Please check the examination details below before entering your candidate information

Candidate surname	Other names

Centre Number | Candidate Number

Pearson Edexcel Level 3 GCE

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

©2022 Pearson Education Ltd.

Pearson

1. A researcher is investigating the number of female cubs present in litters of size 4
He believes that the number of female cubs in a litter can be modelled by B(4, 0.5)
He randomly selects 100 litters each of size 4 and records the number of female cubs.
The results are recorded in the table below.

Number of female cubs	0	1	2	3	4
Observed number of litters	10	33	33	15	9

He calculated the expected frequencies as follows

Number of female cubs	0	1	2	3	4
Expected number of litters	6.25	r	s	r	6.25

(a) Find the value of r and the value of s

(3)

(b) Carry out a suitable test, at the 5% level of significance, to determine whether or not the number of female cubs in a litter can be modelled by B(4, 0.5)
You should clearly state your hypotheses and the critical value used.

(6)

(a) $r = P(X = 3) \times 100$ ①

$r = 0.25 \times 100$ ← use value in binomial stats. table with $n = 4$, $p = 0.5$

$r = 25$ ①

$s = P(X = 2) \times 100$

$s = 0.375 \times 100$

$s = 37.5$ 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

Question 1 continued

(b) To test if a model is suitable, use a Chi-squared test

H_0 $B(4, 0.5)$ is a suitable model

H_1 $B(4, 0.5)$ is not a suitable model. ①

Observed	10	33	33	15	9
Expected	6.25	25	37.5	25	6.25

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

$$\chi^2 = \frac{(10-6.25)^2}{6.25} + \frac{(33-25)^2}{25} + \frac{(33-37.5)^2}{37.5} + \frac{(15-25)^2}{25} + \frac{(9-6.25)^2}{6.25}$$ ①

$\chi^2 = 2.25 + 2.56 + 0.54 + 4 + 1.21$

$\chi^2 = 10.56$ ①

$v = 5 - 1 = 4$ ① ← degrees of freedom

Critical Value $= \chi^2_3(0.05) = 9.488$ ① ← stats. table

↑ significance level

$10.56 > 9.488$

∴ Reject H_0, there is evidence that the researcher's model is not suitable. ①

(Total for Question 1 is 9 marks)

2. The discrete random variable X has probability distribution

x	-5	-1	0	5	b
$P(X = x)$	0.3	0.25	0.1	0.15	0.2

where b is a constant and $b > 5$

(a) Find $E(X)$ in terms of b

(1)

Given that $Var(X) = 34.26$

(b) find the value of b

(4)

(c) Find $P(X^2 < 2 - 3X)$

(4)

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

$E(X) = -5 \times 0.3 + -1 \times 0.5 + 0 + 5 \times 0.15 + b \times 0.2$

$E(X) = 0.2b - 1$ ①

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

$E(X^2) = (-5)^2 \times 0.3 + (-1)^2 \times 0.25 + (5)^2 \times 0.15 + b^2 \times 0.2$

$E(X^2) = 11.5 + 0.2b^2$ ①

$Var(X) = 11.5 + 0.2b^2 - (0.2b - 1)^2$ ①

$34.26 = 11.5 + 0.2b^2 - 0.04b^2 + 0.4b - 1$

$0.16b^2 + 0.4b - 23.76 = 0$ ①

$b = 11$ ① ← ignore other result because $b > 5$

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) Find the values of X that satisfy $X^2 < 2 - 3X$

X	-5	-1	0	1	11	
X^2	25	1	0	1	121	1
$2-3X$	17	5	2	-13	-31	1

only values of X that satisfy $X^2 < 2 - 3x$ are 0 and -1.

$P(X^2 < 2 - 3x) = P(X = 0) + P(X = -1)$ 1

$= 0.1 + 0.25$

$= 0.35$ 1

(Total for Question 2 is 9 marks)

3. During the summer, mountain rescue team A receives calls for help randomly with a rate of 0.4 per day.

(a) Find the probability that during the summer, mountain rescue team A receives at least 19 calls for help in 28 randomly selected days.

(2)

The leader of mountain rescue team A randomly selects 250 summer days from the last few years.
She records the number of calls for help received on each of these days.

(b) Using a Poisson approximation, estimate the probability of the leader finding at least 20 of these days when more than 1 call for help was received by mountain rescue team A.

(4)

Mountain rescue team A believes that the number of calls for help per day is lower in the winter than in the summer. The number of calls for help received in 42 randomly selected winter days is 8

(c) Use a suitable test, at the 5% level of significance, to assess whether or not there is evidence that the number of calls for help per day is lower in the winter than in the summer. State your hypotheses clearly.

(4)

During the summer, mountain rescue team B receives calls for help randomly with a rate of 0.2 per day, independently of calls to mountain rescue team A.

