S2 January 2007

1. (a) Define a statistic.

(2)

A random sample $X_1, X_2, \ldots, X_n$ is taken from a population with unknown mean μ.

(b) For each of the following state whether or not it is a statistic.

(i) $\dfrac{X_1 + X_4}{2}$,

(1)

(ii) $\dfrac{\sum X^2}{n} - \mu^2$.

(1)

---

2. The random variable J has a Poisson distribution with mean 4.

(a) Find $P(J \geqslant 10)$.

(2)

The random variable K has a binomial distribution with parameters $n = 25$, $p = 0.27$.

(b) Find $P(K \leqslant 1)$.

(3)

---

3. For a particular type of plant 45% have white flowers and the remainder have coloured flowers. Gardenmania sells plants in batches of 12. A batch is selected at random.

Calculate the probability that this batch contains

(a) exactly 5 plants with white flowers,

(3)

(b) more plants with white flowers than coloured ones.

(2)

Gardenmania takes a random sample of 10 batches of plants.

(c) Find the probability that exactly 3 of these batches contain more plants with white flowers than coloured ones.

(3)

Due to an increasing demand for these plants by large companies, Gardenmania decides to sell them in batches of 50.

(d) Use a suitable approximation to calculate the probability that a batch of 50 plants contains more than 25 plants with white flowers.

(7)

---

4. (a) State the condition under which the normal distribution may be used as an approximation to the Poisson distribution.
(1)

(b) Explain why a continuity correction must be incorporated when using the normal distribution as an approximation to the Poisson distribution.
(1)

A company has yachts that can only be hired for a week at a time. All hiring starts on a Saturday.
During the winter the mean number of yachts hired per week is 5.

(c) Calculate the probability that fewer than 3 yachts are hired on a particular Saturday in winter.
(2)

During the summer the mean number of yachts hired per week increases to 25.
The company has only 30 yachts for hire.

(d) Using a suitable approximation find the probability that the demand for yachts cannot be met on a particular Saturday in the summer.
(6)

In the summer there are 16 Saturdays on which a yacht can be hired.

(e) Estimate the number of Saturdays in the summer that the company will not be able to meet the demand for yachts.
(2)

---

5. The continuous random variable X is uniformly distributed over the interval $\alpha < x < \beta$.

(a) Write down the probability density function of X, for all x.
(2)

(b) Given that $\mathrm{E}(X) = 2$ and $\mathrm{P}(X < 3) = \frac{5}{8}$ find the value of α and the value of β.
(4)

A gardener has wire cutters and a piece of wire 150 cm long which has a ring attached at one end. The gardener cuts the wire, at a randomly chosen point, into 2 pieces. The length, in cm, of the piece of wire with the ring on it is represented by the random variable X. Find

(c) $\mathrm{E}(X)$,
(1)

(d) the standard deviation of X,
(2)

(e) the probability that the shorter piece of wire is at most 30 cm long.
(3)

---

6. Past records from a large supermarket show that 20% of people who buy chocolate bars buy the family size bar. On one particular day a random sample of 30 people was taken from those that had bought chocolate bars and 2 of them were found to have bought a family size bar.

(a) Test at the 5% significance level, whether or not the proportion p, of people who bought a family size bar of chocolate that day had decreased. State your hypotheses clearly.

(6)

The manager of the supermarket thinks that the probability of a person buying a gigantic chocolate bar is only 0.02. To test whether this hypothesis is true the manager decides to take a random sample of 200 people who bought chocolate bars.

(b) Find the critical region that would enable the manager to test whether or not there is evidence that the probability is different from 0.02. The probability of each tail should be as close to 2.5% as possible.

(6)

(c) Write down the significance level of this test.

(1)

---

7. The continuous random variable X has cumulative distribution function

$$F(x) = \begin{cases} 0, & x < 0, \\ 2x^2 - x^3, & 0 \leqslant x \leqslant 1, \\ 1, & x > 1. \end{cases}$$

(a) Find $P(X > 0.3)$.

(2)

(b) Verify that the median value of X lies between $x = 0.59$ and $x = 0.60$.

