OCR
Oxford Cambridge and RSA

Accredited

A Level Further Mathematics B (MEI)

Y432 Statistics Minor

Sample Question Paper

Date – Morning/Afternoon

Time allowed: 1 hour 15 minutes

Model Answers

OCR supplied materials:
- Printed Answer Booklet
- Formulae Further Mathematics B (MEI)

You must have:
- Printed Answer Booklet
- Formulae Further Mathematics B (MEI)
- Scientific or graphical calculator

INSTRUCTIONS
- Use black ink. HB pencil may be used for graphs and diagrams only.
- Complete the boxes provided on the Printed Answer Booklet with your name, centre number and candidate number.
- Answer **all** the questions.
- **Write your answer to each question in the space provided in the Printed Answer Booklet.**
- Additional paper may be used if necessary but you must clearly show your candidate number, centre number and question number(s).
- Do **not** write in the bar codes.
- You are permitted to use a scientific or graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.

INFORMATION
- The total number of marks for this paper is **60**.
- The marks for each question are shown in brackets **[ ]**.
- You are advised that an answer may receive **no marks** unless you show sufficient detail of the working to indicate that a correct method is used. You should communicate your method with correct reasoning.
- The Printed Answer Booklet consists of **12** pages. The Question Paper consists of **8** pages.

Answer **all** the questions.

1 A darts player is trying to hit the bullseye on a dart board. On each throw the probability that she hits it is 0.05, independently of any other throw.

(i) Find the probability that she hits the bullseye for the first time on her 10th throw. **[2]**

P(9 misses followed by bullseye) $= 0.95^9 \times 0.05$
$= 0.0315$ (3sf)

(ii) Find the probability that she does not hit the bullseye in her first 10 throws. **[1]**

P(10 misses) $= 0.95^{10} = 0.599$

(iii) Write down the expected number of throws which it takes her to hit the bullseye for the first time. **[1]**

$E(X) = \frac{1}{p} = \frac{1}{0.05} = 20$

2 The number of televisions of a particular model sold per week at a retail store can be modelled by a random variable X with the probability function shown in the table.

x	0	1	2	3	4
$P(X = x)$	0.05	0.2	0.5	0.2	0.05

(i) (*A*) Explain why $E(X) = 2$. **[1]**

As the distribution is symmetrical, E(X) lies at the centre thus $E(X) = 2$

(*B*) Find $\text{Var}(X)$. **[3]**

$E(X^2) = \sum x^2 P(X=x) = (0^2 \times 0.05) + (1^2 \times 0.2)$
$+ (2^2 \times 0.5) + (3^2 \times 0.2) + (4^2 \times 0.05) = 4.8$
$\text{Var}X = E(X^2) - [E(X)]^2 = 4.8 - 2^2 = 0.8$

(ii) The profit, measured in pounds made in a week, on the sales of this model of television is given by Y, where $Y = 250X - 80$.
Find

- $E(Y)$ and
- $\text{Var}(Y)$. **[2]**

$E(Y) = E(250X - 80) = 250E(X) - 80$
$= 250(2) - 80 = 420$

$\text{Var}Y = \text{Var}(250X - 80) = 250^2 \text{Var}X$
$= 250^2 \times 0.8 = 50{,}000$

The remote controls for the televisions are quality tested by the manufacturer to see how long they last before they fail.

(iii) Explain why it would be inappropriate to test all the remote controls in this way. **[1]**

If all the remote controls were tested until failure, there wouldn't be any left to sell to customers.

(iv) State an advantage of using random sampling in this context. **[1]**

Random sampling avoids unsuspected sources of bias

3 A website awards a random number of loyalty points each time a shopper buys from it. The shopper gets a whole number of points between 0 and 10 (inclusive). Each possibility is equally likely, each time the shopper buys from the website. Awards of points are independent of each other.

(i) Let X be the number of points gained after shopping once.
Find

- the mean of X

$E(X) = 5 \leftarrow$ exactly between 0 and 10

- the variance of X.

$\text{Var}\,X = \frac{1}{12}(n^2 - 1) = \frac{1}{12}(11^2 - 1) = \frac{120}{12} = 10$

$n = 11$ as includes 0

[3]

(ii) Let Y be the number of points gained after shopping twice.
Find

$Y = X_1 + X_2$

- the mean of Y

$E(Y) = E(X_1 + X_2) = E(X_1) + E(X_2)$
$= 5 + 5 = 10$

- the variance of Y.