The random variable C is the total number of calls for help received by mountain rescue teams A and B during a period of n days in the summer.
On a Monday in the summer, mountain rescue teams A and B each receive a call for help.

Given that over the next n days $P(C = 0) < 0.001$

(d) calculate the minimum value of n

(3)

(e) Write down an assumption that needs to be made for the model to be appropriate.

(1)

(a) Let W = number of calls in 1 day

$0.4 \times 28 = 11.2 \Rightarrow W \sim Po(11.2)$ ← expected / mean number of calls

$P(W \geq 19) = 1 - P(W \leq 18)$ ①

$= 1 - 0.979$ ← using stats. tables

$= 0.021$ 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 3 continued

(b) Let S = number of calls in 1 day

$S \sim \text{Poisson}(0.4)$

$P(S > 1) = 0.0616$ ① ← probability of more than 1 call in a day

Model as a binomial distribution : $X \sim B(250, 0.0616)$ ①

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

$250 \times 0.0616 = 15.39$

$X \sim \text{Poisson}(15.39)$ ①

$P(X \geq 20) = 1 - P(X \leq 19)$

$= 1 - 0.852$

$= 0.148$ ①

(c) $H_0 : \lambda = 16.8$ $\quad H_1 : \lambda < 16.8$ ①

$42 \times 0.4 = 16.8$ $\quad U \sim \text{Poisson}(16.8)$ ①

↑ expected calls in 42 days

$P(U \leq 8) = 0.014$ ①

5% significance level

$0.014 < 0.05$ ∴ Reject H_0, there is evidence to suggest that the number of calls received per day is lower in the winter. ①

Question 3 continued

(d) Combine means for teams A and B.

$0.4 + 0.2 = 0.6$ calls per day

$0.6 \times n = 0.6n$ calls per n days

$C \sim \text{Poisson}(0.6n)$ (1)

$e^{-0.6n} < 0.001$ (1)

$\ln(0.001) - -0.6 = n$

$n = 11.5$

So minimum is $n = 12$ (1)

(e) The rate of calls per day is constant (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 3 continued

(Total for Question 3 is 14 marks)

4. In a game a spinner is spun repeatedly. When the spinner is spun, the probability of it landing on blue is 0.11

(a) Find the probability that the spinner lands on blue

(i) for the first time on the 6th spin,

(2)

(ii) for the first time before the 6th spin,

(2)

(iii) exactly 4 times during the first 6 spins,

(2)

(iv) for the 4th time on or before the 6th spin.

(4)

Zac and Izana play the game. They take turns to spin the spinner. The winner is the first one to have the spinner land on blue. Izana spins the spinner first.

(b) Show that the probability of Zac winning is 0.471 to 3 significant figures.

(3)

(a)(i) $W \sim Geo(0.11)$

$P(W=6) = 0.11(1-0.11)^{6-1}$ ① $\leftarrow P(X=x) = p(1-p)^{x-1}$

$= 0.11(0.89)^5$

$= 0.0614$ ①

(a)(ii) $P(W \leq 5) = 1-(1-0.11)^5$ ① $\leftarrow P(X \leq x) = 1-(1-p)^x$

before the 6th $= 0.442$ ①

(a)(iii) $X \sim B(6, 0.11)$

$P(X=4) = 0.00174$ ① $\leftarrow$ use stats tables or a calculator

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

(a)(iv) $Y \sim NB(4, 0.11)$ ① ← Y trials are needed to achieve 4 successes

$P(Y \geq 4) = P(Y=4) + P(Y=5) + P(Y=6)$ ①

For an NB model $P(X=x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$

$P(Y \geq 4) = \binom{3}{3} 0.11^4 (1-0.11)^0 + 0.11 \times \binom{4}{3} 0.11^3 (1-0.11)^1$

$+ 0.11 \times \binom{5}{3} 0.11^3 (1-0.11)^2$ ①

$P(Y \geq 4) = 0.00183$ ①

(b) We need the case that I loses $(1-p = 0.89)$ and Z wins (0.11) for any pair of spins. Since it is Z's turn every $2n-1$ spins, this is a geometric probability series.

$P(Z \text{ wins}) = 0.11 \times 0.89 + 0.11^{2n-1} \times 0.89 \ldots$

$= 0.11 \times 0.89 + 0.11 \times 0.89^3 + 0.11 \times 0.89^5$ ①

This is a geometric series with $a = 0.11 \times 0.89$ and $d = 0.89^2$, so it has sum:

$S = \dfrac{0.11 \times 0.89}{1-(0.89)^2}$ ① $= 0.471$ ①

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 13 marks)

5. A random sample of 150 observations is taken from a geometric distribution with parameter 0.3

Estimate the probability that the mean of the sample is less than 3.45

(5)

Geo(0 3)

$\mu = \frac{1}{0\ 3} = \frac{10}{3}$ ① $\leftarrow \mu = \frac{1}{p}$

$\sigma^2 = \frac{1-0\ 3}{0\ 3^2} = \frac{70}{9}$ ① $\leftarrow \sigma^2 = \frac{1-p}{p^2}$

CENTRAL LIMIT THEOREM.

