Oxford Cambridge and RSA

Thursday 16 May 2019 – Afternoon

AS Level Further Mathematics A

Y532/01 Statistics

Time allowed: 1 hour 15 minutes

You must have:
- Printed Answer Booklet
- Formulae AS level Further Mathematics A

You may use:
- a scientific or graphical calculator

INSTRUCTIONS
- Use black ink. HB pencil may be used for graphs and diagrams only.
- Answer **all** the questions.
- **Write your answer to each question in the space provided in the Printed Answer Booklet.** If additional space is required, use the lined page(s) at the end of the Printed Answer Booklet. The question number(s) must be clearly shown.
- You are permitted to use a scientific or graphical calculator in this paper.
- Give non-exact numerical answers correct to 3 significant figures unless a different degree of accuracy is specified in the question.
- The acceleration due to gravity is denoted by $g\,\mathrm{m\,s^{-2}}$. Unless otherwise instructed, when a numerical value is needed, use $g = 9.8$.

INFORMATION
- The total mark for this paper is **60**.
- The marks for each question are shown in brackets **[ ]**.
- **You are reminded of the need for clear presentation in your answers.**
- The Printed Answer Booklet consists of **12** pages. The Question Paper consists of **4** pages.

1 When a spinner is spun, the outcome is equally likely to be 1, 2 or 3. In a competition, the spinner is spun twice and the outcomes are added to give a total score T.

(a) Show that the expectation of T is 4. **[3]**

(b) Find the variance of T. **[3]**

a. Let t = score obtained (1, 2 or 3)

t	2	3	4	5	6
$P(T=t)$	$\frac{1}{3}\times\frac{1}{3} = \frac{1}{9}$	$\frac{1}{3}\times\frac{1}{3}\times 2 = \frac{2}{9}$	$\frac{1}{3}\times\frac{1}{3}\times 3 = \frac{3}{9}$	$\frac{1}{3}\times\frac{1}{3}\times 2 = \frac{2}{9}$	$\frac{1}{3}\times\frac{1}{3} = \frac{1}{9}$

$E(X) = \sum x P(X=x)$

$\therefore E(T) = 2\times\frac{1}{9} + 3\times\frac{2}{3} + 4\times\frac{3}{9} + 5\times\frac{2}{3} + 6\times\frac{1}{9}$

$= 4$

b. $Var(X) = E(X^2) - (E(X))^2$

$E(T^2) = 4\times\frac{1}{9} + 9\times\frac{2}{9} + 16\times\frac{2}{3} + 25\times\frac{2}{9} + 36\times\frac{1}{9}$

$= \frac{156}{9}$

$\therefore Var(T) = \frac{156}{9} - 4^2$

$= \frac{4}{3}$

A competitor pays £1.50 to enter the competition and receives £X, where $X = 0.3T$.

(c) **(i)** Find the expectation of the competitor's profit. **[1]**

(ii) Find the variance of the competitor's profit. **[2]**

i. Expected profit = winnings - entry fee

$= 0.3\,E(T) - 1.5$

$= \frac{12}{10} - \frac{3}{2}$

$= £\,-0.30$

30p loss is expected.

ii. Variance of profit $= 0.3^2 \times Var(T)$

$= 0.09 \times \frac{4}{3}$

$= £^2\ 0.12$

2 On any day, the number of orders received in one randomly chosen hour by an online supplier can be modelled by the distribution Po(120).

(a) Find the probability that at least 28 orders are received in a randomly chosen 10-minute period. **[2]**

$\Rightarrow P(x \geq 28)$

$P(x \geq 28) = 1 - P(x \leq 27)$

$= 1 - 0.9475$

$= 0.0525$

(b) Find the probability that in a randomly chosen 10-minute period on one day and a randomly chosen 10-minute period on the next day a total of at least 56 orders are received. **[3]**

λ for 10 min $= \dfrac{\lambda \text{ for 1 hour}}{6} = \dfrac{120}{6} = 20$

$\therefore$ for 2×10 minutes : $Po(40)$

$P(X \geq 56) = 1 - P(X \leq 55)$

$= 1 - 0.99032$

$= 0.00968$

(c) State a necessary assumption for the validity of your calculation in part **(b)**. **[1]**

- Orders on one day are independant of orders on the other day.
- All orders have a constant average rate.

