Write your name here

Surname	Other names

Pearson Edexcel GCE

Centre Number | Candidate Number

A level Further Mathematics
Further Statistics 1
Practice Paper 5

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. The discrete random variable X has the following probability distribution, where p and q are constants.

x	-2	-1	$\frac{1}{2}$	$\frac{3}{2}$	2
$\mathrm{P}(X = x)$	p	q	0.2	0.3	p

(a) Write down an equation in p and q.

(1)

Given that $\mathrm{E}(X) = 0.4$,

(b) find the value of q.

(3)

(c) Hence find the value of p.

(2)

Given also that $\mathrm{E}(X^2) = 2.275$,

(d) find $\mathrm{Var}(X)$.

(2)

(Total 8 marks)

[Mark scheme for Question 1]

[Examiner comment]

---

2. The police are carrying out a check on car tyres. The percentage of cars with tyre defects is assumed to be 40%. Let X represent the number of cars checked, up to and including the first one with tyre defects.

(a) State the distribution that could be used to model X and write down the mean of X.

(2)

The police decide that they are going to change the place where they are carrying out the check after they have found two cars with tyre defects. Let W represent the number of cars checked, up to and including the second one with tyre defects.

(b) Determine the mean and the variance of W.

(4)

(c) Calculate the probability that $W = 5$.

(2)

(Total 8 marks)

[Mark scheme for Question 2]

[Examiner comment]

---

3. A company receives telephone calls at random at a mean rate of 2.5 per hour.

(*a*) Find the probability that the company receives

(i) at least 4 telephone calls in the next hour,

(ii) exactly 3 telephone calls in the next 15 minutes.

(4)*

(*b*) Find, to the nearest minute, the maximum length of time the telephone can be left unattended so that the probability of missing a telephone call is less than 0.2.

(3)

The company puts an advert in the local newspaper. The number of telephone calls received in a randomly selected 2 hour period after the paper is published is 10.

(*c*) Test at the 5% level of significance whether or not the mean rate of telephone calls has increased. State your hypotheses clearly.

(5)

(Total 12 marks)

Mark scheme for Question 3

Examiner comment

---

*Part 3(a)(ii) would be 2 marks in the new specification and 3 marks in the old specification.

4. A factory manufactures batches of an electronic component. Each component is manufactured in one of three shifts. A component may have one of two types of defect, D_1 or D_2, at the end of the manufacturing process. A production manager believes that the type of defect is dependent upon the shift that manufactured the component. He examines 200 randomly selected defective components and classifies them by defect type and shift.

The results are shown in the table below.

Shift \ Defect type	D_1	D_2
First shift	45	18
Second shift	55	20
Third shift	50	12

Stating your hypotheses, test, at the 10% level of significance, whether or not there is evidence to support the manager's belief. Show your working clearly.

(10)

(Total 10 marks)

Mark scheme for Question 4

Examiner comment

5. A drug is claimed to produce a cure to a certain disease in 35% of people who have the disease. To test this claim a sample of 20 people having this disease is chosen at random and given the drug. If the number of people cured is between 4 and 10 inclusive the claim will be accepted. Otherwise the claim will not be accepted.

(a) Write down suitable hypotheses to carry out this test.

(2)

(b) Find the probability of making a Type I error.

(3)

The table below gives the value of the probability of the Type II error, to 4 decimal places, for different values of p where p is the probability of the drug curing a person with the disease.

P(cure)	0.2	0.3	0.4	0.5
P(Type II error)	0.5880	r	0.8565	s

(c) Calculate the value of r and the value of s.

(3)

(d) Calculate the power of the test for $p = 0.2$ and $p = 0.4$

(2)

(e) Comment, giving your reasons, on the suitability of this test procedure.

(2)

(Total 12 marks)

[Mark scheme for Question 5](#)

[Examiner comment](#)

6. The number of goals scored by a football team is recorded for 100 games. The results are summarised in Table 1 below.

Number of goals	**Frequency**
0	40
1	33
2	14
3	8
4	5

Table 1

(a) Calculate the mean number of goals scored per game.

