Please check the examination details below before entering your candidate information

Candidate surname	Other names
MODEL SOLUTIONS	

Centre Number | Candidate Number

Pearson Edexcel Level 3 GCE

Friday 16 June 2023

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 allowed 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 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 paper. The total mark for this part of the examination 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

P74080A

N:1/1/1/1/

Pearson

1. The discrete random variable X has probability distribution

x	-2	-1	0	1	3
$P(X = x)$	0.25	a	b	a	0.30

where a and b are probabilities.

(a) Find $E(X)$

(2)

Given that $\text{Var}(X) = 3.9$

(b) find the value of a and the value of b

(4)

The independent random variables X_1 and X_2 each have the same distribution as X

(c) Find $P(X_1 + X_2 > 3)$

(3)

a) $E[x] = -2(0.25) - 1(a) + 0(b) + 1(a) + 3(0.3)$ 1

$= 0.4$ 1

b) Recall that $\text{Var}(x) = E[x^2] - (E[x])^2$

$E[x^2] = (-2)^2(0.25) + (-1)^2(a) + 0^2(b) + 1^2(a) + 3^2(0.3)$ 1

$\Rightarrow E[x^2] = 3.7 + 2a$

$\Rightarrow \text{Var}(x) = 3.7 + 2a - (0.4)^2$ 1

$\Rightarrow 3.9 = 2a + 3.54$

$\Rightarrow a = 0.18$ 1

The sum of probabilities always add up to 1.

$0.25 + a + b + a + 0.3 = 1$

$\Rightarrow 2a + b = 0.45$

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

$\Rightarrow \quad 0.36 + b = 0.45$

$\Rightarrow \quad b = 0.09$ (1)

c) The list of combinations are 1+3, 3+1 or 3+3. (1)

$Pr(X_1 = 1 \wedge X_2 = 3) = a(0.3)$

$Pr(X_1 = 3 \wedge X_2 = 1) = 0.3(a)$

$Pr(X_1 = 3 \wedge X_2 = 3) = 0.3(0.3)$ (1)

If we add these together, we get

$0.3(0.18) + 0.3(0.18) + 0.3^2$

$= 0.198$ (1)

(Total for Question 1 is 9 marks)

2. Telephone calls arrive at a call centre randomly, at an average rate of 1.7 per minute. After the call centre was closed for a week, in a random sample of 10 minutes there were 25 calls to the call centre.

(a) Carry out a suitable test to determine whether or not there is evidence that the rate of calls arriving at the call centre has changed. Use a 5% level of significance and state your hypotheses clearly. **(4)**

Only 1.2% of the calls to the call centre last longer than 8 minutes.

One day Tiang has 70 calls.

(b) Find the probability that out of these 70 calls Tiang has more than 2 calls lasting longer than 8 minutes. **(3)**

The call centre records show that 95% of days have at least one call lasting longer than 30 minutes.
On Wednesday 900 calls arrived at the call centre and none of them lasted longer than 30 minutes.

(c) Use a Poisson approximation to estimate the proportion of calls arriving at the call centre that last longer than 30 minutes. **(4)**

a) $H_0: \lambda = 1.7$

This is 10×1.7 because we have a random sample of 10 minutes.

$H_1: \lambda \neq 1.7$ 1

Assume that $X \sim Po(17)$ 1

$Pr(X \geqslant 25) = 1 - Pr(X \leqslant 24)$

$= 0.04064$ 1 using a calculator

We have a two tailed test so we compare this with 2.5%:

$0.04064 > 0.025$

Do not reject H_0 as there is insufficient evidence of a change in rate of calls. 1

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

Question 2 continued

b) We have a number of trials and we have a probability so we use a binomial distribution.

