Write your name here

| Surname | Other names |
|---|---|
| | |

**Pearson Edexcel GCE**

Centre Number | Candidate Number

# A level Further Mathematics
# Further Statistics 1
# Practice Paper 4

**You must have:**
Mathematical Formulae and Statistical Tables (Pink)

Total Marks

**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 the 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.
- 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.
- Calculators must not be used for questions marked with a * sign.

**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.
- If you change your mind about an answer, cross it out and put your new answer and any working underneath.

**1.** Jane shoots at a target until she hits it. The random variable J is the number of shots needed by Jane to hit the target.

(a) State a suitable distribution to model J.

**(1)**

(b) Given that the mean of J is 5, calculate the probability of Jane

(i) hitting the target for the first time on her 4th shot,

**(3)**

(ii) taking at least 3 shots to hit the target for the first time.

**(3)**

(c) State any assumptions you have made using this model.

**(2)**

**(Total 9 marks)**

[Mark scheme for Question 1]

[Examiner comment]

---

**2.** The discrete random variable $X$ has probability distribution given by

| $x$ | $-1$ | $0$ | $1$ | $2$ | $3$ |
|---|---|---|---|---|---|
| $P(X = x)$ | $\frac{1}{5}$ | $a$ | $\frac{1}{10}$ | $a$ | $\frac{1}{5}$ |

where $a$ is a constant.

(a) Find the value of $a$.

**(2)**

(b) Write down $E(X)$.

**(1)**

(c) Find $\text{Var}(X)$.

**(3)**

The random variable $Y = 6 - 2X$.

(d) Find $\text{Var}(Y)$.

**(2)**

(e) Calculate $P(X \geq Y)$.

**(3)**

**(Total 11 marks)**

[Mark scheme for Question 2]

[Examiner comment]

**3.** A robot is programmed to build cars on a production line. The robot breaks down at random at a rate of once every 20 hours.

(a) Find the probability that it will work continuously for 5 hours without a breakdown.

**(3)**

Find the probability that, in an 8 hour period,

(b) the robot will break down at least once,

**(3)**

(c) there are exactly 2 breakdowns.

**(2)**

In a particular 8 hour period, the robot broke down twice.

(d) Write down the probability that the robot will break down in the following 8 hour period. Give a reason for your answer.

**(2)**

**(Total 10 marks)**

**Mark scheme for Question 3**

**Examiner comment**

**4.** The probability of Richard winning a coconut in a game at the fair is 0.12.

Richard plays a number of games.

(a) Find

(i) the probability of Richard winning his second coconut on his 8th game,

**(2)**

(ii) the expected number of games Richard will need to play in order to win 3 coconuts.

**(1)**

(b) State two assumptions that you have made in part (*a*).

**(2)**

Mary plays the same game, but has a different probability of winning a coconut. She plays until she has won $r$ coconuts. The random variable $G$ represents the total number of games Mary plays.

(c) Given that the mean and standard deviation of $G$ are 18 and 6 respectively, determine whether Richard or Mary has the greater probability of winning a coconut in a game.

**(5)**

**(Total 10 marks)**

[Mark scheme for Question 4]

[Examiner comment]

**5.** A research station is doing some work on the germination of a new variety of genetically modified wheat.

They planted 120 rows containing 7 seeds in each row.

The number of seeds germinating in each row was recorded. The results are as follows

| Number of seeds germinating in each row | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| Observed number of rows | 2 | 6 | 11 | 19 | 25 | 32 | 16 | 9 |

(a) Write down two reasons why a binomial distribution may be a suitable model.

**(2)**

(b) Show that the probability of a randomly selected seed from this sample germinating is 0.6.

**(2)**

The research station used a binomial distribution with probability 0.6 of a seed germinating. The expected frequencies were calculated to 2 decimal places. The results are as follows:

| Number of seeds germinating in each row | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| Expected number of rows | 0.20 | 2.06 | $s$ | 23.22 | $t$ | 31.35 | 15.68 | 3.36 |

(c) Find the value of $s$ and the value of $t$.

**(2)**

(d) Stating your hypotheses clearly, test, at the 1% level of significance, whether or not the data can be modelled by a binomial distribution.