(2)

The manager claimed that the number of goals scored per match follows a Poisson distribution. He used the answer in part (*a*) to calculate the expected frequencies given in Table 2.

Number of goals	**Expected Frequency**
0	34.994
1	r
2	s
3	6.752
≥ 4	2.221

Table 2

(b) Find the value of r and the value of s giving your answers to 3 decimal places.

(3)

(c) Stating your hypotheses clearly, use a 5% level of significance to test the manager's claim.

(7)

(Total 12 marks)

[Mark scheme for Question 6](#)

[Examiner comment](#)

7. Four torpedoes are fired independently from a ship at a target. Each one has a probability of $\frac{1}{3}$ of hitting the target. The random variable X represents the number of hits and has probability generating function

$$G_X(t) = k(2 + t)^4.$$

(a) Show that $k = \dfrac{1}{81}$.

(3)

(b) Find the mean and the variance of X.

(2)

A second ship fires at the same target and the random variable Y, representing its number of hits, has probability generating function

$$G_Y(\mathrm{t}) = \frac{1}{243}(2 + t)^5.$$

Given that X and Y are independent,

(c) find the probability generating function of $Z = X + Y$.

(2)

(d) Calculate the mean and the variance of Z.

(6)

(Total 13 marks)

[Mark scheme for Question 7](#)

[Examiner comment](#)

TOTAL FOR PAPER: 75 MARKS

A level Further Mathematics – Further Statistics 1 – Practice Paper 05 – Mark scheme –

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

Question	Scheme	Marks
1(a)	$p + q + 0.2 + 0.3 + p = 1$ or $2p + q = 0.5$ (o.e)	**B1**
		(1)
(b)	[E(X) =] $-2p - q + ½\times0.2+1.5\times0.3+ 2p$ [=0.4] or $-q + 0.1 + 0.45$ [0.4]	**M1A1**
	$q = 0.15$	**(A1**
		(3)
(c)	$2p$ + "0.15" = 0.5 (o.e)	**M1**
	$p = 0.175$	**A1**
		(2)
(d)	[Var(X) =] $2.275 - (0.4)^2$	**M1**
	= 2.115 **(Accept 2.12)**	**A1**
		(2)
		(8 marks)

Mark scheme for Question 2

(Examiner comment) (Return to Question 2)

Question	Scheme	Marks
2(a)	$X \sim \text{Geo}(0.4)$	B1
	$\text{E(X)} = \dfrac{1}{p} = \dfrac{1}{0.4} = 2.5$	B1
		(2)
(b)	$W \sim \text{Neg.Bin}(r = 2 \text{ etc})$	
	$\text{E(W)} = \dfrac{r}{p} = \dfrac{2}{0.4} = 5$	M1A1
	$\text{Var(W)} = \dfrac{r(1-p)}{p^2} = \dfrac{2 \times 0.6}{0.4^2} = 7.5$	M1A1
		(4)
(c)	$\text{P(W = 5)} = \begin{pmatrix} 5-1 \\ 2-1 \end{pmatrix} 0.4^2 \times 0.6^3 = 0.13824 \quad (0.138)$	M1A1
		(2)
		(8 marks)