Let $Y \sim B(70, 0.012)$ ①

$Pr(Y > 2) = 1 - Pr(Y \leq 2)$ ①

$= 0.0523$ ① using a calculator.

c) $C \sim B(900, p)$

So $C \approx \sim Po(900p)$ ①

Recall the pdf for poisson is

$Pr(X = x) = \dfrac{\lambda^x e^{-\lambda}}{x!}$

$Pr(C = 0) = 0.05$

$\Rightarrow e^{-900p} = 0.05$ ①

$\Rightarrow -900p = \ln(0.05)$ ①

$\Rightarrow p = \dfrac{-\ln(0.05)}{900}$

$\Rightarrow p = 0.00333$ ① approximately

3. In a class experiment, each day for 170 days, a child is chosen at random and spins a large cardboard coin 5 times and the number of heads is recorded.
The results are summarised in the following table.

Number of heads	0	1	2	3	4	5
Frequency	3	10	45	62	38	12

Marcus believes that a B(5, 0.5) distribution can be used to model these data and he calculates expected frequencies, to 2 decimal places, as follows

Number of heads	0	1	2	3	4	5
Expected frequency	r	26.56	s	s	26.56	r

(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 B(5, 0.5) distribution is a good model for these data.
You should state clearly your hypotheses, the test statistic and the critical value used.

(6)

Nima believes that a better model for these data would be B(5, p)

(c) Find a suitable estimate for p

(1)

To test her model, Nima uses this value of p, to calculate expected frequencies as follows

Number of heads	0	1	2	3	4	5
Expected frequency	2.07	14.65	41.44	58.63	41.47	11.74

The test statistic for Nima's test is 1.62 (to 3 significant figures)

(d) State,
(i) giving your reasons, the degrees of freedom
(ii) the critical value
that Nima should use for a test at the 5% significance level.

(3)

(e) With reference to Marcus' and Nima's test results, comment on
(i) the probability of the coin landing on heads,
(ii) the independence of the spins of the coin.
Give reasons for your answers.

(2)

DO NOT WRITE IN THIS AREA

Question 3 continued

a) Let $X \sim Bin(5, 0.5)$

$Pr(X=0) = 0.03125$ 1

$r = 170 \times 0.03125 = 5.3125$ 1 As the table is for 170 days.

$Pr(X=2) = 0.3125$

$s = 170 \times 0.3125 = 53.125$ 1

b) H_0: $Bin(5, 0.5)$ is a suitable model.

H_1: $Bin(5, 0.5)$ is not a suitable model. 1

Recall that our test statistic

$$X^2 = \sum \frac{O_i^2}{E_i} - N$$

We can use this formula as none of our E_i values are less than 5.

$$X^2 = \frac{3^2}{r} + \frac{10^2}{26.56} + \frac{45^2}{s} + \frac{62^2}{s} + \frac{38^2}{26.56} + \frac{12^2}{r} - 170$$ 1

$\Rightarrow X = 27.4$ 1

$n = 6$ so we have 5 degrees of freedom. 1

Compare this with the X^2 table and we have a critical value of 11.07. 1

$27.4 > 11.07$

Question 3 continued

So we reject the null hypothesis as there is evidence that Marcus' model is not a good fit. (1)

c) Recall that if $X \sim Bin(n,p)$, $E[X] = np$

and $E[X] = \sum x \Pr(X = x)$

$$5p = \left(0 \times \frac{3}{170}\right) + \left(1 \times \frac{1}{17}\right) + \left(2 \times \frac{45}{170}\right) + \left(3 \times \frac{62}{170}\right) +$$

$$\left(4 \times \frac{38}{170}\right) + \left(5 \times \frac{12}{170}\right) = \frac{498}{170}$$

$\Rightarrow p = 0.586$ (1)

d) i) Because the expected frequency of '0' is less than 5, we have to collapse the first two columns to make a '0 or 1' column with an expected frequency of 16.72. (1)

We also have an estimator for p, so this is another constraint.