**(7)**

**(Total 13 marks)**

[Mark scheme for Question 5](#)

[Examiner comment](#)

**6.** The random variable $X$ has probability generating function

$$G_X(t) = k[t^3(2 + 3t) + (1 + t)^4],$$

where $k$ is a positive constant.

(a) Show that $k = \dfrac{1}{21}$.

**(2)**

Find

(b) $E(X)$,

**(3)**

(c) $Var(X)$,

**(4)**

(d) $P(X = 3)$.

**(2)**

**(Total 11 marks)**

[Mark scheme for Question 6](#)

[Examiner comment](#)

**7.** A poultry farm produces eggs which are sold in boxes of 6. The farmer believes that the proportion, $p$, of eggs that are cracked when they are packed in the boxes is approximately 5%. She decides to test the hypotheses

$$H_0: p = 0.05 \quad \text{against} \quad H_1: p > 0.05.$$

To test these hypotheses she randomly selects a box of eggs and rejects $H_0$ if the box contains 2 or more eggs that are cracked. If the box contains 1 egg that is cracked, she randomly selects a second box of eggs and rejects $H_0$ if it contains at least 1 egg that is cracked. If the first or the second box contains no cracked eggs, $H_0$ is immediately accepted and no further boxes are sampled.

(a) Show that the power function of this test is

$$1 - (1 - p)^6 - 6p(1 - p)^{11}.$$

**(3)**

(b) Calculate the size of this test.

**(2)**

Given that $p = 0.1$,

(c) find the expected number of eggs inspected each time this test is carried out, giving your answer correct to 3 significant figures,

**(3)**

(d) calculate the probability of a Type II error.

**(2)**

Given that $p = 0.1$ is an unacceptably high value for the farmer,

(e) use your answer from part (*d*) to comment on the farmer's test.

**(1)**

**(Total 11 marks)**

Mark scheme for Question 7

Examiner comment

**TOTAL FOR PAPER: 75 MARKS**

**Mark scheme for Question 1** (Examiner comment) (Return to Question 1)

| Question | Scheme | Marks |
|---|---|---|
| **1(a)** | Geometric | **B1** |
| | | **(1)** |
| **(b)(i)** | $p = \frac{1}{5}$ | **B1** |
| | $\therefore\ \mathrm{P}(J = 4) = \left(\frac{4}{3}\right)^3 \left(\frac{1}{5}\right) = \frac{64}{625} = 0.1024$ | **M1A1** |
| | | **(3)** |
| **(ii)** | $\mathrm{P}(J \geqslant 3) = (1-p)^2 = \left(\frac{4}{5}\right)^5 = \frac{16}{25} = 0.64$ | **M1A1** <br> **A1** |
| | | **(3)** |
| **(c)** | Assume **shots** are **independent** and | **B1** |
| | **probability** of a **hit** is **constant**. | **B1** |
| | | **(2)** |
| | | **(9 marks)** |

**Mark scheme for Question 2** (Examiner comment) (Return to Question 2)

| Question | Scheme | Marks |
|---|---|---|
| **2(a)** | $2a+\frac{2}{5}+\frac{1}{10}=1$ | **M1** |
| | $a=\frac{1}{4}$ **or 0.25** | **A1** |
| | | **(2)** |
| **(b)** | $E(X)=\mathbf{1}$ | **B1** |
| | | **(1)** |
| **(c)** | $E(X^2)=1\times\frac{1}{5}+1\times\frac{1}{10}+4\times\frac{1}{4}+9\times\frac{1}{5}$ $(=3.1)$ | **M1** |
| | $\text{Var}(X)=3.1-1^2$, **= 2.1 or** $\frac{21}{10}$ **oe** | **M1A1** |
| | | **(3)** |
| **(d)** | $\text{Var}(Y)=(-2)^2\text{Var}(X)$, $=$ **8.4 or** $\frac{42}{5}$ **oe** | **M1A1** |
| | | **(2)** |
| **(e)** | $X\geq Y$ when $X=3$ or 2, so probability $=$ "$\frac{1}{4}$"$+\frac{1}{5}$ | **M1** **A1ft** |
| | $=\frac{9}{20}$ **oe** | **A1** |
| | | **(3)** |
| | | **(7 marks)** |

