Please check the examination details below before entering your candidate information

Candidate surname	Other names

Pearson Edexcel Level 3 GCE

Centre Number

Candidate Number

Specimen Paper

(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, calculator

Total Marks

Candidates may use any calculator permitted by Pearson regulations. Calculators must not have the facility for algebraic 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.
- 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

1/1/1/1/1/

S 6 1 2 9 7 A 0 1 2 0

Pearson

Answer ALL questions. Write your answers in the spaces provided.

1. In a randomly selected week, a camera recorded the number of speeding drivers on a particular stretch of motorway.

	Sunday	**Monday**	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Saturday**
Frequency	35	30	28	24	40	51	37

Jeremy believes drivers are equally likely to be recorded speeding on any day of the week.

Carry out a hypothesis test, at the 5% level of significance, to see if the data support Jeremy's belief.

You should state your hypotheses, the degrees of freedom and the critical value used for this test.

(8)

(Total for Question 1 is 8 marks)

---

2. The discrete random variable Y has probability distribution given by

y	a	2	7
$\mathrm{P}(Y = y)$	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{5}$

where a is a constant.

Given that $\mathrm{Var}(Y) = 28$ and $\mathrm{E}(Y) < 0$

(*a*) find the value of a.

(6)

(*b*) Find $\mathrm{E}\left(\frac{1}{3-Y}\right)$

(2)

(Total for Question 2 is 8 marks)

---

1. Hypotheses

H_0: there is no association between day of the week and speeding

H_1: there is an association between day of the week and speeding B1

If we assume there is no association, the probability and hence expected value is same for all days.

total speeding recordings, $35+30+28+\ldots+37=245$

$\therefore E_i = \frac{245}{7} = 35$ M1A1

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

$$\chi^2 = \frac{35^2}{35} + \frac{30^2}{35} + \frac{28^2}{35} + \frac{24^2}{35} + \frac{40^2}{35} + \frac{51^2}{35} + \frac{37^2}{35} - 245$$ M1

$= 13.714$ A1

Degrees of Freedom:

DoF $= 7 - 1 = 6$ B1

Get critical value from tables

$\chi^2_6(5\%) = 12.592 < 13.714$ $\therefore$ falls in the critical region, sufficient evidence to reject H_0 and that there is an association between day of the week and speeding. B1A1

2. (a) 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 M1

y	a	2	7
y^2	a^2	4	49
$P(Y=y)$	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{5}$

Calculate $E(Y)$ and $E(Y^2)$

$E(Y) = \frac{1}{2}a + 2 \times \frac{3}{10} + 7 \times \frac{1}{5}$ B1

$= \frac{1}{2}a + 2$

$E(Y^2) = \frac{1}{2}a^2 + 4 \times \frac{3}{10} + 49 \times \frac{1}{5}$ B1

$= \frac{1}{2}a^2 + 11$

Substitute $E(Y)$ and $E(Y^2)$ into $Var(Y) = 28$

$Var(Y) = \frac{1}{2}a^2 + 11 - (\frac{1}{2}a + 2)^2 = 28$ M1

$\frac{1}{2}a^2 + 11 - \frac{1}{4}a^2 - 2a - 4 = 28$ expand bracket

M1 $\frac{1}{4}a^2 - 2a - 21 = 0$ solve quadratic, factorize

$a^2 - 8a - 84 = 0$

$(a+6)(a-14) = 0$

$a = -6$ $a = 14$

Since $E(Y) < 0$, $a = -6$ A1

(b)

from (a)

y	-6	2	7
$\frac{1}{3-y}$	$\frac{1}{3--6} = \frac{1}{9}$	$\frac{1}{3-2} = 1$	$\frac{1}{3-7} = -\frac{1}{4}$
$P(Y=y)$	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{5}$

$E\left(\frac{1}{3-Y}\right) = \frac{1}{9} \times \frac{1}{2} + 1 \times \frac{3}{10} + -\frac{1}{4} \times \frac{1}{5}$ M1

$= \frac{1}{18} + \frac{3}{10} - \frac{1}{20}$

$= \frac{11}{36}$ A1

3. Tim and Sue are each typing a manuscript and they make errors at random.

On average, Tim makes 0.45 errors per page and Sue makes 0.2 errors per page.

Rate ∴ Poisson

A random sample of n pages of Tim's typing is taken. The probability that these n pages contain no errors is less than 0.05

(*a*) Find the smallest possible value of n.

