OCR
Oxford Cambridge and RSA

Accredited

A Level Further Mathematics A
Y542 Statistics
Sample Question Paper

Date – Morning/Afternoon
Time allowed: 1 hour 30 minutes

OCR supplied materials:
- Printed Answer Booklet
- Formulae A Level Further Mathematics A

You must have:
- Printed Answer Booklet
- Formulae A Level Further Mathematics A
- Scientific or graphical calculator

* 0 0 0 0 0 0 *

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.
- The acceleration due to gravity is denoted by $g\,\text{m}\,\text{s}^{-2}$. Unless otherwise instructed, when a numerical value is needed, use $g = 9.8$.

INFORMATION

- The total number of marks for this paper is **75**.
- 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 The table below shows the typical stopping distances d metres for a particular car travelling at v miles per hour.

v	20	30	40	50	60	70
d	13	24	36	52	72	94

(i) State each of the following words that describe the variable v.

Independent Dependent Controlled Response

[1]

(ii) Calculate the equation of the regression line of d on v. **[2]**

(iii) Use the equation found in part **(ii)** to estimate the typical stopping distance when this car is travelling at 45 miles per hour. **[1]**

It is given that the product moment correlation coefficient for the data is 0.990 correct to three significant figures.

(iv) Explain whether your estimate found in part **(iii)** is reliable. **[2]**

i. independent & controlled

ii. use calculator: $d = 1.61v - 24.1$ (3 s.f.)

iii. $d = (1.61 \times 45) - 24.1 = 48$

$v = 45$

iv. it is reliable as r is close to 1 & 45 is within the data range

2 The mass J kg of a bag of randomly chosen Jersey potatoes is a normally distributed random variable with mean 1.00 and standard deviation 0.06. The mass K kg of a bag of randomly chosen King Edward potatoes is an independent normally distributed random variable with mean 0.80 and standard deviation 0.04.

(i) Find the probability that the total mass of 6 bags of Jersey potatoes and 8 bags of King Edward potatoes is greater than 12.70 kg. **[3]**

(ii) Find the probability that the mass of one bag of King Edward potatoes is more than 75% of the mass of one bag of Jersey potatoes. **[3]**

i. $J \sim N(1, 0{\cdot}06^2)$ $\quad K \sim N(0{\cdot}8, 0{\cdot}04^2)$

$(6J+4K) \sim N(12{\cdot}4, 0{\cdot}0344)$

$P(6J+8K > 12{\cdot}7) = 1 - 0{\cdot}9471$

$= 0{\cdot}0529$

ii. $\sigma^2 = 0{\cdot}04^2 + \left(\frac{3}{4}\right)^2 \times 0{\cdot}06^2$

$= 0{\cdot}003625$

$(K - 0{\cdot}75J) \sim N(0{\cdot}05, 0{\cdot}003625)$

$P(K - 0{\cdot}75J > 0) = \Phi(0{\cdot}08305)$

$= 0{\cdot}7969$

3 A game is played as follows. A fair six-sided dice is thrown once. If the score obtained is even, the amount of money, in £, that the contestant wins is half the score on the dice, otherwise it is twice the score on the dice.

(i) Find the probability distribution of the amount of money won by the contestant. [3]

(ii) The contestant pays £5 for every time the dice is thrown.
Find the standard deviation of the loss made by the contestant in 120 throws of the dice. [5]

i.

(roll)	(2)	(1,4)	(6)	(3)	(5)
x	1	2	3	6	10
$P(X=x)$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

ii. $\sum xP(x) = \frac{1}{6} + \frac{4}{6} + \frac{3}{6} + \frac{6}{6} + \frac{10}{6} = 4$

$$\sum x^2 P(x) - \mu^2 = \frac{1}{6} + \frac{8}{6} + \frac{9}{6} + \frac{36}{6} + \frac{100}{6} - \mu^2$$

$$= \frac{77}{3} - \mu^2$$

$$= \frac{29}{3}$$

Variance of 120 games $= 120\sigma^2$

$\Rightarrow$ s.d. $= \sqrt{120 \times \frac{29}{3}} = 34.1$

4 A psychologist investigated the scores of pairs of twins on an aptitude test.
Seven pairs of twins were chosen randomly, and the scores are given in the following table.

