Write your name here

Surname	Other names

Pearson Edexcel Level 3 GCE

Centre Number | Candidate Number

Further Mathematics

Advanced
Further Mathematics Option 1
Paper 3: Further Statistics 1
Further Mathematics Option 2
Paper 4: Further Statistics 1

Sample Assessment Material for first teaching September 2017

Time: 1 hour 30 minutes

Paper Reference **9FM0/3B** **9FM0/4B**

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 ▶

S54441A

1/1/1/1/

S 5 4 4 4 1 A 0 1 2 3

Pearson

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

1. Bacteria are randomly distributed in a river at a rate of 5 per litre of water. A new factory opens and a scientist claims it is polluting the river with bacteria. He takes a sample of 0.5 litres of water from the river near the factory and finds that it contains 7 bacteria. Stating your hypotheses clearly test, at the 5% level of significance, whether there is evidence that the level of pollution has increased.

(5)

Question 1 continued

(Total for Question 1 is 5 marks)

2. A call centre routes incoming telephone calls to agents who have specialist knowledge to deal with the call. The probability of a caller, chosen at random, being connected to the wrong agent is p.

The probability of at least 1 call in 5 consecutive calls being connected to the wrong agent is 0.049

The call centre receives 1000 calls each day.

(a) Find the mean and variance of the number of wrongly connected calls a day.

(7)

(b) Use a Poisson approximation to find, to 3 decimal places, the probability that more than 6 calls each day are connected to the wrong agent.

(2)

(c) Explain why the approximation used in part (b) is valid.

(2)

The probability that more than 6 calls each day are connected to the wrong agent using the binomial distribution is 0.8711 to 4 decimal places.

(d) Comment on the accuracy of your answer in part (b).

(1)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 2 continued

Question 2 continued

Question 2 continued

(Total for Question 2 is 12 marks)

3. Bags of £1 coins are paid into a bank. Each bag contains 20 coins.

The bank manager believes that 5% of the £1 coins paid into the bank are fakes. He decides to use the distribution $X \sim B(20, 0.05)$ to model the random variable X, the number of fake £1 coins in each bag.

The bank manager checks a random sample of 150 bags of £1 coins and records the number of fake coins found in each bag. His results are summarised in Table 1. He then calculates some of the expected frequencies, correct to 1 decimal place.

Number of fake coins in each bag	0	1	2	3	4 or more
Observed frequency	43	62	26	13	6
Expected frequency	53.8	56.6		8.9	

Table 1

(a) Carry out a hypothesis test, at the 5% significance level, to see if the data supports the bank manager's statistical model. State your hypotheses clearly.

(10)

The assistant manager thinks that a binomial distribution is a good model but suggests that the proportion of fake coins is higher than 5%. She calculates the actual proportion of fake coins in the sample and uses this value to carry out a new hypothesis test on the data. Her expected frequencies are shown in Table 2.

Number of fake coins in each bag	0	1	2	3	4 or more
Observed frequency	43	62	26	13	6
Expected frequency	44.5	55.7	33.2	12.5	4.1

Table 2

(b) Explain why there are 2 degrees of freedom in this case.

(2)

(c) Given that she obtains a χ^2 test statistic of 2.67, test the assistant manager's hypothesis that the binomial distribution is a good model for the number of fake coins in each bag. Use a 5% level of significance and state your hypotheses clearly.

(2)

Question 3 continued

Question 3 continued

Question 3 continued

(Total for Question 3 is 14 marks)

4. A random sample of 100 observations is taken from a Poisson distribution with mean 2.3

Estimate the probability that the mean of the sample is greater than 2.5

(4)

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 4 marks)

5. The probability of Richard winning a prize in a game at the fair is 0.15

Richard plays a number of games.

(a) Find the probability of Richard winning his second prize on his 8th game,

(2)

(b) State two assumptions that have to be made, for the model used in part (a) to be valid.

(2)

Mary plays the same game, but has a different probability of winning a prize. She plays until she has won r prizes. 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 prize in a game.

(4)

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 8 marks)

6. The probability generating function of the discrete random variable X is given by

$$G_X(t) = k(3 + t + 2t^2)^2$$

(a) Show that $k = \frac{1}{36}$

(2)

(b) Find $P(X = 3)$

(2)

(c) Show that $\text{Var}(X) = \frac{29}{18}$

(8)

(d) Find the probability generating function of $2X + 1$

(2)

Question 6 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 6 continued

Question 6 continued

(Total for Question 6 is 14 marks)

7. Sam and Tessa are testing a spinner to see if the probability, p, of it landing on red is less than $\frac{1}{5}$.
They both use a 10% significance level.