(3)

The random variable X represents the total number of errors in a random sample of 5 pages of Tim's typing and 5 pages of Sue's typing.

(*b*) Find $P(X = 2)$, stating a necessary assumption.

(3)

Random samples, each consisting of 5 pages of Tim's typing and 5 pages of Sue's typing, are selected.

(*c*) Find the probability that in 10 of these samples there are at least 2 with no errors.

(4)

(Total for Question 3 is 10 marks)

3. (a) $T \to$ # of errors Tim makes in n pages define variable

$T \sim Po(0.45n)$ M1

Formula for Poisson Distribution:

$P(X=x) = \frac{e^{-\lambda} \times \lambda^x}{x!}$

We want $P(T=0) < 0.05$.

$P(T=0) = \frac{e^{-0.45n} \times \cancel{0.45n^0}^{1}}{\cancel{0!}^{1}}$

$= e^{-0.45n} < 0.05$ M1

Solve: $e^{-0.45n} < 0.05$

$-0.45n \cancel{\ln e}^{1} < \ln 0.05$

$n > \frac{\ln 0.05}{-0.45}$ flip the inequality sign as we divided by a negative

$n > 6.657$

$\therefore n > 7$ A1

(b) $X \to$ total errors in 5 Tim and 5 Sue pages define variable

$T \sim Po(0.45)$ and $S \sim Po(0.2)$ Individual distributions for 1 page

Assuming that Tim's and Sue's errors happen independently: B1

$X = T + S$, $X \sim Po(5 \times 0.45 + 5 \times 0.2) \to X \sim Po(3.25)$ M1

$P(X=2) = 0.20478 \to 0.205$ A1

(c) Use $X \sim Po(3.25)$ to model the # of errors

$P(X=0) = 0.03877$ M1

$Y \to$ # of samples with no errors define variable

$Y \sim B(10, 0.03877)$ M1

$P(Y \geq 2) = 1 - P(Y \leq 0)$ M1

$= 0.055$ A1

4. The discrete random variable $X \sim B(n, p)$ has probability generating function given by

$$G_X(t) = (0.4 + 0.6t)^2$$

(*a*) Write down the value of n and the value of p

(2)

Using the probability generating function, find

(*b*) (i) $P(X = 1)$

(2)

(ii) $E(X)$

(3)

Two independent observations, X_1 and X_2, are taken from the distribution of X
The random variable $Y = X_1 + X_2$

(*c*) Use calculus to show that $E(Y^2) = 6.72$

(7)

(Total for Question 4 is 14 marks)

4.(a) Formula for the PGF of a Binomial Distribution:

For $X \sim B(n,p)$: $G_X(t) = (1-p+pt)^n$

$G_X(t) = (0.4 + \underbrace{0.6}_{p}t)^{\overbrace{2}^{n}}$

$\therefore$ $p = 0.6$ and $n = 2$ B1 B1

(b) i. $P(X=1)$ is the coefficient of t^1 M1

$G_X(t) = (0.4 + 0.6t)^2$ expand

$= 0.16 + 0.48t + 0.36t^2$

$\therefore$ $P(X=1) = 0.48$ A1

ii. For PGFs: $E(X) = G'_X(1)$ M1

We need to differentiate $G_X(t)$:

$G_X(t) = 0.16 + 0.48t + 0.36t^2$

$G'_X(t) = 0.48 + 0.72t$ M1

$G'_X(1) = 0.48 + 0.72 = 1.2$

$\therefore$ $E(X) = 1.2$ A1

(c) For $Y = X_1 + X_2$, $G_Y(t) = G_X(t) \times G_X(t)$ B1

$G_Y(t) = ((0.4+0.6t)^2)^2 = (0.4+0.6t)^4$ B1

We know that $Var(X) = E(X^2) - [E(X)]^2$.