Mark scheme for Question 3 (Examiner comment) (Return to Question 3)

| Question | Scheme | Marks |
|---|---|---|
| 3(a) | $Y \sim \text{Po}(0.25)$ | B1 |
| | P($Y$=0) $= e^{-0.25}$ | M1 |
| | $= 0.7788$ | A1 |
| | | (3) |
| (b) | $X \sim \text{Po}(0.4)$ | B1 |
| | P(Robot will break down) $= 1 - \text{P}(X = 0)$ | |
| | $= 1 - e^{-0.4}$ | M1 |
| | $= 1 - 0.067032$ | |
| | $= 0.3297$ | A1 |
| | | (3) |
| (c) | $P(X = 2) = \dfrac{e^{-0.4}(0.4)^2}{2}$ | M1 |
| | $= 0.0536$ | A1 |
| | | (2) |
| (d) | 0.3297 or answer to part (b) | B1ft |
| | as Poisson events are <u>independent</u> | B1dep |
| | | (2) |
| | | (10 marks) |

Mark scheme for Question 4 (Examiner comment) (Return to Question 4)

| Question | Scheme | Marks |
|---|---|---|
| 4(a)(i) | $\binom{7}{1}(0.12)^2(0.88)^6 = 0.0468$ | M1A1 |
| | | (2) |
| (ii) | $\frac{3}{0.12} = 25$ | B1 |
| | | (1) |
| (b) | Probability constant; games independent of each other | B1B1 |
| | | (2) |
| (c) | $\frac{r}{p} = 18;\quad \frac{r(1-p)}{p^2} = 36$ | B1B1 |
| | $18(1-p) = 36p$ Substitute $\frac{r}{p}$ | M1 |
| | $18 = 54p$ | A1 |
| | $p = \frac{1}{3}$; Mary | A1 A1ft |
| | | (5) |
| | | (10 marks) |

**Mark scheme for Question 5** (Examiner comment) (Return to Question 5)

| Question | Scheme | Marks |
|---|---|---|
| **5(a)** | The seeds are **independent** / There are a **fixed number** of seeds in a row / There are only **two outcomes** to the seed germinating – either it germinates or it does not / The **probability** of a seed germinating is **constant** | **B1B1** |
| | | **(2)** |
| **(b)** | $\dfrac{(0\times2)+(1\times6)+(2\times11)+(3\times19)+(4\times25)+(5\times32)+(6\times16)+(7\times9)}{120\times7}$ $=\dfrac{504}{840}$ | **M1** |
| | $= 0.6$ ** | **A1cso** |
| | | **(2)** |
| **(c)** | $p = 0.6 \quad q = 0.4$ | |
| | $s = 120\times\ 21q^5p^2 = 120\times21 \text{ x } 0.4^5 \text{ x } 0.6^2 = 9.29$ | **B1** |
| | $t = 120\times35q^3p^4 = 120\times35 \text{ x } 0.4^3 \text{ x } 0.6^4 = 34.84$ | **B1** |
| | | **(2)** |
| **(d)** | H0 : A binomial distribution is a suitable model. | |
| | $H_1$: A binomial distribution is not a suitable model. | **B1** |
| | (table below) | **M1** |

| | | | | | |
|---|---|---|---|---|---|
| Observed number of rows | 19 | 19 | 25 | 32 | 25 |
| Expected number of rows | 11.55 | 23.22 | 34.84 | 31.35 | 19.04 |
| $\dfrac{(O-E)^2}{E}$ | 4.81 | 0.77 | 2.78 | 0.013 | 1.87 |
| $\dfrac{O^2}{E}$ | 31.26 | 15.55 | 17.94 | 32.66 | 32.83 |

| Question | Scheme | Marks |
|---|---|---|
| | $\nu = 5 - 2 = 3$ | **B1ft** |
| | Critical value for $\chi^2 =\ 11.345$ | **B1ft** |
| | $\sum\dfrac{(O-E)^2}{E} = 10.23$ or $\sum\dfrac{O^2}{E} - N = 130.23 - 120 = 10.23$ | **M1A1** |
| | $10.23 < 11.345$ therefore do not reject $H_0$<br>A binomial is a suitable model. | **A1** |
| | | **(7)** |
| | | **(14 marks)** |