Sam decides to spin the spinner 20 times and record the number of times it lands on red.

(a) Find the critical region for Sam's test.

(2)

(b) Write down the size of Sam's test.

(1)

Tessa decides to spin the spinner until it lands on red and she records the number of spins.

(c) Find the critical region for Tessa's test.

(6)

(d) Find the size of Tessa's test.

(1)

(e) (i) Show that the power function for Sam's test is given by

$$(1 - p)^{19}(1 + 19p)$$

(ii) Find the power function for Tessa's test.

(4)

(f) With reference to parts (b), (d) and (e), state, giving your reasons, whether you would recommend Sam's test or Tessa's test when $p = 0.15$

(4)

Question 7 continued

Question 7 continued

Question 7 continued

(Total for Question 7 is 18 marks)

TOTAL FOR PAPER IS 75 MARKS

Paper 3B/4B: Further Statistics 1 Mark Schemes

Question	Scheme		Marks	AOs
1	$H_0 : \lambda = 5$ $(\lambda = 2.5)$ $H_1 : \lambda > 5$ $(\lambda > 2.5)$		B1	2.5
	$X \sim \text{Po}(2.5)$		B1	3.3
	Method 1:	**Method 2:**		
	$\text{P}(X \geqslant 7) = 1 - \text{P}(X \leqslant 6)$ $= 1 - 0.9858$	$\text{P}(X \geqslant 5) = 0.1088$ $\text{P}(X \geqslant 6) = 0.042$	M1	1.1b
	$= 0.0142$	CR $X \geqslant 6$	A1	1.1b
	$0.0142 < 0.05$ $7 \geqslant 6$ or 7 is in critical region or 7 is significant Reject H_0. There is evidence at the 5% significance level that the level of pollution has increased. **or** There is evidence to support the scientists claim is justified		A1cso	2.2b
			(5 marks)	

Notes:

B1: Both hypotheses correct using λ **or** μ **and** 5 **or** 2.5

B1: Realising that the model Po(2.5) is to be used. This may be stated or used

M1: Using or writing $1 - \text{P}(X \leqslant 6)$ **or** $1 - \text{P}(X < 7)$
a correct CR **or** $\text{P}(X \geqslant 5) =$ awrt 0.109 **and** $\text{P}(X \geqslant 6) =$ awrt 0.042

A1: awrt 0.0142 **or** CR $X \geqslant 6$ **or** $X > 5$

M1: A fully correct solution and drawing a correct inference in context

Question	Scheme	Marks	AOs
2(a)	$P(X \geqslant 1) = 1 - P(X = 0)$ $1 - P(X = 0) = 0.049$	B1	3.1b
	$P(X = 0) = 0.951$	B1	1.1b
	$x^5 = 0.951$ $x = 0.99$	M1	3.1b
	$p = 0.01$	A1	1.1b
	$X \sim B(1000, 0.01)$	M1	3.3
	Mean $= np = 10$	A1ft	1.1b
	Variance $= np(1 - p) = 9.9$	A1ft	1.1b
		(7)	
(b)	$X \sim$ Po("10") then require: $P(X > 6) = 1 - P(X \leqslant 6)$	M1	3.4
	$= 1 - 0.1301$		
	$= 0.870$	A1	1.1b
		(2)	
(c)	The approximation is valid as : the number of calls is large	B1	2.4
	The probability of connecting to the wrong agent is small	B1	2.4
		(2)	
(d)	The answer is accurate to 2 decimal place	**B1**	3.2b
		(1)	
		(12 marks)	

Notes:

(a)
B1: Realising that the P(at least 1 call) $= 1 - P(X = 0)$
B1: Calculating $P(X = 0) = 0.951$
M1: Forming the equation $x^5 = \text{"their } 0.951\text{"}$ may be implied by $p = 0.01$
A1: 0.01 only
M1: Realising the need to use the model B(1000, 0.01) This may be stated or used
A1: Mean =10 or ft their p but only if $0 < p < 1$
A1: Var = 9.9 or ft their p but only if $0 < p < 1$

