



Edexcel AS Further Mathematics

8FM0/23 - Further Statistics 1

Thursday 16 May 2019

Questions and Fully Worked Model Solutions

This booklet places the **question and the solution together**. The mathematics is set out using proper fractions, powers, roots and exact notation, with extra written explanation so that a student can see *why* each step is being taken.

Solutions have been checked against the Pearson Edexcel Summer 2019 mark scheme. The aim is a clear **full-mark model response**, not merely a list of final answers.

Contents

Question 1	2
Question 2	4
Question 3	5
Question 4	7

www.mathswithmatt.co.uk

Question 1

Question 1

A leisure club offers a choice of badminton, bowls or snooker to its 150 members on a Tuesday evening. The data collected are

Age a years	Badminton	Bowls	Snooker
$a < 20$	9	3	3
$20 \leq a < 40$	10	10	14
$40 \leq a < 50$	16	15	5
$50 \leq a < 60$	15	13	11
$a \geq 60$	4	19	3

- Write down suitable hypotheses for a test of whether activity choice is associated with age. (1)
- Calculate the expected frequency of members aged 60 or over who choose snooker. (1)
- Explain why 6 degrees of freedom are used in this test. (2)
- The test statistic is 19.583. Using a 5% significance level, complete the test. (2)

Model solution

(a) The hypotheses for a chi-squared test of association are

H_0 : there is no association between a member's age and chosen activity,

H_1 : there is an association between a member's age and chosen activity.

H_0 : no association; H_1 : an association.

(1)

(b) The total aged 60 or over is $4+19+3 = 26$, and the snooker total is $3+14+5+11+3 = 36$. Hence

$$E = \frac{26 \times 36}{150} = 6.24.$$

6.24

(1)

(c) The expected frequency for the $a < 20$, snooker cell is

$$\frac{15 \times 36}{150} = 3.6 < 5.$$

Therefore the first two age rows must be amalgamated. The resulting table has 4 rows and 3 columns, so

$$\nu = (4 - 1)(3 - 1) = 6.$$

6 degrees of freedom.

(2)

(d) For $\nu = 6$, the 5% critical value is

$$\chi^2_{6,0.05} = 12.592.$$

Since $19.583 > 12.592$, reject H_0 . There is significant evidence, at the 5% level, of an association between age and chosen activity; this supports the manager's belief.

Reject H_0 ; there is evidence of an association.

(2)

www.mathswithmatt.co.uk

Question 2

Question 2

A spinner is designed to give the following probabilities:

Score	1	2	3	4	6
Probability	$\frac{3}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{2}{5}$	$\frac{1}{10}$

When spun 80 times, the corresponding observed frequencies are 15, 4, 12, 41, 8. Test, at the 10% significance level, whether the spinner is giving scores as designed. State hypotheses and show working. (7)

Model solution

Use a chi-squared goodness-of-fit test.

H_0 : the spinner gives scores with the designed distribution, H_1 : the spinner does not give scores

The expected frequencies are 80 times the stated probabilities:

Score	1	2	3	4	6
O_i	15	4	12	41	8
E_i	24	8	8	32	8

Thus

$$\begin{aligned}\chi^2 &= \sum \frac{(O_i - E_i)^2}{E_i} \\ &= \frac{81}{24} + \frac{16}{8} + \frac{16}{8} + \frac{81}{32} + 0 \\ &= 9.90625.\end{aligned}$$

There are $5 - 1 = 4$ degrees of freedom. The 10% critical value is

$$\chi_{4,0.10}^2 = 7.779.$$

Since $9.90625 > 7.779$, reject H_0 . There is significant evidence at the 10% level that the spinner is not giving scores as it is designed to do.

Reject H_0 : the spinner is not operating as designed.

(7)

Question 3

Question 3

Andreia's secretary makes random errors at an average rate of 1.7 errors per 100 words.

- (a) Find the probability of fewer than 2 errors in the next 100-word piece. (2)
- (b) Find the probability of exactly 5 errors in a 250-word article. (2)
- (c) From a random sample of 40 pieces, each of 100 words, scheme A gives a bonus if more than 10 pieces contain no errors. Scheme B gives a bonus if the total number of errors is fewer than 56. Explain, with calculations, which scheme should be chosen. (5)
- (d) After the bonus scheme, a randomly selected 500-word piece is used to test whether the error rate has decreased. State hypotheses and, at the 5% significance level, find the critical region. (4)

Model solution

(a) Let X be the number of errors in a 100-word piece. Then $X \sim \text{Po}(1.7)$, so

$$P(X < 2) = P(X \leq 1) = e^{-1.7}(1 + 1.7) = 0.49324 \dots$$

$$0.493 \text{ (3 s.f.)}$$

(2)