If $X_1 \quad X_n$ is a random sample with μ and σ^2, then:

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

$\frac{\sigma^2}{n} = \frac{\frac{70}{9}}{150} = \frac{7}{135}$

$\bar{X} \approx N\left(\frac{10}{9}, \frac{7}{135}\right)$ ① ①

$P(\bar{X} < 3\ 45) = 0\ 696$ ①

from stats. tables or calculator

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 5 marks)

6. The discrete random variable V has probability distribution

v	2	3	4
$P(V = v)$	$\frac{9}{25}$	$\frac{12}{25}$	$\frac{4}{25}$

(a) Show that the probability generating function of V is

$$G_V(t) = t^2\left(\frac{2}{5}t + \frac{3}{5}\right)^2$$

(2)

The discrete random variable W has probability generating function

$$G_W(t) = t\left(\frac{2}{5}t + \frac{3}{5}\right)^5$$

(b) Use calculus to find

(i) $E(W)$

(2)

(ii) $Var(W)$

(4)

Given that V and W are independent,

(c) find the probability generating function of $X = V + W$ in its simplest form.

(2)

The discrete random variable $Y = 2X + 3$

(d) Find the probability generating function of Y

(2)

(e) Find $P(Y = 15)$

(2)

(a) $G_V(t) = \sum P(V=v)t^v$

$G_V(t) = \frac{9}{25}t^2 + \frac{12}{25}t^3 + \frac{4}{25}t^4$ 1

$G_V(t) = t^2\left(\frac{9}{25} + \frac{12}{25}t + \frac{4}{25}t^2\right)$

$G_V(t) = t^2\left(\frac{2}{5}t + \frac{3}{5}\right)^2$ 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

(b)(i) $G_W(t) = t\left(\frac{2}{5}t + \frac{3}{5}\right)^5$ and $E(W) = G_W'(1)$

Using product rule $(fg)' = f'g + fg'$

$f = t \implies f' = 1$

$g = \left(\frac{2}{5}t + \frac{3}{5}\right)^5 \implies g' = \frac{2}{5} \times 5 \times \left(\frac{2}{5}t + \frac{3}{5}\right)^{5-1}$

$g' = 2\left(\frac{2}{5}t + \frac{3}{5}\right)^4$

$G_W'(t) = 2t\left(\frac{2}{5}t + \frac{3}{5}\right)^4 + \left(\frac{2}{5}t + \frac{3}{5}\right)^5$ 1

$G_W'(1) = 2 \times 1 \times \left(\frac{2}{5} \times 1 + \frac{3}{5}\right)^4 + \left(\frac{2}{5} \times 1 + \frac{3}{5}\right)^5$

$G_W'(1) = 2 \times 1^4 + 1^5$

$G_W'(1) = 3$ 1

(b)(ii) $Var(W) = G''_W(1) + G_W'(1) - [G_W'(1)]^2$

$G_W'(t) = 2t\left(\frac{2}{5}t + \frac{3}{5}\right)^4 + \left(\frac{2}{5}t + \frac{3}{5}\right)^5$

$f = 2t \implies f' = 2$

$g = \left(\frac{2}{5}t + \frac{3}{5}\right)^4 \implies g' = 4 \times \frac{2}{5} \times \left(\frac{2}{5}t + \frac{3}{5}\right)^{4-1}$

Question 6 continued

$$g' = \frac{8}{5}\left(\frac{2}{5}t + \frac{3}{5}\right)^3$$

$$G_W''(t) = 2t \times \frac{8}{5}\left(\frac{2}{5}t + \frac{3}{5}\right)^3 + 2\left(\frac{2}{5}t + \frac{3}{5}\right)^4$$

$$+ \frac{2}{5} \times 5 \times \left(\frac{2}{5}t + \frac{3}{5}\right)^{5-1}$$

$$G_W''(t) = 2\left(\frac{2}{5}t + \frac{3}{5}\right)^4 + \frac{16}{5}t\left(\frac{2}{5}t + \frac{3}{5}\right)^3 + 2\left(\frac{2}{5}t + \frac{3}{5}\right)^4 \quad 1$$

$$G_W''(t) = 4\left(\frac{2}{5}t + \frac{3}{5}\right)^4 + \frac{16}{5}t\left(\frac{2}{5}t + \frac{3}{5}\right)^3$$

$$G_W''(1) = 4\left(\frac{2}{5} \times 1 + \frac{3}{5}\right)^4 + \frac{16}{5} \times 1 \times \left(\frac{2}{5} \times 1 + \frac{3}{5}\right)^3$$

$$G_W''(1) = 4 \times 1^4 + \frac{16}{5} \times 1^3$$

$$G_W''(1) = \frac{36}{5} \quad 1$$

$$Var(W) = \frac{36}{5} + 3 - (3)^2 \quad 1$$

$$Var(W) = \frac{6}{5} \quad 1$$

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 6 continued

(c) $G_X(t) = G_V(t) \times G_W(t)$

$G_X(t) = t^2\left(\frac{2}{5}t + \frac{3}{5}\right)^2 \times t\left(\frac{2}{5}t + \frac{3}{5}\right)^5$ ①