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

Question	Scheme	Marks
3(a)(i)	$X \sim \text{Po}(2.5)$	
	$\text{P}(X \geq 4) = 1 - \text{P}(X \leq 3)$ $= 1 - 0.7576$	M1
	$= 0.2424$	A1
(ii)	$X \sim \text{Po}(0.625)$	B1
	$\text{P}(X = 3) = \dfrac{e^{-0.625}0.625^3}{3!}$	M1
	$= 0.02177...$	A1
		(4)*
	*Part (ii) would be 2 marks in the new specification and 3 marks in the old specification	
(b)	1 - $\text{P}(X = 0) < 0.2$ $\text{P}(X = 0) > 0.8$	M1
	$e^{-2.5t} > 0.8$ $t < 0.089\ldots$ hours = 5.36 mins	M1
	[t <] 5 mins	A1cso
		(3)
(c)	$\text{H}_0: \lambda = 2.5 \ (\lambda = 5)$ $\text{H}_1: \lambda > 2.5 \ (\lambda > 5)$	B1
	$\text{P}(X \geq 10) = 1 - \text{P}(X \leq 9)$ $= 1 - 0.9682$	M1
	$= 0.0318$	A1
	Sufficient evidence to reject H_0, Accept H_1, significant. 10 does lie in the Critical region.	M1d
	There is sufficient evidence that the mean rate of telephone **calls** has increased **(oe)**	A1cso
		(5)
		(12 marks)

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

Question	Scheme	Marks
4(a)	(see table below)	M1A1

Defect Type / Shift	D_1	D_2	
First Shift	47.25	15.75	63
Second Shift	56.25	18.75	75
Third Shift	46.5	15.5	62
	150	50	200

Scheme	Marks
H_0 : Type of defect is independent of Shift (no association) H_1 : Type of defect is not independent of Shift (association)	B1
(see table below)	M1A1

O	E	$\frac{(O-E)^2}{E}$	$\frac{O_i^2}{E_i}$
45	47.25	0.1071…	42.857...
18	15.75	0.3214…	20.571..
55	56.25	0.02777…	53.777...
20	18.75	0.0833…	21.333...
50	46.5	0.2634…	53.763...
12	15.5	0.7903…	9.290...

Scheme	Marks
$\frac{(O-E)^2}{E}$ =1.5934.. or $\frac{O_i^2}{E_i}$ -200=201.5934-200=1.5934.. awrt1.59	A1
$\nu = (3-1)(2-1) = 2$	B1
$\chi_2^2(0.10) = 4.605$	B1ft
1.59<4.605 so insufficient evidence to reject H_0	M1
Insufficient evidence to support manager's belief /claim.	A1
	(10)

(10 marks)

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

Question	Scheme	Marks
5(a)	H_0 $p = 0.35$ H_1 $p \neq 0.35$	**B1B1**
		(2)
(b)	Let X = Number cured then $X \sim B(20, 0.35)$	**B1**
	α = P(Type 1 error) = P($x \leqslant 3$) + P($x \geqslant 11$) given $p = 0.35$ = 0.0444 + 0.0532	**M1**
	= 0.0976	**A1**
		(3)
(c)	$\beta = P$(Type 11 error) = P($4 \leqslant x \leqslant 10$)	**M1**
	p 0.2 0.3 0.4 0.5 β 0.5880 **0.8758** 0.8565 **0.5868**	**A1A1**
		(3)
(d)	Power = $1 - \beta$	**M1**
	0.4120 0.1435	**A1**
		(2)
(e)	Not a good procedure. Better further away from 0.35 or	**B1**
	This is not a very powerful test (power = 1 - β)	**B1dep**
		(2)
		(12 marks)