$\text{Var}\,Y = \text{Var}(X_1 + X_2) = \text{Var}\,X_1 + \text{Var}\,X_2$
$= 10 + 10 = 20$

[3]

(iii) Find the probability of the most likely number of points gained after shopping twice. Justify your answer. **[4]**

There are 11 options for the number of points awarded in each shop (0 - 10) so for 2 shops, the total number of combinations is $11 \times 11 = 121$ possibilities.
The most likely total is 10 which can be achieved in the following ways:
(0, 10), (1, 9), (2, 8), (3, 7), (4, 6), (5, 5), (6, 4), (7, 3), (8, 2), (9, 1), (10, 0) so 11 ways
Probability $= \frac{11}{121} = \frac{1}{11}$

4 **(i)** State the conditions under which the Poisson distribution is an appropriate model for the number of emails received by one person in a day. **[2]**

Emails are recieved independently, randomly and at a constant average rate

Jane records the number of junk emails which she receives each day. During working hours (9am to 5pm, Monday to Friday) the mean number of junk emails is 7.4 **per day**. Outside working hours (5pm to 9am), the mean number of junk emails is 0.3 **per hour**.

For the remainder of this question, you should assume that Poisson models are appropriate for the number of junk emails received during each of "working hours" and "outside working hours".

(ii) Find the probability that the number of junk emails which she receives between 9am and 5pm on a Monday is $X \sim Po(7.4)$

(*A*) exactly 10, **[1]**

$$P(X=10) = \frac{e^{-7.4} \times 7.4^{10}}{10!} = 0.0829$$

(*B*) at least 10. **[2]**

$$P(X \geq 10) = 1 - P(X \leq 9) = 1 - 0.7877$$
$$= 0.2123$$

(iii) (*A*) What assumption must you make to calculate the probability that the number of junk emails which she receives from 9am Monday to 9am Tuesday is at most 20? **[1]**

The number of junk emails recieved are independent.

(*B*) Find the probability. **[2]**

$\lambda = 7.4 + 16(0.3) = 12.2$. $Y \sim Po(12.2)$
$P(Y \leq 20) = 0.9863$

5 Each contestant in a talent competition is given a score out of 20 by a judge. The organisers suspect that the judge's scores are associated with the age of the contestant. Table 5.1 and the scatter diagram in Fig. 5.2 show the scores and ages of a random sample of 7 contestants.

Contestant	A	B	C	D	E	F	G
Age	66	51	39	29	9	22	14
Score	12	11	15	17	16	18	9

Table 5.1

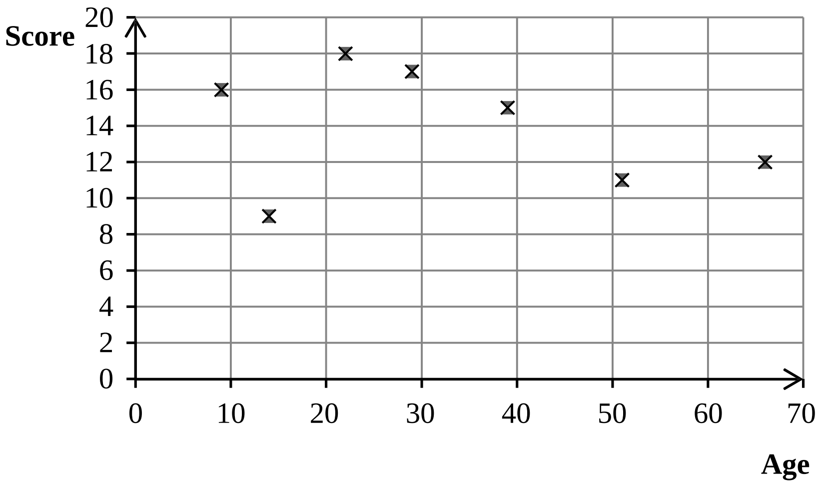

Fig. 5.2

Contestant G did not finish her performance, so it is decided to remove her data.

(i) Spearman's rank correlation coefficient between age and score, including all 7 contestants, is −0.25. Explain why Spearman's rank correlation coefficient becomes more negative when the data for contestant G is removed. **[1]**

Spearman's will become more negative as there is a stronger tendency for score to go down as age goes up

(ii) Calculate Spearman's rank correlation coefficient for the 6 remaining contestants. **[3]**

Contestant	A	B	C	D	E	F
Age rank	6	5	4	3	1	2
Score rank	2	1	3	5	4	6
Difference	4	4	1	2	3	4

$\Sigma d^2 = 4^2 + 4^2 + 1^2 + 2^2 + 3^2 + 4^2$
$= 62$

$r_s = 1 - \frac{6\Sigma d^2}{n(n^2-1)} = 1 - \frac{6\times62}{6(6^2-1)}$
$= -0.7714$ (4sf)

(iii) Using this value of Spearman's rank correlation coefficient, carry out a hypothesis test at the 5% level to investigate whether there is any association between age and score. **[5]**

H_0: There is no association between age and score
H_1: There is an association between age and score
For $n=6$ and $p=5\%$, critical value $=0.8857$.
Since $0.7714 < 0.8857$, this result is not significant so insufficient evidence to reject H_0, which would suggest there is no association between age and score.