(b)
M1: Using the model Po("their 10") (this may be written or used) and $1 - P(X \leqslant 6)$
A1: awrt 0.870 Award M1 A1 for awrt 0.870 with no incorrect working

(c)
B1: Explaining why approximation is valid - need the context of number and calls
B1: Need the context connecting, wrong agent

(d)
B1: Evaluating the accuracy of their answer in (b). Allow 2 significant figures

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th><th>AOs</th></tr>
<tr><td>3(a)</td><td>Expected value for 2 $= 150 \times \mathrm{P}(X = 2)$</td><td>M1</td><td>3.4</td></tr>
<tr><td></td><td>$= 28.3015\ldots$</td><td>A1</td><td>1.1b</td></tr>
<tr><td></td><td>Expected value for 4 or more = 150 – (53.8 + 56.6 + 28.3 + 8.9)
= 2.4</td><td>A1ft</td><td>1.1b</td></tr>
<tr><td></td><td>H_0: Bin(20, 0.05) is a suitable model
H_1: Bin(20, 0.05) is not a suitable model</td><td>B1</td><td>2.5</td></tr>
<tr><td></td><td>Combining last two groups
<table><tr><th></th><th>$\geqslant$ 3</th></tr><tr><td>Observed frequency</td><td>19</td></tr><tr><td>Expected frequency</td><td>11.3</td></tr></table></td><td>M1</td><td>2.1</td></tr>
<tr><td></td><td>$\nu = 4 - 1 = 3$</td><td>B1</td><td>1.1b</td></tr>
<tr><td></td><td>Critical value, χ^2 (0.05) = 7.815</td><td>B1</td><td>1.1a</td></tr>
<tr><td></td><td>Test statistic = $\frac{(43-53.8)^2}{53.8} + \frac{(62-56.6)^2}{56.6} + \ldots$</td><td>M1</td><td>1.1b</td></tr>
<tr><td></td><td>= 8.117</td><td>A1</td><td>1.1b</td></tr>
<tr><td></td><td>In critical region, sufficient evidence to reject H_0, accept H_1
Significant evidence at 5% level to reject the manager's model</td><td>A1</td><td>3.5a</td></tr>
<tr><td></td><td></td><td>(10)</td><td></td></tr>
<tr><td>(b)</td><td>$\nu = 4 - 2 = 2$</td><td></td><td></td></tr>
<tr><td></td><td>4 classes due to pooling</td><td>B1</td><td>2.4</td></tr>
<tr><td></td><td>2 restrictions (equal total and mean/proportion)</td><td>B1</td><td>2.4</td></tr>
<tr><td></td><td></td><td>(2)</td><td></td></tr>
<tr><td>(c)</td><td>H_0: Binomial distribution is a good model
H_1: Binomial distribution is not a good model</td><td>B1</td><td>3.4</td></tr>
<tr><td></td><td>Critical value, χ^2 (0.05) = 5.991
Test statistic is not in critical region, insufficient evidence to reject H_0
There is evidence that the Binomial distribution is a good model</td><td>B1</td><td>3.5a</td></tr>
<tr><td></td><td></td><td>(2)</td><td></td></tr>
<tr><td colspan="4" align="right">(14 marks)</td></tr>
</table>

Question 3 notes:	
(a)	
M1:	Using the binomial model $150 \times p^2 \times (1-p)^{18}$ may be implied by 28.3
A1:	awrt 28.3
A1:	awrt 2.4 or ft their "28.3"
B1:	Both hypotheses correct using the correct notation or written out in full
M1:	For recognising the need to combine groups
B1:	Number of degrees of freedom = 3 may be implied by a correct CV
B1:	awrt 7.82
M1:	Attempting to find $\sum \frac{(O_i - E_i)^2}{E_i}$ or $\sum \frac{O_i^2}{E_i} - N$ may be implied by awrt 8.12
A1:	awrt 8.12
A1:	Evaluating the outcome of a model by drawing a correct inference in context
(b)	
B1:	Explaining why there are 4 classes
B1:	Explanation of why 2 is subtracted
(c)	
B1:	Correct hypotheses for the refined model
B1:	The CV awrt 5.99 and drawing the correct inference for the refined model