Mark scheme for Question 6 (Examiner comment) (Return to Question 6)

| Question | Scheme | | Marks |
|---|---|---|---|
| 6(a) | G(1)=1 $\Rightarrow$, k $[1\times5 + 2^4] = 1$<br>$k = \frac{1}{21}$ * | | M1A1 |
| | | | (2) |
| (b) | $E(X) = G'_x(1)$. $G'_x(t) = k[\ 6t^2 + 12t^3 + 4(1+t)^3]$ | $G'_x(t)$ | M1 |
| | $G'_x(1) = k[\ 6 + 12 + 4(8)]$ | $G'_x(t)$ | M1 |
| | $\therefore \quad E(X) = \frac{50}{21}$ o.e | | A1 |
| | | | (3) |
| (c) | $G'_x(t) = k[12t + 36t^2 + 12(1+6)^2]$ | | M1 |
| | $G''_x(1) = 96k$ | $G''_x(1)$ | M1 |
| | $Var(X) = G''_x(1) + G'_x(1) - G'_x(1)^2$ | | M1 |
| | $= \frac{96}{21} + \frac{50}{21} - \frac{2500}{21^2} = \frac{566}{441}$ or 1.28 | | A1 |
| | | | (4) |
| (d) | $P(X=3)$ = coefficient of $t^3$ in expansion of $G_x(t)$ | | |
| | $= k[2 + \binom{4}{3}], = \frac{6}{21} = \underline{\frac{2}{7}}$ | | M1A1 |
| | | | (2) |
| | | | (11 marks) |

**Mark scheme for Question 7** (Examiner comment) (Return to Question 7)

| Question | Scheme | Marks |
|---|---|---|
| 4(a) | Power function $= \mathrm{P}(\mathrm{H}_0 \text{ rejected}) = \mathrm{P}(X_1 \geq 2) + \mathrm{P}(X_1 = 1) \times \mathrm{P}(X_2 \geq 1)$ | |
| | $= 1-(1-p)^6 - 6p(1-p)^5 + 6p(1-p)^5 \times (1-(1-p)^6)$<br>$= 1-(1-p)^6 - 6p(1-p)^5 + 6p(1-p)^5 - 6p(1-p)^{11}$ | M1A1 |
| | $= 1-(1-p)^6 - 6p(1-p)^{11}$ | A1cso |
| | | (3) |
| (b) | Size of test is value of power function when $p = 0.05$ | |
| | Size of test $= 1 - 0.95^6 - 6 \times 0.05 \times 0.95^{11} = 0.094268\ldots$ (awrt 0.0943) | M1A1 |
| | | (2) |
| (c) | $\mathrm{E}[\text{number of eggs inspected}] = 12 \times \mathrm{P}(X_1 = 1) + 6 \times \mathrm{P}(X_1 \neq 1)$ | M1 |
| | $= 12 \times 6 \times 0.1 \times 0.9^5 + 6 \times (1 - (6 \times 0.1 \times 0.9^5))$ | A1 |
| | $= 8.1257\ldots$ (awrt 8.13) | A1 |
| | | (3) |
| (d) | $\mathrm{P}(\text{Type II error} \mid p = 0.1) = 1 - (\text{value of power function when } p = 0.1)$ | M1 |
| | $\mathrm{P}(\text{Type II error} \mid p = 0.1) = 1 - (1 - 0.9^6 - 6 \times 0.1 \times 0.9^{11}) = 0.7197\ldots$ (awrt 0.720) | A1 |
| | | (2) |
| (e) | Prob of Type II error, accepting $p = 0.05$ when it is actually 0.1, unacceptably high, is large, therefore not a good **test**. | B1 |
| | | (1) |
| | | **(11 marks)** |

**A level Further Mathematics – Further Statistics 1 – Practice Paper 04 – Examiner report –**

**Examiner comment for Question 1** [(Mark scheme)] [(Return to Question 1)]

1. This proved to be a good start for many candidates, with many correct answers seen to parts (a) and (b). However, responses often did not refer to the content of the question in part (c) and gained no credit as a result.