For PGFs:

$E(X) = G'_X(1)$

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

We need to differentiate:

$G'_Y(t) = 4(0.4+0.6t)^3 \times \underbrace{0.6}$ use chain rule! M1

$= 2.4(0.4+0.6t)^3$

$G'_Y(1) = 2.4(0.4+0.6)^3 = 2.4 \times 1 = 2.4$ A1

$G''_Y(t) = 3 \times 2.4(0.4+0.6t)^2 \times \underbrace{0.6}$ use chain rule! M1

$= 4.32(0.4+0.6t)^2$

$G''_Y(1) = 4.32(0.4+0.6)^2 = 4.32 \times 1 = 4.32$ A1

$\therefore$ $G''_Y(1) + G'_Y(1) - \cancel{[G'_Y(1)]^2} = E(Y^2) - \cancel{[G'_Y(1)]^2}$

$G''_Y(1) + G'_Y(1) = E(Y^2)$ M1

Substitute: $E(Y^2) = 4.32 + 2.4 = 6.72$

$E(Y^2) = 6.72$ A1

5. A company claims that the proportion of visitors to its website who make a purchase is 0.03

Nina believes that the proportion of visitors to the website who make a purchase is less than 0.03 and asks the company for some data to test this.

Geometric Distribution

The company tells Nina, for a particular day, the number of visits to its website until a purchase is made. Nina assumes that visits to the website are made independently and that each visitor has the same probability of making a purchase.

(*a*) Using a 5% level of significance, find the critical region for this test.
State your hypotheses clearly.

(5)

(*b*) Find the actual level of significance for this test.

(2)

The 94th visitor to the website was the first person to make a purchase.

(*c*) Test, at the 5% level of significance, whether or not there is evidence that the proportion of visitors who make a purchase is less than 0.03.

(2)

(Total for Question 5 is 9 marks)

5. (a) Hypotheses

$H_0: p = 0.03$

$H_1: p < 0.03$ B1

$X \to$ # of visitors on the website until the first purchase   define variable

$X \sim Geo(0.03)$ M1

$P(X \geq c) < 0.05$ we want $\geq c$ since for geometric the larger p is, the smaller the # of trials needed until a success occurs!

$P(X \geq c) = (1-0.03)^{c-1} < 0.05$ Solve for c M1

$(c-1)\log 0.97 < \log 0.05$

M1 $c-1 > \frac{\log 0.05}{\log 0.97}$ flip inequality sign as we divided by $\log 0.97$ and it's negative

$c > \frac{\log 0.05}{\log 0.97} + 1 \rightarrow c > 99.35$

$\therefore$ CR is $X \geq 100$ A1

(b) Use $X \sim Geo(0.03)$

$P(X \geq 100) = 0.97^{0.99}$ M1

$= 0.0490$ to 3sf A1

(c) Since the critical region is $X \geq 100$, 94 does not fall in the critical region. Insufficient evidence to reject H_0, insufficient evidence to suggest that the proportion of visitors making a purchase is less than 0.03. M1A1

6. The weights of bars of soap are known to be normally distributed.

The standard deviation of the weights is 3 grams. σ

A label on the bars of soap states that the mean weight is 120 grams. μ

Gizel believes that the mean weight of the bars of soap is greater than 120 grams.
She takes a random sample of 10 bars of soap and finds the mean weight of her sample.

Gizel then tests, at the 5% level of significance, whether there is evidence that the bars of soap weigh more than 120 grams.

(*a*) Write down the probability of a Type I error.

(1)

Given that the true mean weight of the bars of soap is 122 grams,

(*b*) show that the power of Gizel's test is 0.68 to 2 significant figures.

(5)

Alex decides to carry out the same test at the 1% level of significance.

(*c*) Without carrying out the test, compare the power of Alex's test with the power of Gizel's test.
Give a reason for your answer.

(2)

Joseph decides to increase the sample size and carry out the test at the 5% level of significance.

Assuming that the true mean weight of the bars of soap is 122 grams,

(*d*) calculate the smallest sample size that will produce a power of at least 0.9

(5)

(*e*) State which of Gizel's and Joseph's tests should be used.
Give a reason for your answer.

(2)

(Total for Question 6 is 15 marks)

6. (a) P(Type I error) = 0.05 B1

(b) Hypotheses

$H_0: \mu = 120$

$H_1: \mu > 120$

$X \sim N(120, 3) \rightarrow \bar{X} \approx N(120, \frac{3^2}{10})$ M1

We need to find the critical region at 5% significance

Convert to Standard Normal

$X \sim N(\mu, \sigma^2) \rightarrow Z \sim N(0,1)$

$P(X = x),\ x \rightarrow \frac{x-\mu}{\sigma} = z$

$P(\bar{X} > c) < 0.05$

$\hookrightarrow P(\bar{x} < c) > 0.95$ — code to standard normal → $\frac{c-120}{\frac{3}{\sqrt{10}}} > \text{InvN}(0.95)$ M1