(3)

(c) Find the probability density function $f(x)$.

(2)

(d) Evaluate $E(X)$.

(3)

(e) Find the mode of X.

(2)

(f) Comment on the skewness of X. Justify your answer.

(2)

TOTAL FOR PAPER: 75 MARKS

January 2007
6684 Statistics S2
Mark Scheme

Question Number	Scheme	Marks
1. (a)	A random variable; function of known observations (from a population). data OK	B1 B1 (2)
(b) (i)	Yes	B1 (1)
(ii)	No	B1 (1)
		Total 4
2. (a)	$P(J \geq 10) = 1 - P(J \leq 9)$ or $= 1 - P(J < 10)$	M1
	$= 1 - 0.9919$ implies method	
	$= 0.0081$ awrt 0.0081	A1 (2)
(b)	$P(K \leq 1) = P(K = 0) + P(K = 1)$ both, implied below even with ‘25’ missing	M1
	$= (0.73)^{25} + 25(0.73)^{24}(0.27)$ clear attempt at ‘25’ required	M1
	$= 0.00392$ awrt 0.0039 implies M	A1 (3)
		Total 5

Question Number	Scheme	Marks
3. (a)	Let W represent the number of white plants.	
	$W \sim B(12,0.45)$ use of	B1
	$P(W=5) = P(W \le 5) - P(W \le 4)$ ${}^{12}C_5 0.45^5 0.55^7$ or equivalent award B1M1	M1
	$= 0.5269 - 0.3044$ values from correct table implies B	
	$= 0.2225$ awrt 0.222(5)	A1
		(3)
(b)	$P(W \ge 7) = 1 - P(W \le 6)$ or $=1-P(W<7)$	M1
	$= 1 - 0.7393$ implies method	
	$= 0.2607$ awrt 0.261	A1
		(2)
(c)	P(3 contain more white than coloured)$= \frac{10!}{3!7!}(0.2607)^3(1-0.2607)^7$ use of B,n=10	M1A1∫
	$= 0.256654\ldots$ awrt 0.257	A1
		(3)
(d)	mean = np = 22.5 ; var = npq = 12.375	B1B1
	$P(W>25) \approx P\left(Z > \frac{25.5-22.5}{\sqrt{12.375}}\right)$ ± standardise with σ and μ; ±0.5 c.c.	M1;M1
	$\approx P(Z > 0.8528..)$ awrt 0.85	A1
	$\approx 1 - 0.8023$ 'one minus'	M1
	≈ 0.1977 awrt 0.197 or 0.198	A1
		(7)
		Total 15

Question Number	Scheme		Marks
4. (a)	$\lambda > 10$ or large	μ ok	**B1** **(1)**
(b)	The Poisson is discrete and the normal is continuous.		**B1** **(1)**
(c)	Let Y represent the number of yachts hired in winter		
	$P(Y<3) = P(Y \leqslant 2)$	$P(Y \leqslant 2)$ & Po(5)	**M1**
	$= 0.1247$	awrt 0.125	**A1** **(2)**
(d)	Let X represent the number of yachts hired in summer $X \sim Po(25)$.		
	N(25,25)	all correct, can be implied by standardisation below	**B1**
	$P(X > 30) \approx P\left(Z > \frac{30.5 - 25}{5}\right)$	± standardise with 25 & 5; ±0.5 c.c.	**M1;M1**
	$\approx P(Z > 1.1)$	1.1	**A1**
	$\approx 1 - 0.8643$	'one minus'	**M1**
	≈ 0.1357	awrt 0.136	**A1** **(6)**
(e)	no. of weeks $= 0.1357 \times 16$	ANS (d)x16	**M1**
	$= 2.17$ or 2 or 3	ans>16 M0A0	**A1∫** **(2)**
			Total 12