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

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th></tr>
<tr><td rowspan="2">6(a)</td><td>$\lambda = \dfrac{0\times 40 + 1\times 33 + 2\times 14 + 3\times 8 + 4\times 5}{100} = 1.05$</td><td>M1A1</td></tr>
<tr><td></td><td>(2)</td></tr>
<tr><td rowspan="4">(b)</td><td>Using Expected frequency $= 100 \times \mathrm{P}(X = x) = 100 \times \dfrac{e^{-1.05}1.05^x}{x!}$ gives</td><td>M1</td></tr>
<tr><td>$r = 36.743$ awrt 36.743 or 36.744</td><td>A1</td></tr>
<tr><td>$s = 19.290$ 19.29 or awrt 19.290</td><td>A1</td></tr>
<tr><td></td><td>(3)</td></tr>
<tr><td rowspan="8">(c)</td><td>H_0 : Poisson distribution is a suitable model
H_1 : Poisson distribution is not a suitable model</td><td>B1</td></tr>
<tr><td><table>
<tr><th>Number of goals</th><th>Frequency</th><th>Expected frequency</th><th></th></tr>
<tr><td>0</td><td>40</td><td>34.994</td><td></td></tr>
<tr><td>1</td><td>33</td><td>36.743</td><td></td></tr>
<tr><td>2</td><td>14</td><td>19.290</td><td></td></tr>
<tr><td>3</td><td>8</td><td>6.752</td><td rowspan="2">8.972443</td></tr>
<tr><td>≥ 4</td><td>5</td><td>2.221</td></tr>
</table></td><td>M1</td></tr>
<tr><td>$\upsilon = 4 - 1 - 1 = 2$</td><td>B1ft</td></tr>
<tr><td>CR : $\chi_2^2(0.05) > 5.991$</td><td>B1</td></tr>
<tr><td>$\sum \dfrac{(\mathrm{O}-\mathrm{E})^2}{\mathrm{E}} = \dfrac{(40-34.9937)^2}{34.9937} + \ldots + \dfrac{(13-8.972443)^2}{8.972443}$</td><td>M1</td></tr>
<tr><td>[=0.7161…+0.3813…+1.4508…+1.80789..]
= 4.356. (ans in range 4.2 – 4.4)</td><td>A1</td></tr>
<tr><td>Not in critical region
Number of goals scored can follow a Poisson distribution / managers claim is justified</td><td>A1ft</td></tr>
<tr><td></td><td>(7)</td></tr>
<tr><td colspan="3">(12 marks)</td></tr>
</table>

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

Question	Scheme	Marks
7(a)	$G_X(1) = 1$	**M1**
	$k\,(2+1)^4 = 1$	**M1**
	$\therefore\ k = \dfrac{1}{81}$	**A1**
		(3)
(b)	$E(X)\ = \dfrac{4}{3}$; $Var(X) = \dfrac{8}{9}$	**B1B1**
		(2)
(c)	$G_Z(t) = G_X(t) \times G_Y(t)$	**M1**
	$= \dfrac{1}{81}(2 + t)^4 \times \dfrac{1}{243}(2 + t)^5$	
	$= \dfrac{1}{19683}(2 + t)^5$	**A1**
		(2)
(d)	$G'_Z(t) = \dfrac{9}{19683}(2 + t)^8 \Rightarrow E(Z) = G'_Z(1) = \dfrac{9\times 3^8}{19683} = 3$	**M1A1**
	$G''_Z(t) = \dfrac{72}{19683}(2 + t)^7 \Rightarrow G''_Z(1) = \dfrac{72\times 3^7}{19683} = 8$	**M1A1**
	$Var(Z) = G''_Z(1) + G'_Z(1)\ldots\{G'_Z(1)\}^2$ $= 8 + 3 - 9$	**M1**
	$= 2$	**A1**
		(6)
		(13 marks)

A level Further Mathematics – Further Statistics 1 – Practice Paper 05 – Examiner report –

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

1. Most candidates found the first four parts of this question accessible and appeared to have been well prepared for it. Almost everyone was able to use the sum of the probabilities condition to obtain a first equation in part (a) and there was often a correct expression for E(X) seen too, although sometimes this was set equal to 1 not 0.4.

Errors in solving the equations in part (b) sometimes led to negative values for the probabilities of p and q and it is disappointing to see candidates still using these impossible values in later stages of their working. One would hope that they would realise that an error had been made and then try to correct it.

The variance calculation in part (d) was usually correct with only a small number of candidates making slips, such as subtracting the mean instead of the mean squared, or dividing by 5.

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

2. This question proved to be quite a difficult one for many candidates, but almost all of them gained some marks. Too often the answer related to the advantages or disadvantages of modelling, rather than explaining the process involved. There were 4 marks for the question and these corresponded to 4 key words – observe; devise/collect; compare and refine. Candidates were not expected to use these specific words, but to embody these ideas in their answer. It was good to see that some candidates did this and gained full marks.