Question	Scheme	Marks	AOs
4	Po(2.3) $n = 100$ $\mu = 2.3$ $\sigma^2 = 2.3$		
	$\text{CLT} \Rightarrow \bar{X} \approx \text{N}\left(2.3, \frac{2.3}{100}\right)$	M1 A1	3.1a 1.1b
	$\text{P}(\bar{X} > 2.5) = \text{P}\left(Z > \frac{2.5 - 2.3}{\sqrt{0.023}}\right)$	M1	3.4
	$= \text{P}(Z > 1.318..)$		
	$= 0.09632\ldots$	A1	1.1b
		(4)	
			(4 marks)

Notes:

M1: For realising the need to use the CLT to set $\bar{X} \approx$ normal with correct mean
May be implied by using the correct normal distribution

A1: For fully correct normal stated or used

M1: Use of the normal model to find $\text{P}(\bar{X} > 2.5)$. Can be awarded for $\frac{2.5-2.3}{\sqrt{0.023}}$ **or** awrt 1.32

A1: awrt 0.0963

Question	Scheme	Marks	AOs
5(a)	$\binom{7}{1} \times 0.15^2 \times (0.85)^6$	M1	3.3
	$= 0.05940\ldots \quad =$ awrt **<u>0.0594</u>**	A1	1.1b
		(2)	
(b)	The model is only valid if:		
	the games (trials) are **independent**	B1	3.5b
	the probability of winning a prize, 0.15, is **constant** for each game	B1	3.5b
		(2)	
(c)	$18 = \frac{r}{p}$ and $6^2 = \frac{r(1-p)}{p^2}$	M1 A1	3.1b 1.1b
	Solving: $2p = 1 - p$	M1	1.1b
	$p = \frac{1}{3}$ (> 0.15) so Mary has the greater chance of winning a prize	A1	3.2a
		(4)	
	(8 marks)		

Notes:

5(a)

M1: For selecting an appropriate model negative binomial or B(7, 0.15) with an extra success in 8th trial e.g.

$\binom{7}{1} 0.15 \times (0.85)^6 \times 0.15$ Allow $\binom{7}{1} 0.85 \times (0.15)^6 \times 0.85$ may be implied by awrt 0.0594

A1: awrt 0.0594

(b)

B1: Stating the first assumption that games are independent

B1: Stating the second assumption that the probability remains constant

(c)

M1: Forming an equation for the mean or for the standard deviation

A1: Both equations correct

M1: Solving the 2 equations leading to $2p = 1 - p$

A1: For $p = \frac{1}{3}$ followed by a correct deduction

Question	Scheme	Marks	AOs
6(a)	$G_X(1)=1$ gives	M1	2.1
	$k\times 6^2=1$ so $k=\frac{1}{36}$ *	A1*cso	1.1b
		(2)	
(b)	$P(X=3)$ = coefficient of t^3 so $G_X(t)=k(...+4t^3...)$	M1	1.1b
	[$P(X=3)=$] $\frac{1}{9}$	A1	1.1b
		(2)	
(c)	$G'_X(t)=2k(3+t+2t^2)\times(1+4t)$	M1	2.1
	$E(X)=G'_X(1)=2k(3+1+2)\times(1+4)$	M1	1.1b
	$=\frac{5}{3}$	A1	1.1b
	$G''_X(t)=2k\left[(3+t+2t^2)\times4+(1+4t)^2\right]$	M1 A1	2.1 1.1b
	$G''_X(1)=2k[6\times4+5^2]$ $\left\{=\frac{49}{18}\right\}$	M1	1.1b
	$\text{Var}(X)=G''_X(1)+G'_X(1)-[G'_X(1)]^2=\frac{49}{18}+\frac{5}{3}-\frac{25}{9}$	M1	2.1
	$=\frac{29}{18}$ *	A1*cso	1.1b
		(8)	
(d)	$G_{2X+1}(t)=\frac{t}{36}\left(3+t^2+2(t^2)^2\right)^2$ [$\times t$ or sub t^2 for t]	M1	3.1a
	$=G_{2X+1}(t)=\frac{t}{36}(3+t^2+2t^4)^2$	A1	1.1b
		(2)	
			(14 marks)

Notes:

(a)

M1: Stating $G_X(1)=1$

A1*: Fully correct proof with no errors cso

(b)

M1: Attempting to find the coefficient of t^3. May be implied by obtaining $\frac{1}{9}$ or awrt 0.11