3 **(a)** Shula calculates the value of Spearman's rank correlation coefficient r_s for 9 pairs of rankings.

Find the largest possible value of r_s that Shula can obtain that is less than 1. **[4]**

$r_s = 1 - \dfrac{6 \Sigma d^2}{n(n^2-1)}$

for 9 pairings, the ranking that gives the smallest non-zero Σd^2 value is:

R_x	1	2	3	4	5	6	7	8	9
R_y	2	1	3	4	5	6	7	8	9
d	-1	1	0	0	0	0	0	0	0
d^2	1	1	0	0	0	0	0	0	0

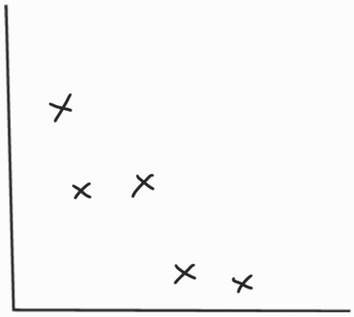
$\therefore$ Smallest possible non-zero $\Sigma d^2 = 2$

$\therefore$ largest r_s (<1) $= 1 - \dfrac{6 \times 2}{9(81-1)}$

$= \dfrac{59}{60}$

(b) A set of bivariate data consists of 5 pairs of values. It is known that for this data the value of Spearman's rank correlation coefficient is -1 but the value of Pearson's product-moment correlation coefficient is not -1.

Sketch a possible scatter diagram illustrating the data. **[2]**



5 pairs → 5 points
Spearman's = −1 → perfect negative
PMCC ≠ −1 → not perfectly linear

← Points strictly decreasing

4 The members of a team stand in a random order in a straight line for a photograph. There are four men and six women.

(a) Find the probability that all the men are next to each other. **[3]**

Possible configurations:
MMMM WWWWWW
WMMMMWWWWW
WWMMMMWWWW
WWWMMMMWWW
WWWWMMMMWW
WWWWWMMMMW
WWWWWWMMMM

$$\text{Probability} = \frac{7! \times 4!}{10!}$$

7! ← for men's positions
4! ← for women's positions
10! ← total

$$= \frac{120960}{3628800} = \frac{1}{30}$$

(b) Find the probability that no two men are next to one another. **[4]**

Women are placed in $6!$ ways

Men are placed in $4!$ ways

4 of the slots must be men, each with a woman between:

MWMWMWM $= 4M + 3W = 7$ total

$\Rightarrow$ probability $= {}^7C_4 \times \frac{4! \times 6!}{10!}$

$= \frac{1}{6}$

5 Sixteen candidates took an examination paper in mechanics and an examination paper in statistics.

(a) For all sixteen candidates, the value of the product moment correlation coefficient r for the marks on the two papers was 0.701 correct to 3 significant figures.

Test whether there is evidence, at the 5% significance level, of association between the marks on the two papers. **[4]**

$H_0: \rho = 0$

$H_1: \rho \neq 0$

(where ρ is population PMCC)

does not specify positive or negative, so H_1 is $\rho \neq 0$ rather than $\rho <$ or > 0.

Critical values, 5% two tailed $= \pm 0.4973$

$0.701 > 0.4973$ $\therefore$ significant

Accept H_1, there is significant evidence of association between the marks on the two papers.

(Reject H_0)

(b) A teacher decided to omit the marks of the candidates who were in the top three places in mechanics and the candidates who were in the bottom three places in mechanics. The marks for the remaining 10 candidates can be summarised by

$$n = 10,\ \Sigma x = 750,\ \Sigma y = 690,\ \Sigma x^2 = 57\,690,\ \Sigma y^2 = 49\,676,\ \Sigma xy = 50\,829.$$

(i) Calculate the value of r for these 10 candidates. **[2]**

(ii) What do the two values of r, in parts **(a)** and **(b)(i)**, tell you about the scores of the sixteen candidates? **[1]**

i. $r = \dfrac{S_{XY}}{\sqrt{S_{XX} S_{YY}}} = \dfrac{50\,829 - \frac{750 \times 690}{10}}{\sqrt{\left(57690 - \frac{750^2}{10}\right) \times \left(49676 - \frac{690^2}{10}\right)}}$

$= -0.534$

ii. The 6 candidates removed in (b) had very high or very low scores in both papers.
The middle 10 tended to do well on one paper and badly on the other. There is not much correlation in this section.