Elder twin	65	37	60	79	39	40	88
Younger twin	58	39	61	62	50	26	84

(i) Carry out an appropriate Wilcoxon test at the 10% significance level to investigate whether there is evidence of a difference in test scores between the elder and the younger of a pair of twins. **[6]**

(ii) Explain the advantage in this case of a Wilcoxon test over a sign test. **[1]**

i.

difference E−Y	7	−2	−1	17	−11	14	4
rank	4	2	1	7	6	5	3
+/−	+	−	−	+	−	+	+

H_0 : population median difference $= 0$

H_1 : population median difference $\neq 0$

$P = 4+7+5+3 = 19$

$Q = 1+2+6 = 9$

$T = 8$

$T_{crit} = 3$

$8 > 3$ ∴ do not reject H_0 as insufficient evidence of a difference in test scores

ii. it uses the magnitudes of the diferences

5 The number of goals scored by the home team in a randomly chosen hockey match is denoted by X.

(i) In order for X to be modelled by a Poisson distribution it is assumed that goals scored are random events. State two other conditions needed for X to be modelled by a Poisson distribution in this context. **[2]**

Assume now that X can be modelled by the distribution $\mathrm{Po}(1.9)$.

(ii) (a) Write down an expression for $\mathrm{P}(X = r)$. **[1]**

(b) Hence find $\mathrm{P}(X = 3)$. **[1]**

(iii) Assume also that the number of goals scored by the away team in a randomly chosen hockey match has an independent Poisson distribution with mean λ between 1.31 and 1.32. Find an estimate for the probability that more than 3 goals are scored altogether in a randomly chosen match. **[4]**

i. 1. goals are scored independently of each other

2. goals are scored at a uniform (average) rate

ii. a) $P(X=r) = e^{-1.9} \times \frac{1.9^r}{r!}$

b) $r=3 \Rightarrow P(X=3) = 0.171$

iii. total $\sim P_o(1.9+\lambda)$

$\lambda = (1.9 + 1.31): P(>3) = 0.399\ldots$

$\lambda = (1.9 + 1.32): P(>3) = 0.401\ldots$

$0.399\ldots < 0.4$ & $0.401 > 0.4$

$\therefore$ a reasonable estimate $= 0.4$

6 A bag contains 3 green counters, 3 blue counters and w white counters. Counters are selected at random, one at a time, with replacement, until a white counter is drawn.
The total number of counters selected, including the white counter, is denoted by X.

(i) In the case when $w = 2$,

(a) write down the distribution of X, **[1]**

(b) find $P(3 < X \le 7)$. **[2]**

(ii) In the case when $E(X) = 2$, determine the value of w. **[2]**

(iii) In the case when $w = 2$ and $X = 6$, find the probability that the first five counters drawn alternate in colour. **[2]**

i. a) $X \sim Geo\left(\frac{1}{4}\right)$

b) $P(X>3) = 1-\left(\frac{3}{4}\right)^3$ $\quad P(X\le 7) = 1-\left(\frac{3}{4}\right)^7$

$$P(X\le 7) - P(X>3) = \left(\frac{3}{4}\right)^3 - \left(\frac{3}{4}\right)^7$$

$$= \frac{4725}{16384}$$

$$= 0.288$$

ii. $E(X) = \frac{1}{p}$ $\therefore$ $p = \frac{1}{2}$

$$\Rightarrow w = \frac{1}{2}(3+3+w)$$

$$\frac{1}{2}w = 3$$

$$w = 6$$

iii. $$\frac{2 \times \left(\frac{3}{8}\right)^5 \times \frac{1}{4}}{\left(\frac{3}{4}\right)^5 \times \frac{1}{4}} = \frac{1}{16}$$

