A2 Statistics Practice Paper C
60 Marks

60 Minutes

1. A certain forest contains only trees of a particular species. Dipak wished to take a random sample of 5 trees from the forest. He numbered the trees from 1 to 784. Then, using his calculator, he generated the random digits 14 781 049. Using these digits, Dipak formed 5 three-digit numbers. He took the first, second and third digits, followed by the second, third and fourth digits and so on. In this way he obtained the following list of numbers for his sample.

147 478 781 104 49

a. Explain why Dipak omitted the number 810 from his list. **(2)**

b. Explain why Dipak's sample is not random. **(2)**

The mean height of all trees of this species is known to be 4.2 m. Dipak wishes to test whether the mean height of trees in the forest is less than 4.2 m. He uses a correct method to choose a random sample of 50 trees and finds that their mean height is 4.0 m. It is given that the standard deviation of trees in the forest is 0.8 m.

c. Carry out a hypothesis test at the 2% significance level. **(7)**

(Total marks: 11)

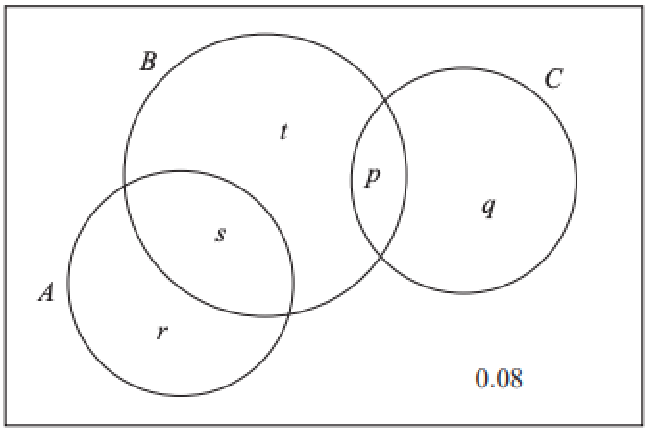
2. The random variable $X\ sim\ N(\mu, 5^2)$ and $P(X < 23) = 0.9192$

a. Find the value of μ **(4)**

b. Write down the value of $P(\mu < X < 23)$. **(1)**

(Total marks: 5)

3. The Venn diagram shows three events A, B and C, where p, q, r, s and t are probabilities.



$P(A) = 0.5$, $P(B) = 0.6$ and $P(C) = 0.25$ and the events B and C are independent.

a. Find the value of p and the value of q. **(2)**

b. Find the value of r. **(2)**

c. Hence write down the value of s and the value of t. **(2)**

d. State, giving a reason, whether or not the events A and B are independent. **(2)**

e. Find $P(B\ I\ A \cup C)$ **(3)**

(Total marks: 11)

4. Bhim and Joe play each other at badminton and for each game, independently of all others, the probability that Bhim loses is 0.2

a. Find the probability that, in 9 games, Bhim loses exactly 3 of the games. **(3)**

b. Find the probability that, in 9 games, Bhim loses fewer than half of the games. **(2)**

c. Calculate the mean and variance for the number of these 60 games that Bhim loses. **(2)**

d. Using a suitable approximation calculate the probability that Bhim loses more than 4 games. **(3)**

(Total marks: 10)

5. The volume of a sample of gas is kept constant. The gas is heated and the pressure, p, is measured at 10 different temperatures, t. The results are summarised below.

$$\Sigma p = 445 \qquad \Sigma p^2 = 38125 \qquad \Sigma t = 240 \qquad \Sigma t^2 = 27520 \qquad \Sigma pt = 26830$$

a. Given that $S_{pp} = 18322.5$, $S_{pt} = 16150$ and $S_{tt} = 21760$, calculate the product moment correlation coefficient.

(2)

b. Give an interpretation of the product moment correlation coefficient. **(1)**

(Total marks: 3)

6. A sample of 13 values of x and y are taken.

For those values, the PMCC is $r = -0.43$.

Using a 5% level of significance, test whether the correlation between x and y is negative. **(2)**

(Total marks: 2)

7. Yuto works in the quality control department of a large company. The time, T minutes, it takes Yuto to analyse a sample is normally distributed with mean 18 minutes and standard deviation 5 minutes.

a. Find the probability that Yuto takes longer than 20 minutes to analyse the next sample. **(3)**

The company has a large store of samples analysed by Yuto with the time taken for each analysis recorded.

Serena is investigating the samples that took Yuto longer than 15 minutes to analyse.