6 A bag contains a mixture of blue and green beads, in unknown proportions. The proportion of green beads in the bag is denoted by p.

(a) Sasha selects 10 beads at random, with replacement.

Write down an expression, in terms of p, for the variance of the number of green beads Sasha selects. **[1]**

Variance $= np(1-p)$
$= 10p - 10p^2$

Freda selects one bead at random from the bag, notes its colour, and replaces it in the bag. She continues to select beads in this way until a green bead is selected. The first green bead is the X th bead that Freda selects.

(b) Assume that $p = 0.3$.

Find

(i) $P(X \geqslant 5)$, **[2]**

(ii) $\mathrm{Var}(X)$. **[1]**

i. $P(x \geq 5) = 1 - P(x \leq 4)$
$= 1 - 0.3(1 + 0.7 + 0.7^2 + 0.7^3)$
$= 1 - 0.7599$
$= 0.240$ (3sf)

ii. $Var(x) = \dfrac{(1-p)}{p^2} = \dfrac{0.7}{0.09} = \dfrac{70}{9}$

(c) In fact, on the basis of a large number of observations of X, it is found that $P(X = 3) = \frac{4}{25} \times P(X = 1)$.

Estimate the value of p. **[5]**

$\Rightarrow (1-p^2)p = \frac{4}{25} p$

$\Rightarrow p = 0$ or $(1-p)^2 = \frac{4}{25}$

$p \neq 0 \therefore (1-p)^2 = \frac{4}{25}$

$1 - p = \pm \frac{2}{5}$

$p = \frac{7}{5}$ or $\frac{3}{5}$

$p < 1$, $\therefore$ $p = 3/5$

7 In a standard model from genetic theory, the ratios of types a, b, c and d of a characteristic from a genetic cross are predicted to be 9:3:3:1. Andrei collects 120 specimens from such a cross, and the numbers corresponding to each type of the characteristic are given in the table.

Type	a	b	c	d
Frequency	51	33	30	6

Andrei tests, at the 1% significance level, whether the observed frequencies are consistent with the standard model.

(a) State appropriate hypotheses for the test. **[1]**

(b) Carry out the test. **[6]**

a. H_0: population frequencies in ratio 9:3:3:1

H_1: population frequencies not in ratio 9:3:3:1

b.

Type	a	b	c	d
Frequency	51	33	30	6
Expected frequency	$\frac{9}{16}\times 120 = 67.5$	$\frac{3}{16}\times 120 = 22.5$	$\frac{3}{16}\times 120 = 22.5$	$\frac{1}{16}\times 120 = 7.5$

$$\sum \frac{(O-E)^2}{E} = \frac{(51-67.5)^2}{67.5} + \frac{(33-22.5)^2}{22.5} + \frac{(30-22.5)^2}{22.5} + \frac{(6-7.5)^2}{7.5}$$

$$= 4.033 + 4.9 + 2.5 + 0.3$$
$$= 11.73$$

Critical value $= 11.34$

$11.73 > 11.34$ ∴ significant

Accept H_1. There is significant evidence that the results are not consistent with the theory. (Reject H_0)

(c) State with a reason which one of the frequencies is least consistent with the standard model. **[1]**

(d) Suggest a different, improved model by changing exactly two of the ratio values. **[1]**

c. b, as it provides the largest value in the χ^2 calculation.

d. Increase b and c, e.g. 9 : 5 : 5 : 1

8 Alex claims that he can read people's minds. A volunteer, Jane, arranges the integers 1 to n in an order of Jane's own choice and Alex tells Jane what order he believes was chosen.

They agree that Alex's claim will be accepted if he gets the order completely correct or if he gets the order correct apart from two numbers which are the wrong way round.

They use a value of n such that, if Alex chooses the order of the integers at random, the probability that Alex's claim will be accepted is less than 1%.

Determine the smallest possible value of n. **[7]**

Probability of two or less numbers correct $= \dfrac{1 + {}^nC_2}{n!}$

Need probability less than 1% :

$$\frac{1 + {}^nC_2}{n!} < 0.01$$

if $n = 6$, $p = \dfrac{1 + {}^6C_2}{6!} = 0.0222\ldots > 0.01$

if $n = 7$, $p = \dfrac{1 + {}^7C_2}{7!} = 0.00437\ldots < 0.01$

$\therefore$ Smallest possible $n = 7$.