**Examiner comment for Question 2** [(Mark scheme)] [(Return to Question 2)]

2. Finding the correct value of $a$ in the first part of the question proved to be relatively straightforward for most candidates. Few errors were seen although some candidates provided very little in the way of working out and did not always make it explicit that they were using the fact that the sum of the probabilities equals one. Similarly, most candidates were able to obtain the correct value of $E(X)$, though not many deduced this fact by recognising the symmetry of the distribution.

   The majority opted to use the formula to calculate $E(X)$, which resulted in processing errors in some cases. Common errors seen in calculating Var ($X$) included forgetting to subtract $[E(X)]^2$ from $E(X^2)$ or calculating $E(X^2) - E(X)$, although on the whole the correct formula was successfully applied.

   Most candidates were able to correctly apply $\text{Var}(aX + b) = a^2 \text{Var}(X)$ to deduce $\text{Var}(Y) = 4 \text{Var}(X)$, although $\text{Var}(Y) = 6 - 2\text{Var}(X)$ was a typical error. Quite a number of candidates attempted to calculate $E(Y^2) - [E(Y)]^2$ with varying degrees of success. Occasionally, candidates divided their results in part (b), part (c) and part (d) by 5.

   The final part of the question proved to be the most challenging of all and was often either completely omitted or poorly attempted with little or no success. Only a minority of candidates knew they would need to equate $6 - 2X$ to $X$ in order to obtain the corresponding values of $X$ and of those who did, only a small number scored full marks, as candidates were generally unable to identify the correct values of $X$.

**Examiner comment for Question 3** [(Mark scheme)] [(Return to Question 3)]

3. Although there were a minority of candidates who were unable to identify the correct distribution to use the majority of candidates achieved full marks to parts (a) (b) and (c). Part (d) seemed to cause substantial difficulty. In part (a) the majority of candidates identified that a Poisson (rather than the Binomial) distribution was appropriate but some calculated the parameter as 2.5 or 4 rather than 0.4. A few used Po(1) and calculated $P(Y=5)$.

   In part (b) and part (c) the most common error was to use Po(2.5). The majority of candidates were able to work out $P(X > 1)$ and $P(X = 2)$ using the correct Poisson formula. Many thought that their answer to part (c) was the correct solution while others used or multiplied their answers to both (b) and (c). Whether stating a correct or incorrect solution only a minority used the statistical term "independence" as the reason for their answer.

**Examiner comment for Question 4** (Mark scheme) (Return to Question 4)

4. This question was well done with many candidates gaining eight of the 10 marks available. In part (b) many of the candidates did not use the context of the question in giving their assumptions.

**Examiner comment for Question 5** (Mark scheme) (Return to Question 5)

5. Many students failed to give Binomial conditions in context in part (a) but parts (b) and (c) were done well. Many students still wanted to include the probability in their hypotheses but most then went on to perform the test satisfactorily and arrived at the correct conclusion. There were very few students who failed to combine some classes, usually correctly. Also, most students managed to arrive at the correct degrees of freedom. Too many students lost marks at the end of this question by failing to show adequate working when their test statistic lacked accuracy.

**Examiner comment for Question 6** (Mark scheme) (Return to Question 6)

6. This question was a good source of marks for many candidates. Attempts were well presented and many solutions were clear and accurate.

**Examiner comment for Question 7** (Mark scheme) (Return to Question 7)

7. A minority of candidates tried to fiddle their answers in part (a). Candidates who wrote down $\mathrm{P}(X_1 \geq 2) + \mathrm{P}(X_1 = 1) \times \mathrm{P}(X_2 \geq 1)$ or $1 - \left(\mathrm{P}(X_1 = 0) + \mathrm{P}(X_1 = 1) \times \mathrm{P}(X_2 = 0)\right)$ were usually able to proceed through the required steps to complete the proof accurately. Those who chose to start by writing the calculation in terms of $p$ were less successful, as they were unclear in their mind what they were trying to calculate.

In part (b) many candidates did not realise the connection to part (a) and started again. Even those who were unsuccessful in proving part (a) were able to get this part correct by starting from the beginning.
Part (c) proved to be quite a challenge to most candidates. The most common errors were to work out the number of boxes required rather than the number of eggs or using 0.95 instead of 0.9.

Part (d) was well answered with the majority of candidates gaining the right answer here even if they had struggled with the first 3 parts.

The answer to part (e) was answered eloquently by many candidates.