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

3. Part (a)(i) was answered extremely well overall with virtually all of the students gaining full marks. However, a few students attempted to find P(X > 4) or P(X = 4) in error. The vast majority of students also earned full marks in part (a)(ii), though more errors were seen than in part (a)(i). Some students used an incorrect value for λ, usually 2.5, having not realised that the new mean was 0.625. However, even when an incorrect value for λ was used, the correct method for finding P(X = 3) was generally applied.

In complete contrast to part (a), very few students gained any marks in part (b). This part of the question proved to be extremely challenging to all but the most able students, with many clearly not understanding the meaning of the question and hence being unable to access it at all. Of those that did show some semblance of knowing how to approach this question, by far the most common error was to associate P(X = 0) with 0.2, usually P(X = 0) < 0.2 or P(X = 0) = 0.2 and sometimes P(X = 0) > 0.2. These students would then either be unable to proceed any further or continued with $e^{-\lambda} < 0.2$, $e^{-\lambda} = 0.2$ or $e^{-\lambda} > 0.2$, accordingly.

Most students who began with P($X \geqslant 1$) < 0.2 or P(X = 0) > 0.8, did ultimately achieve the correct answer, even when students correctly obtained $e^{-\lambda} > 0.8$, some did not use an appropriate λ, some forgot to convert to minutes and in a few instances, an answer of 5.36 was incorrectly rounded up to 6 minutes.

Part (c) was answered well overall, with many students gaining full marks. Hypotheses were generally stated clearly and correctly, with a mean rate of 2.5, or more commonly 5. Occasionally, p (or no letter) was used incorrectly in place of λ or μ, or 10 used as the mean rate. Other errors included, finding $P(X \leqslant 10)$; $1 - P(X \leqslant 10)$; $P(X = 10)$; using Po(2.5) with $P(X \geqslant 5)$; and stating the critical region as $X > 9$ or just 10. Attempting to identify the critical region was generally less successful as an approach. Correct non-contextual statements tended to be obtained, although occasionally students included contradictory non-contextual statements or incorrectly interpreted (or made incorrect) comparisons with 0.05, 0.95 or 10.

Whilst, on occasion, students neglected to place these conclusions into context, on the whole students were able to draw the correct contextual conclusions. However, in a few instances, these conclusions were not fully contextual and did not refer to calls.

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

4. This was done very well by most candidates. Calculations were accurate and well set out, hypotheses clearly stated and a correct critical value was common. Some candidates used the easier calculation method and very few confused the solution with correlation.

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

5. This question proved to be the most challenging for many candidates. In part (a) they did not understand what was required and stated B(20.0.35) as their answer. In part (b) the most common error was to use $P(X \leqslant 4) + P(X \geqslant 10)$. Parts (c) and (d) were generally answered correctly but candidates showed a lack of understanding in part (e). Many thought this test was suitable and did not realise that the power of the test gives an indication of its suitability.

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

6. Parts (a) and (b) were answered very well and most scored full marks on these two parts but part (c) proved more challenging. Many insisted on including the mean of 1.05 in their hypotheses even though this was incompatible with their correct treatment of the degrees of freedom. The pooling of the last two groups was usually carried out and the calculation of the test statistic was often correct. There was some confusion over the calculation of the degrees of freedom though: many subtracted 2 but others only 1 and some were not sure whether to subtract from the number of classes before or after the pooling. A number failed to score the final mark because their conclusion was not given in context: comments such as "there is evidence to support the manager's claim" or "there is evidence that the number of goals scored in football matches does follow a Poisson distribution" are fine; "the data follows a Poisson distribution" is not.

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

7. This was perhaps the question that was answered best by the candidates. Many recognised the probability generating function of the binomial distribution and thus found the question relatively easy. Even those that did not recognise it scored most if not all of the marks.