A1: $\frac{1}{9}$, allow awrt 0.111

Question 6 notes continued:
(c) **M1:** Attempting to find $G'_X(t)$. Allow Chain rule or multiplying out the brackets and differentiating **M1:** Substituting $t = 1$ into $G'_X(t)$ **A1:** $\frac{5}{3}$, allow awrt 1.67 **M1:** Attempting to find $G''_X(t)$ **A1:** $2k\left[\left(3+t+2t^2\right)\times 4+\left(1+4t\right)^2\right]$ **or** $k(48t^2+24t+26)$ o.e. **A1:** $2k[6\times 4+5^2]$ o.e. **M1:** Using $G''_X(1)+G'_X(1)-[G'_X(1)]^2$ to find the Variance **A1*:** $\frac{29}{18}$ cso
(d) **M1:** Realising the need to $\times t$ **or** sub t^2 for t **A1:** $\frac{t}{36}\left(3+t^2+2t^4\right)^2$, or $\frac{t}{36}\left(9+6t^2+13t^4+4t^6+4t^8\right)$ o.e.

Question	Scheme	Marks	AOs
7(a)	$X \sim B(20, 0.2)$ and seek c such that $P(X \leqslant c) < 0.10$	M1	3.3
	$[P(X \leqslant 1) = 0.0692]$ CR is $X \leqslant 1$	A1	1.1b
		(2)	
(b)	Size = **<u>0.0692</u>**	B1ft	1.2
		(1)	
(c)	Y = no. of spins until red obtained so $Y \sim Geo(0.2)$	M1	3.3
	$\mu = \frac{1}{p}$ so if $p < 0.2$ then mean is <u>larger</u> so seek d so that $P(Y \geqslant d) < 0.10$	M1	2.4
	$P(Y \geqslant d) = (0.8)^{d-1}$	M1	3.4
	$(0.8)^{d-1} < 0.10 \Rightarrow d - 1 > \frac{\log(0.1)}{\log(0.8)}$	M1	1.1b
	$d > 11.3..$	A1	1.1b
	CR is $\boldsymbol{Y \geqslant 12}$	A1	2.2b
		(6)	
(d)	Size = $[0.8^{11} = 0.085899\ldots]$ = **<u>0.0859</u>**	B1	1.1b
		(1)	
(e)(i)	Power = P(reject H_0 when it is false) = $P(X \leqslant 1 \mid X \sim B(20, p))$	M1	2.1
	$= (1-p)^{20} + 20(1-p)^{19} p$	M1	1.1b
	$= (1-p)^{19}(1+19p)$ *	A1*cso	1.1b
(ii)	Power = $(1-p)^{11}$	B1	1.1b
		(4)	
(f)	Sam's test has smaller P(Type I error) (or size) so is better	B1	2.2a
	Power of Sam's test = 0.1755…	B1	1.1b
	Power of Tessa's test = 0.85^{11} = 0.1673…	B1	1.1b
	So for $p = 0.15$ **Sam's test** is recommended	B1	2.2b
		(4)	
			(18 marks)

Question 7 notes:	
(a)	
M1:	Realising the need to use the model Using B(20,0.2) with method for finding the CR or implied by a correct CR
A1:	$X \leqslant 1$ or $X < 2$
(b)	
B1:	awrt 0.0692
(c)	
M1:	Realising that the model Geo(0.2)is needed. This may be written or used
M1:	Realising the key step that they need to find $P(Y \geqslant d) < 0.10$
M1:	Using the model $(0.8)^{d-1}$
M1:	Using the model $(0.8)^{d-1} < 0.10$ and finding a method to solve leading to a value/range of values for d
A1:	For $d > 11.3..$
A1:	For $Y \geqslant 12$ or $Y > 11$ (a correct inference)
(d)	
B1ft:	awrt 0.0692. ft their answer to part (c)
(e)(i)	
M1:	Using B(20, p) and realizing they need to find $P(X \leqslant 1)$ o.e. This may be used or written
M1:	Using $P(X = 0) + P(X = 1)$
A1*:	Fully correct proof (no errors) cso
(ii)	
B1:	For $(1 - p)^{11}$
(f)	
B1:	Making a deduction about the tests using the answers to part(b) and (d)
B1:	awrt 0.0176
B1:	awrt 0.167
B1:	A correct inference about which test is recommended