(iv) Briefly explain why it may be inappropriate to carry out a hypothesis test based on Pearson's product moment correlation coefficient using these data. **[1]**

You cannot tell if the data is from a Bivariate Normal population so you cannot do a pmcc test.

6 At a bird feeding station, birds are captured and ringed. If a bird is recaptured, the ring enables it to be identified. The table below shows the number of recaptures, x, during a period of a month, for each bird of a particular species in a random sample of 40 birds.

Number of recaptures, x	0	1	2	3	4	5	6	7	8	9	10
Frequency	2	5	5	9	10	4	3	1	0	1	0

(i) The sample mean of x is 3.4. Calculate the sample variance of x. **[2]**

$\Sigma fx^2 = (1^2 \times 5) + (2^2 \times 5) + (3^2 \times 9) + (4^2 \times 10) + (5^2 \times 4) + (6^2 \times 3) + (7^2 \times 1) + (9^2 \times 1) = 604$

$S_x = \Sigma fx^2 - n\bar{x}^2 = 604 - 40(3.4^2) = 3.63$

(ii) Briefly comment on whether the results of part **(i)** support a suggestion that a Poisson model might be a good fit to the data. **[1]**

The sample mean is similar to the sample variance so Poisson model might be a good fit.

The screenshot below shows part of a spreadsheet for a χ^2 test to assess the goodness of fit of a Poisson model. The sample mean of 3.4 has been used as an estimate of the Poisson parameter. Some values in the spreadsheet have been deliberately omitted.

	A	B	C	D	E
1	Number of recaptures	Observed frequency	Poisson probability	Expected frequency	Chi-squared contribution
2	0 or 1	7	0.1468	5.8737	0.2160
3	2	5			0.9560
4	3	9	0.2186	8.7447	0.0075
5	4	10	0.1858	7.4330	0.8865
6	5	4	0.1264	5.0544	
7	≥6	5	0.1295	5.1783	0.0061

(iii) State the null and alternative hypotheses for the test. **[1]**

H_0: Poisson model is a good fit
H_1: Poisson model is not a good fit

(iv) Calculate the missing values in cells

- C3 $= \frac{e^{-3.4} \times 3.4^2}{2!} = 0.1929$
- D3 $= 0.1929 \times 40 = 7.7159$
- E6. $= \frac{(4-5.0544)^2}{5.0544} = 0.2200$

[4]

(v) Complete the test at the 10% significance level.

$\chi^2 = 0.2160 + 0.9560 + 0.0075 + 0.8865 + 0.2200 + 0.0061 = 2.2921$

Critical value at $p = 10\%$ for $v = 6-1-1 = 4$ is 7.779. As $2.2921 < 7.779$, this result is not significant so insufficient evidence to reject H_0, which would suggest that the Poisson model is a good fit.

[5]

(vi) The screenshot below shows part of a spreadsheet for a χ^2 test for a different species of bird. Find the value of the Poisson parameter used.

	A	B	C	D	E
1	Number of recaptures	Observed frequency	Poisson probability	Expected frequency	Chi-squared contribution
3	1	10	0.25716	12.8579	0.6352
4	2	7	0.27002	13.5008	3.1302
5	3	15	0.18901	9.4506	3.2587
6	≥4	11	0.16136	8.0679	1.0656

[3]

$$P(X=1)=\frac{e^{-\lambda}\lambda}{1!}=0.25716 \text{ and}$$

$$P(X=2)=\frac{e^{-\lambda}\lambda^2}{2!}=0.27002$$

$$e^{-\lambda}\lambda=0.25716 \text{ and } e^{-\lambda}\lambda=\frac{2(0.27002)}{\lambda}$$

$$0.25716=\frac{2(0.27002)}{\lambda}$$

$$\Rightarrow \lambda = 2.\underline{1}$$

7 A fair coin has +1 written on the heads side and -1 on the tails side. The coin is tossed 100 times. The sum of the numbers showing on the 100 tosses is the random variable *Y*. Show that the variance of *Y* is 100. **[4]**

Let X be the number of heads in 100 tosses so $X \sim B(100, 0.5)$.

$Y = X-(100-X)=2X-100$

$\text{Var}X = npq = 100\times0.5\times0.5=25.$

$\text{Var}(2X-100)=4\text{Var}X=4\times25=100$

END OF QUESTION PAPER