(b) Let R be the number of errors in 250 words. Then $R \sim \text{Po}(2.5 \times 1.7) = \text{Po}(4.25)$.
Hence

$$P(R = 5) = e^{-4.25} \frac{4.25^5}{5!} = 0.16482 \dots$$

$$0.165 \text{ (3 s.f.)}$$

(2)

(c) For scheme A, the probability that a 100-word piece contains no errors is

$$P(X = 0) = e^{-1.7}.$$

Let A be the number of error-free pieces in the 40-piece sample. Then

$$A \sim B(40, e^{-1.7}).$$

Therefore

$$P(\text{bonus under A}) = P(A > 10) = 1 - P(A \leq 10) = 0.0995591 \dots$$

For scheme B, let B be the total number of errors in all 40 pieces. Since Poisson means add,

$$B \sim \text{Po}(40 \times 1.7) = \text{Po}(68).$$

Thus

$$P(\text{bonus under B}) = P(B < 56) = P(B \leq 55) = 0.061133\dots$$

Since $0.0995591\dots > 0.061133\dots$, scheme A gives the greater probability of receiving a bonus.

Choose scheme A.

(5)

(d) Let E be the number of errors in the 500-word piece. The hypotheses are

$H_0 : \lambda = 1.7$ errors per 100 words (equivalently $\mu = 8.5$ errors per 500 words),

$H_1 : \lambda < 1.7$ errors per 100 words (equivalently $\mu < 8.5$ errors per 500 words).

Under H_0 , $E \sim \text{Po}(8.5)$. For the lower tail,

$$P(E \leq 3) = 0.0301\dots \leq 0.05, \quad P(E \leq 4) = 0.0744\dots > 0.05.$$

The largest lower-tail critical region whose probability is at most 5% is therefore $E \leq 3$.

Critical region: $E \leq 3$.

(4)

Question 4

Question 4

The discrete random variable X has distribution

x	-3	-1	1	2	4
$P(X = x)$	q	$\frac{7}{30}$	$\frac{7}{30}$	q	r

where q and r are probabilities.

(a) Write down, in terms of q , $P(X < 0)$. (1)

(b) Show that $E(X^2) = \frac{7}{15} + 13q + 16r$. (2)

(c) Given that $E(X^3) = E(X^2) + E(6X)$, find q and r . (7)

(d) Hence find $P(X^3 > X^2 + 6X)$. (4)

Model solution

(a)

$$P(X < 0) = q + \frac{7}{30}.$$

$$q + \frac{7}{30}$$

(1)

(b)

$$\begin{aligned} E(X^2) &= (-3)^2q + (-1)^2\frac{7}{30} + 1^2\frac{7}{30} + 2^2q + 4^2r \\ &= 9q + \frac{7}{30} + \frac{7}{30} + 4q + 16r \\ &= \frac{7}{15} + 13q + 16r. \end{aligned}$$

$$E(X^2) = \frac{7}{15} + 13q + 16r.$$

(2)

(c) First,

$$E(X) = -3q - \frac{7}{30} + \frac{7}{30} + 2q + 4r = 4r - q,$$

and

$$E(X^3) = -27q - \frac{7}{30} + \frac{7}{30} + 8q + 64r = 64r - 19q.$$

Using the given condition gives

$$64r - 19q = \left(\frac{7}{15} + 13q + 16r \right) + 6(4r - q) = \frac{7}{15} + 7q + 40r,$$

so

$$24r - 26q = \frac{7}{15}.$$

Also, the probabilities sum to 1:

$$2q + r + \frac{7}{30} + \frac{7}{30} = 1, \quad \text{so} \quad 2q + r = \frac{8}{15}.$$

Substituting $r = \frac{8}{15} - 2q$ into the first equation,

$$24\left(\frac{8}{15} - 2q\right) - 26q = \frac{7}{15},$$

which gives $q = \frac{1}{6}$. Therefore

$$r = \frac{8}{15} - 2\left(\frac{1}{6}\right) = \frac{1}{5}.$$

$$q = \frac{1}{6}, \quad r = \frac{1}{5}.$$

(7)

(d)

$$X^3 > X^2 + 6X \iff X^3 - X^2 - 6X > 0 \iff X(X-3)(X+2) > 0.$$

The inequality holds for $-2 < X < 0$ or $X > 3$. Of the possible values of X , these are -1 and 4 . Hence

$$P(X^3 > X^2 + 6X) = P(X = -1) + P(X = 4) = \frac{7}{30} + r = \frac{7}{30} + \frac{1}{5} = \frac{13}{30}.$$

$$\frac{13}{30}.$$

(4)