$G_X(t) = t^3\left(\frac{2}{5}t + \frac{3}{5}\right)^7$ ①

(d) $G_Y(t) = t^3 \times (t^2)^3 \times \left(\frac{2}{5}t^2 + \frac{3}{5}\right)^7$ ①

$G_Y(t) = t^9\left(\frac{2}{5}t^2 + \frac{3}{5}\right)^7$ ① $Y = aX + b$ $G_Y(t) = t^b G_X(t^a)$

(e) For $P(Y = 15)$ find coefficient of t^{15}

$t^{15} = t^9 \times t^6$ so use binomial expansion for t^6

$t^9 \times {}^7C_3\left(\frac{2}{5}t^2\right)^3\left(\frac{3}{5}\right)^4 = \left(35 \times \frac{8}{125} \times \frac{81}{625}\right)t^9$ ①

$= 0.2903t^9$

$\therefore P(Y=15) = 0.2903$ ①

(Total for Question 6 is 14 marks)

7. A machine fills bags with flour. The weight of flour delivered by the machine into a bag, X grams, is normally distributed with mean μ grams and standard deviation 30 grams. To check if there is any change to the mean weight of flour delivered by the machine into each bag, Olaf takes a random sample of 10 bags. The weight of flour, x grams, in each bag is recorded and $\bar{x} = 1020$

(a) Test, at the 5% level of significance, $H_0 : \mu = 1000$ against $H_1 : \mu \neq 1000$

2-tailed

(4)

Olaf decides to alter the test so that the hypotheses are $H_0 : \mu = 1000$ and $H_1 : \mu > 1000$ but keeps the level of significance at 5%

1-tailed

He takes a second sample of size n and finds the critical region, $\bar{X} > c$

(b) Find an equation for c in terms of n

(2)

When the true value of μ is 1020 grams, the probability of making a Type II error is 0.0050, to 2 significant figures.

(c) Calculate the value of n and the value of c

(5)

(a) $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

$\frac{\sigma^2}{n} = \frac{30^2}{10} = \frac{900}{10} = 90$

$\bar{X} \sim N(1000, 90)$ ①

$P(\bar{X} > 1020) = 0.0175$ ① ← from stats tables or calculator

more extreme than observed mean

5% significance, 2-tailed test ⇒ $0.05 \div 2 = 0.025$

$0.0175 < 0.025$ ∴ reject H_0. ①

There is evidence to suggest a change in the mean weights of bags of flour. 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

(b) $\bar{y} \sim N\left(1000, \frac{900}{n}\right)$ ← as n is no longer 10, the distribution changes.

Convert to standard normal distribution:

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

$z = 1.6449$ ← from standard normal stats tables with $z \sim N(0, 1^2)$ and 0.05 / 5% significance level

$$1.6449 = \frac{c - 1000}{30/\sqrt{n}} \quad ①$$

$$\frac{49.347}{\sqrt{n}} = c - 1000$$

$$c = 1000 + \frac{49.347}{\sqrt{n}} \quad ①$$

(c) $P(\bar{x} \leq c \mid \mu = 1020) = 0.005$

$\bar{x} \sim N\left(1020, \frac{900}{n}\right)$

$\frac{c - 1020}{30/\sqrt{n}} = -2.5758$ ← from stats tables $z \sim N(0, 1^2)$ and 5%

$$\frac{1000 + \frac{49.347}{\sqrt{n}} - 1020}{30/\sqrt{n}} = -2.5758 \quad ①$$

↑ sub in value of c from (b).

Question 7 continued

$$1000 + \frac{49\,347}{\sqrt{n}} - 1020 = \frac{-77.274}{\sqrt{n}}$$

$$20 = \frac{126\,621}{\sqrt{n}} \quad (1)$$

$$\sqrt{n} = \frac{126\,621}{20}$$

$$n = 40 \quad (1) \leftarrow n \text{ must be an integer}$$

$$C = 1000 + \frac{49\,347}{\sqrt{40}}$$

$$C = 1007.8 \quad (1)$$

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 7 continued

Question 7 continued

(Total for Question 7 is 11 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