$c - 120 > 1.64485 \times \frac{3}{\sqrt{10}}$, $c > 121.56$

$\bar{X} > 121.56$ critical region

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

$P(\bar{X} > 121.56 \mid \mu = 122) = 0.6786 \rightarrow$ 0.68 to 2 sf A1M1A1

(c) Alex's test's Power is smaller than that of Gizele's test, since the null hypothesis is less likely to be rejected. B1B1

(d) $\bar{X} \approx \sim N(120, \frac{3}{\sqrt{n}})$

We need to find the critical region at 5% significance for his test

convert to standard normal again:

$$\frac{c-120}{\frac{3}{\sqrt{n}}} > 1.6449 \quad \text{M1}$$

$$c > 1.6449 \times \frac{3}{\sqrt{n}} + 120 \quad \text{A1}$$

Now to calculate power:

$P(\bar{x} > c \mid \mu = 122) > 0.9$

$\hookrightarrow P(\bar{x} < c \mid \mu = 122) < 0.1 \rightarrow$ convert to standard normal:

$$\frac{c-122}{\frac{3}{\sqrt{n}}} < -1.2816$$ substitute in c from above:

$$\frac{(1.6449 \times \frac{3}{\sqrt{n}} + 120) - 122}{\frac{3}{\sqrt{n}}} < -1.2816 \quad \text{M1}$$

$$1.6449 \times \frac{3}{\sqrt{n}} - 2 < -1.2816 \times \frac{3}{\sqrt{n}}$$

$$2.9265 \times \frac{3}{\sqrt{n}} < 2$$

$$\frac{8.7795}{2} < \sqrt{n}$$

$$\sqrt{n} > 4.38975 \longrightarrow n > 19.26\ldots \Rightarrow n = 20 \quad \text{M1A1}$$

(e) Both have the same Type I Error (is equal to significance level in continuous distributions)

Joseph's test has higher power ∴ Joseph's test is better M1A1

7. A radio station is running a contest each day for 30 days.

Each day it awards a prize to each of the first 12 callers who phone in and answer a question correctly. Once 12 prizes are awarded, no more calls are taken that day.

It can be assumed that each caller has a $\frac{3}{4}$ chance of answering a question correctly, independently of all other callers.

(*a*) Find the probability that there are exactly 15 calls taken on the first day.

(2)

(*b*) Find the probability that there are more than 13 calls taken on the first day.

(3)

(*c*) Estimate the probability that the mean number of calls taken over the 30 days is more than 15.5

(6)

(Total for Question 7 is 11 marks)

TOTAL FOR PAPER IS 75 MARKs

7. This question is about the Negative Binomial distribution

(a) $X \rightarrow$ # of callers to achieve 12 correct answers

$X \sim NB(12, \frac{3}{4})$

$P(X=15) = \binom{15-1}{12-1} \times \left(\frac{3}{4}\right)^{12} \times \left(1-\frac{3}{4}\right)^{15-12}$

$= \binom{14}{11} \times \left(\frac{3}{4}\right)^{12} \times \left(\frac{1}{4}\right)^{3}$ M1

$= \boxed{0.180 \text{ to 3sf}}$ A1

(b) $P(X>13) = 1 - (P(X=12)+P(X=13))$ we subtract the possibilities of 12 correct answers within 12 or 13 calls B1

$= 1 - \left[\binom{11}{11} \times \left(\frac{3}{4}\right)^{12} \times \left(\frac{1}{4}\right)^{0} + \binom{12}{11} \times \left(\frac{3}{4}\right)^{12} \times \left(\frac{1}{4}\right)^{1}\right]$ M1

$= 1 - 0.163$

$= \boxed{0.873 \text{ to 3sf}}$ A1

(c) For Negative Binomial from Formula Booklet:

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

Substitute:

$E(X) = \frac{12}{\frac{3}{4}} = 16 = \mu$ M1

$Var(X) = \frac{12\left(\frac{1}{4}\right)}{\left(\frac{3}{4}\right)^2} = \frac{16}{3} = \sigma^2$ A1

So now we can use central limit theorem, where $\bar{X} \approx\sim N\left(\mu, \frac{\sigma^2}{n}\right)$

$\bar{X} \approx\sim N\left(16, \frac{\frac{16}{3}}{30}\right)$ enter $\sqrt{\frac{\frac{16}{3}}{30}}$ into your calculator M1A1

n=30

$P(\bar{X} > 15.5) = \boxed{0.883 \text{ to 3sf}}$ M1A1