She selects, at random, one of the samples that took Yuto longer than 15 minutes to analyse. **(3)**

b. Find the probability that this sample took Yuto more than 20 minutes to analyse. **(4)**

Serena can identify, in advance, the samples that Yuto can analyse in under 15 minutes and in future she will assign these to someone else.

c. Estimate the median time taken by Yuto to analyse samples in future. **(5)**

(Total marks: 12)

8. The discrete random variable X has probability function

$\boldsymbol{x}$	**0**	**1**	**2**	**3**
$\boldsymbol{P(X=x)}$	$3a$	$2a$	x	b

a. Find $P(X = 2)$ in terms of a and b and complete the table below. **(1)**

Given that the P(X = 2) is half the probability of P(X = 1) and $4a + 3b = 1.6$, find the value of a and b **(5)**

(Total marks: 6)

Total Marks for Paper: 60

Mark Scheme

1a	There are only 784 trees	**M1**
	810 > 784	**A1**
1b	Not dependent	**M1**
	As each number is related to the next	**A1**

1c					
	H_0: $\mu = 4.2$	**B1**	**1.1**	Allow other letters except X or $\overline{X}$	
	H_1: $\mu < 4.2$ where μ is mean height of trees (in the wood)	**B1**	**2.5**	One error, eg undefined μ or 2-tail: B0B1	
	$\overline{X} \sim N(4.2, \frac{0.8^2}{50})$ and $\overline{X} < 4.0$ or $\overline{X} \leq 4.0$	**M1**	**3.3**	Stated or implied Allow $\overline{X} > 4.0$ or $\overline{X} = 4.0$	$\phi^{-1}(0.98)$ $(= 2.054)$
	$P(\overline{X} < 4.0) = 0.038549....$ or 0.039	**A1**	**3.4**	**BC** Allow 0.038 NB 0.038... implies M1A1	$4.2 - 2.054 \times \frac{0.8}{\sqrt{50}}$ $(= 3.968)$
	Compare 0.02	**A1**	**1.1**	dep $P(\overline{X} < 4.0)$ attempted	comp their 3.968 with 4.0
	Do not reject H_0	**M1**	**2.2b**	Allow Accept H_0 dep $P(\overline{X} < 4.0)$ attempted	Can be implied by conclusion
	There is insufficient evidence that mean height of these trees in the wood is less than 4.2m.	**A1f**	**3.5a**	In context, not definite; eg "Mean height not less than 4.2m": A0	

2a	$\frac{23-\mu}{5} = 1.40$	**B1** **M1** **A1**
	$\mu = 16$	**A1**
2b	$P(\mu < X < 23) = p(X < 23) - P(X > \mu)$	

3a	$p = P(BnC) = P(B) \times P(C) = 0.6 \times 0.25 = 0.15$	**M1**
	$q = 0.10$	**A1**
3b	$r = 1 - 0.08 - [P(B) + q]$ $r = 1 - 0.08 - 0.6 - 0.1$	**M1**
	$r = 0.22$	**A1**
3c	$s = P(A) - r = 0.28$	**B1**
	$t = 0.6 \times 0.75 - 0.28$ $t = 0.17$	**B1**
3d	$P(A) \times P(B) = 0.5 \times 0.6 = 0.3$	**M1**
	This is not equal to S, therefore A and B are not independent	**A1**
3e	$\frac{(0.28+0.15)}{0.5+0.25}$	**M1** **A1**
	$= \frac{43}{75}$	**A1**

4a	$X \sim B(9, 0.2)$	**B1**
	$P(X \leq 3) - P(X \leq 2) = 0.9144 - 0.7382$	**B1**
	= 0.1762	**A1**
4b	$P\ (X \leq 4)$	**M1**
	= 0.9804	**A1**
4c	Mean = 3	**B1**
	Variance = 2.85	**B1**
4d	$P(X > 4) = 1 - P(X \leq 4)$	**M1**
	1 – 0.8153	**M1**
	0.1847	**A1**

5a	$r = \frac{16150}{\sqrt{18322.5 \times 21760}}$	**M1**
	$r = 0.8088$	**A1**

	$r = 0.809$	
5b	As the temperate increases the pressure increases	**B1**

6	-0.4762	**M1**
	Accept	**M1**

7a	$[P(T > 20) =]P\left(z > \frac{20-18}{5}\right)$	**M1**
	$P(Z > 0.4) = 1 - 0.6554$	**M1**
	= 0.3446	**A1**
7b	$P(T > 20)I(T > 15)$	**M1**
	$= \frac{0.3445}{0.726}$	**M1** **A1**
	= 0.47485… = 0.475	**A1**
7c	$P(T > d\ I\ T > 15) = 0.5$	**M1**
	P(T > d) = 0.5 x 0.7257	**A1**
	P(T < d) = 0.63715	**M1**
	$\frac{d-18}{5} = 0.35$	**A1**
	$d = 19.754\ ...$ $d = 19.8$	**A1**

8a	P(0) = 3a P(1) = 2a P(2) = a P(3) = b	**B1**
8b	Attempt to create and solve simultaneous equations	**M1**
	$3a + 2a + a + b = 1$	**M1**
	14a = 1.4	**M1**
	$a = 0.1$	**B1**
	$b = 0.4$	**B1**