Question Number	Scheme		Marks
5.			
(a)	$f(x) = \begin{cases} \dfrac{1}{\beta-\alpha}, & \alpha < x < \beta, \\ 0, & \text{otherwise.} \end{cases}$	function including inequality, 0 otherwise	**B1,B1**
			(2)
(b)	$\dfrac{\alpha+\beta}{2} = 2$, $\dfrac{3-\alpha}{\beta-\alpha} = \dfrac{5}{8}$	or equivalent	**B1,B1**
	$\alpha + \beta = 4$ $3\alpha + 5\beta = 24$		
	$3(4-\beta) + 5\beta = 24$ $2\beta = 12$ $\beta = 6$	attempt to solve 2 eqns	**M1**
	$\alpha = -2$	both	**A1**
			(4)
(c)	$E(X) = \dfrac{150+0}{2} = 75\,\text{cm}$	75	**B1**
			(1)
(d)	Standard deviation = $\sqrt{\dfrac{1}{12}(150-0)^2}$		**M1**
	$= 43.30127\ldots\text{cm}$	$25\sqrt{3}$ or awrt 43.3	**A1**
			(2)
(e)	$P(X < 30) + P(X > 120) = \dfrac{30}{150} + \dfrac{30}{150}$	1st or at least one fraction, + or double	**M1,M1**
	$= \dfrac{60}{150}$ or $\dfrac{2}{5}$ or 0.4 or equivalent fraction		**A1**
			(3)
			Total 12

Question Number	Scheme	Marks
6. (a)	$H_0 : p = 0.20$, $H_1: p < 0.20$	**B1,B1**
	Let X represent the number of people buying family size bar. $X \sim B(30, 0.20)$	
	$P(X \leqslant 2) = 0.0442$ or $P(X \leqslant 2) = 0.0442$ awrt 0.044 $P(X \leqslant 3) = 0.1227$ CR $X \leqslant 2$	**M1A1**
	0.0442 < 5%, so significant. Significant	**M1**
	There is evidence that the no. of family size bars sold is lower than usual.	**A1** **(6)**
(b)	$H_0 : p = 0.02$, $H_1: p \neq 0.02$ $\lambda = 4$ etc ok both	**B1**
	Let Y represent the number of gigantic bars sold.	
	$Y \sim B(200, 0.02) \Rightarrow Y \sim Po(4)$ can be implied below	**M1**
	$P(Y = 0) = \mathbf{0.0183}$ and $P(Y \leqslant 8) = \mathbf{0.9786} \Rightarrow P(Y \geqslant 9) = \mathbf{0.0214}$ first, either	**B1,B1**
	Critical region $Y = 0 \cup Y \geqslant 9$ $Y \leq 0$ ok	**B1,B1**
	N.B. Accept exact Bin: 0.0176 and 0.0202	
(c)	Significance level = 0.0183 + 0.0214 = 0.0397 awrt 0.04	**B1** **(1)**
		Total 13

Question Number	Scheme	Marks
7. (a)	$1 - F(0.3) = 1 - (2 \times 0.3^2 - 0.3^3)$ $= 0.847$ ‘one minus’ required	M1 A1 (2)
(b)	$F(0.60) = 0.5040$ $F(0.59) = 0.4908$ both required awrt 0.5, 0.49	M1A1
	0.5 lies between therefore median value lies between 0.59 and 0.60.	B1 (3)
(c)	$f(x) = \begin{cases} -3x^2 + 4x, & 0 \le x \le 1, \\ 0\ , & \text{otherwise.} \end{cases}$ attempt to differentiate, all correct	M1A1 (2)
(d)	$\int_0^1 x\text{f}(x)\text{d}x = \int_0^1 -3x^3 + 4x^2 \text{d}x$ attempt to integrate $x\text{f}(x)$	M1
	$= \left[\frac{-3x^4}{4} + \frac{4x^3}{3}\right]_0^1$ sub in limits	M1
	$= \frac{7}{12}$ or $0.58\dot{3}$ or 0.583 or equivalent fraction	A1 (3)
(e)	$\frac{\text{df}(x)}{\text{d}x} = -6x + 4 = 0$ attempt to differentiate f(x) and equate to 0	M1
	$x = \frac{2}{3}$ or $0.\dot{6}$ or 0.667	A1 (2)
(f)	mean < median < mode, therefore negative skew. Any pair, cao	B1,B1 (2)
		Total 14