7 Sweet pea plants grown using a standard plant food have a mean height of 1.6 m. A new plant food is used for a random sample of 49 randomly chosen plants and the heights, x metres, of this sample can be summarised by the following.

$$n = 49$$
$$\Sigma x = 74.48$$
$$\Sigma x^2 = 120.8896$$

Test, at the 5% significance level, whether, when the new plant food is used, the mean height of sweet pea plants is less than 1.6 m. **[9]**

$\mu = \bar{x} = 1.52$ unbiased

$$\hat{\sigma}^2 = \frac{49}{48}\left(\frac{120.8896}{49} - 1.52^2\right) = 0.16$$

$H_0: \mu = 1.6$ $H_1: \mu < 1.6$

$$P(X < 1.52) = 0.0808$$

$0.0808 > 0.05 \therefore$ do not reject H_0 as insufficient evidence that the mean plant height has decreased

8 A continuous random variable X has probability density function given by

$$f(x) = \begin{cases} 0.8e^{-0.8x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

(i) Find the mean and variance of X. **[4]**

The lifetime of a certain organism is thought to have the same distribution as X. The lifetimes in days of a random sample of 60 specimens of the organism were found. The observed frequencies, together with the expected frequencies correct to 3 decimal places, are given in the table.

Range	$0 \leq x < 1$	$1 \leq x < 2$	$2 \leq x < 3$	$3 \leq x < 4$	$x \geq 4$
Observed	24	22	10	3	1
Expected	33.040	14.846	6.671	2.997	2.446

(ii) Show how the expected frequency for $1 \leq x < 2$ is obtained. **[4]**

(iii) Carry out a goodness of fit test at the 5% significance level. **[7]**

i. $E(X) = \int_0^\infty 0.8xe^{-0.8x}\,dx = 1.25$

$E(X^2) = \int_0^\infty 0.8x^2e^{-0.8x}\,dx = 3.125$

$Var(X) = 3.125 - 1.25^2$
$= 1.5625$

ii. $P(1 \leq x < 2) = \int_1^2 0.8e^{-0.8x}\,dx$

$= 0.247432$ (6 s.f.)

60 Specimens $\Rightarrow$ expected frequency =
$0.247432 \times 60 =$
14.846

H_0: data consistent with distribution

H_1: data not consistent

combine last 2 cells:

O	E	$(O-E)^2/E$
24	33.040	2.4734
22	14.846	3.4474
10	6.6707	1.6613
4	5.4431	0.3826

$$\sum \frac{(O-E)^2}{E} = 7.965$$

$\chi^2_3(0.95) = 7.815$ ← critical value

$7.965 > 7.815$ $\therefore$ reject H_0 as there is sufficient evidence that the data is not consistent with the distribution

9 The continuous random variable X has cumulative distribution function given by

$$\mathrm{F}(x) = \begin{cases} 0 & x < 0, \\ \frac{1}{16}x^2 & 0 \le x \le 4, \\ 1 & x > 4. \end{cases}$$

(i) The random variable Y is defined by $Y = \frac{1}{X^2}$. Find the cumulative distribution function of Y. **[5]**

(ii) Show that $\mathrm{E}(Y)$ is not defined. **[4]**

END OF QUESTION PAPER

i. $P(Y \le y) = P\left(\frac{1}{X^2} \le y\right)$

$= P\left(X \ge \frac{1}{\sqrt{y}}\right)$

$= 1 - F\left(\frac{1}{\sqrt{y}}\right)$

$= \begin{cases} 1 - \frac{1}{16y} & y > \frac{1}{16} \\ 0 & \text{else} \end{cases}$

ii. PDF of y is $\frac{1}{16y^2}$

$\int_{\frac{1}{16}}^{\infty} \frac{y}{16y^2}\,dy = \left[\frac{1}{16}\ln y\right]_{\frac{1}{16}}^{\infty}$

$\ln y$ is undefined as $y \to \infty$