So $5 - 2 = 3$. (1)

ii) We again compare with the χ^2 distribution, and we have a critical value of 7.815. (1)

e) i) Nina's model is a very good fit as the test statistic was a lot lower than the critical value. So the probability of a head is approximately 0.6. (1)

Question 3 continued

ii) Nima's test suggests binomial is a good model and therefore independence of spins is a reasonable assumption. 1

(Total for Question 3 is 15 marks)

4. There are 32 students in a class.

Each student rolls a fair die repeatedly, stopping when their total number of sixes is 4
Each student records the total number of times they rolled the die.

Estimate the probability that the mean number of rolls for the class is less than 27.2

(6)

This is a negative binomial.

$X \sim \text{NBin}(4, 1/6)$ ①

From the formula booklet, we have that

$E[X] = \frac{r}{p}$, $\text{Var}(X) = \frac{r(1-p)}{p^2}$

①

$\mu = \frac{4}{(1/6)} = 24$, $\sigma^2 = \frac{4(1-1/6)}{(1/6)^2} = 120$ ①

By the Central Limit Theorem,

$\bar{X} \sim N(\mu, \sigma^2/n)$ ①

$\Rightarrow \bar{X} \sim N(24, 120/32)$

$\Rightarrow \bar{X} \sim N(24, 3.75)$ ①

$\Pr(\bar{X} < 27.2) = 0.951$ ① using calculator.

5. A machine fills cartons with juice.

The amount of juice in a carton is normally distributed with mean μ ml and standard deviation 8 ml.

A manager wants to test whether or not the amount of juice in the cartons, X ml, is less than 330 ml. The manager takes a random sample of 25 cartons of juice and calculates the mean amount of juice $\bar{x}$ ml.

(a) Using a 5% level of significance, find the critical region of $\bar{X}$ for this test. State your hypotheses clearly.

(4)

The Director is concerned about the machine filling the cartons with more than 330 ml of juice as well as less than 330 ml of juice. The Director takes a sample of 55 cartons, records the mean amount of juice $\bar{y}$ ml and uses a test with a critical region of

$$\{\bar{Y} < 328\} \cup \{\bar{Y} > 332\}$$

(b) Find P(Type I error) for the Director's test.

(2)

When $\mu = 325$ ml

(c) find P(Type II error) for the test in part (a)

(2)

a) $H_0: \mu = 330$
$H_1: \mu < 330$ 1

$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ by the Central Limit Theorem.

$\bar{X} \sim N(330, 64/25)$ 1

$\Rightarrow \bar{X} \sim N(330, 2.56)$

$Pr(\bar{X} < z) = 0.05$

$\Rightarrow z = 327.368$ 1 Using a calculator (Inverse Normal)

$\Rightarrow \bar{X} < 327$ 1

b) Reject Ho when Ho is true.

$\bar{Y} \sim N\left(330, \frac{8^2}{55}\right)$

$Pr(\bar{Y} < 328) + Pr(\bar{Y} < 332)$

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

$= 2\Pr(\bar{y} < 328)$ ①

$= 0.0637$ ① using a calculator

c) Fail to reject Ho when Ho is false.

$\bar{X} \sim N\left(325, \frac{8^2}{25}\right)$

$\Pr(\bar{x} > 327.368 \mid \mu = 325) = 1 - \Pr(\bar{x} < 327.368)$ ①

$= 0.0694$ ①

6. The discrete random variable X has probability generating function

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

(a) Specify the distribution of X

(2)

A fair die is rolled repeatedly.

(b) Describe an outcome that could be modelled by the random variable X

(1)

(c) Use calculus and $G_X(t)$ to find

(i) $E(X)$

(ii) $Var(X)$

(7)

The discrete random variable Y has probability generating function

$$G_Y(t) = \frac{t^{10}}{(3-2t^3)^2}$$

(d) Find the exact value of $P(Y = 19)$

(3)

a) Look at the 'P.G.F' column in the formula booklet.

$G_X(t)$ is of the form $\left(\frac{pt}{1-(1-p)t}\right)^r$

where $r = 2$.

