that x is in the critical region.

$= (1-0.011)^{32}$ ①

$= 0.7019$ 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 5 continued

(Total for Question 5 is 18 marks)

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 $k = \frac{1}{243}$ **(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 $\text{Var}(U)$ **(6)**

(a) $G_X(1) = 1$ ① ← for all probability-generating functions.

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

$k \times 3^5 = 1$

$k = \frac{1}{243}$ ①

(b) $P(X=2)$ is given by the coefficient of t^2 in the expansion of $G_X(t)$ ← use binomial expansion to find t^2 term

$k\left[{}^5C_2 \times 1^{5-2} \times (2t)^2\right]$ ① $= \frac{1}{243} \times 40$

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

Question 6 continued

(c) $Y = aX + b \rightarrow G_Y(t) = t^b G_X(t^a)$

$W = 2X + 3$ so $a = 2$, $b = 3$

$G_W(t) = t^3 \times G_X(t^2)$ replace t with t^2

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

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

(d) If $U = X + Y$, $G_U(t) = G_X(t) \times G_Y(t)$

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

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

$G_U(t) = \frac{t(1+2t)^7}{2187}$ ①

(e) $Var(U) = G''_U(1) + G'_U(1) - [G'_X(1)]^2$

$G'_U(t) = \frac{1}{2187}\left[1 \times (1+2t)^7 + t \times 7 \times 2(1+2t)^6\right]$

$G'_U(t) = \frac{(1+2t)^7}{2187} + \frac{14t(1+2t)^6}{2187}$ ① product rule: $f' \cdot g + f \cdot g'$

$\therefore G'_U(1) = \frac{3^7 + 14 \times 3^6}{2187} = \frac{17}{3}$ ①

Question 6 continued

$$G_u''(t) = \frac{14(1+2t)^6}{2187} + \frac{14(1+2t)^6}{2187} + \frac{14t \times 6 \times 2(1+2t)^5}{2187}$$

$$G_u''(t) = \frac{28(1+2t)^6}{2187} + \frac{168t(1+2t)^5}{2187}$$

$$G_u''(1) = \frac{28(1+2(1))^6}{2187} + \frac{168(1)(1+2(1))^5}{2187} \quad 1$$

$$G_u''(1) = 28 \quad 1$$

$$\text{Var}(u) = 28 + \frac{17}{3} - \left(\frac{17}{3}\right)^2 \quad 1 \quad = \frac{14}{9} \quad 1$$

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 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 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)

$t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$ $\quad z = \dfrac{x - \mu}{\sigma}$

(a) size of test = 0.01 ①
(significance level)

(b)(i) Let the critical region be $\bar{x} < k$ ①

Using stats. tables and $p = 0.01$, $z = 2.3263$.

$$z = \frac{\bar{x} - \mu}{\sigma} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

$$-2.3263 = \frac{k - 15}{\frac{0.2}{\sqrt{n}}} \quad ①$$

$$k = 15 - \frac{0.46526}{\sqrt{n}} \quad ①$$

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

DO NOT WRITE IN THIS AREA

Question 7 continued

sub in k

$$\frac{15 - \frac{0\ 46526}{\sqrt{n}} - 14\ 9}{\frac{0\ 2}{\sqrt{n}}} > 1.6449 \quad (1)$$

one-tailed test so use stats table to find $p = 0.05$, $z = 1\ 6449$

$$15 - \frac{0\ 46526}{\sqrt{n}} - 14\ 9 > \frac{0.32898}{\sqrt{n}}$$

$$0\ 1 > \frac{0\ 79424}{\sqrt{n}} \quad (1)$$

$$\sqrt{n} > 7\ 9424$$

$$n = 64 \quad (1)$$

since we want min. value of n such that $\sqrt{n} > 7.9424$

(ii) The probability of a type II error would decrease. (1)

Question 7 continued

(Total for Question 7 is 8 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