$X \sim NegBin(2, p)$ 1

$$\frac{pt}{1-(1-p)-} = \frac{t}{3-2t}$$

$$\Rightarrow \frac{3pt}{3-3(1-p)t} = \frac{t}{3-2t}$$

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

Question 6 continued

So $3p = 1$ and $3(1-p) = 2$

$\Rightarrow p = 1/3$

$X \sim \text{NegBin}(2, 1/3)$ (1)

b) i) Model number of rolls of a fair die to achieve 5 or 6 twice! (1)

c) i) Recall that

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

$G_X(t) = t^2(3-2t)^{-2}$ (1)

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

$\Rightarrow G'_X(1) = 2(3-2) + 4(3-2) = 6$

by Product Rule

Hence $E[X] = 6$ (1)

ii) $G''_X(t) = 2(3-2t)^{-2} + 8t(3-2t)^{-3} + 8t(3-2t)^{-3} + 24t^2(3-2t)^{-4}$ (1)

$\Rightarrow G''_X(1) = 2(3-2) + 8(1)(3-2) + 8(1)(3-2) + 24(1)(3-2)$

$= 42$ (1)

(1)

$\Rightarrow Var(X) = 42 + 6 - 6^2 = 12$ (1)

d) $G_Y(t) = t^{10}(3 - 2t^3)^{-2}$

$= t^{10}\left(3\left(1 - \frac{2}{3}t^3\right)\right)^{-2}$ (1)

Question 6 continued

$$= \frac{1}{9}t^{10}\left(1 - \frac{2}{3}t^3\right)^{-2}$$

$$= \frac{1}{9}t^{10}\left(1 + at^3 + bt^6 + \frac{(-2)(-3)(-4)}{6}\left(-\frac{2}{3}t^3\right)^3 + \dots\right) \quad 1$$

Here we have used a binomial expansion to find the t^{19} coefficient.

So the coefficient of $t^{19} = \frac{1}{9} \times \frac{32}{27}$

So $P_r(Y = 19) = \frac{32}{243}$ 1

7. Each time a spinner is spun, the probability that it lands on red is 0.2

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

(i) for the 1st time on the 4th spin Fail, Fail, Fail, Happens **(2)**

(ii) for the 3rd time on the 8th spin **(2)**

(iii) exactly 4 times during 10 spins **(2)**

Each time the spinner is spun, the probability that it lands on yellow is 0.4

In a game with this spinner, a player must choose one of two events

R is the event that the spinner lands on **red** for the **1st** time in at most 4 spins

Y is the event that the spinner lands on **yellow** for the **3rd** time in at most 7 spins

(b) Showing your calculations clearly, determine which of these events has the greater probability. **(7)**

a) i) $0.8 \times 0.8 \times 0.8 \times 0.2 = 0.1024$ 1 1

ii) Need to choose 2 times out of the first 7 times where it happens.

$\binom{7}{2} \times 0.2^2 \times 0.8^5 = 0.2755$ 1

Then it happens on the 8th time

$0.2755 \times 0.2 = 0.0511$ 1

iii) Let $X \sim Bin(10, 0.2)$ 1

$Pr(X = 4) = 0.0881$ 1

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

Question 7 continued

b) Let $X \sim Geom(0.2)$ (1)

$Pr(R) = Pr(X \leq 4)$

$= 1 - Pr(X > 4)$

$= 1 - 0.8^4$ (1) $= 0.5904$ (1)

Let $N \sim NegBin(3, 0.4)$ (1)

$Pr(Y) = Pr(N \leq 7)$ Using the formula booklet,

(1) $= 0.4^3 + \binom{3}{2} 0.4^3 \times 0.6 + \binom{4}{2} 0.4^3 \times 0.6^2 +$

$\binom{5}{2} 0.4^3 \times 0.6^3 + \binom{6}{2} 0.4^3 \times 0.6^4 = 0.5801$ (1)

$Pr(R) = 0.5904 > 0.5801 = Pr(Y)$

So R has a greater probability. (1)