Centre No.						Paper Reference							Surname	Initial(s)
Candidate No.						6	6	8	4	/	0	1	Signature MODEL ANSWERS	

Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Advanced/Advanced Subsidiary

Monday 22 June 2015 – Morning

Time: 1 hour 30 minutes

Examiner's use only

Team Leader's use only

Question Number	Leave Blank
1	
2	
3	
4	
5	
6	
7	
Total	

Materials required for examination
Mathematical Formulae (Pink)

Items included with question papers
Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation or symbolic differentiation/integration, or have retrievable mathematical formulae stored in them.

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initials and signature.
Check that you have the correct question paper.
Answer ALL the questions.
You must write your answer to each question in the space following the question.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
There are 7 questions in this question paper. The total mark for this paper is 75.
There are 24 pages in this question paper. Any blank pages are indicated.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You should show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

Printer's Log. No.
P44846A
W850/R6684/57570 5/1/1/

P44846A0124

Turn over

PEARSON

Leave blank

1. In a survey it is found that barn owls occur randomly at a rate of 9 per 1000 km^2.

(a) Find the probability that in a randomly selected area of 1000 km^2 there are at least 10 barn owls.
(2)

(b) Find the probability that in a randomly selected area of 200 km^2 there are exactly 2 barn owls.
(3)

(c) Using a suitable approximation, find the probability that in a randomly selected area of 50 000 km^2 there are at least 470 barn owls.
(6)

a) Random rate of 9/1000 km^2 $\Rightarrow$ Poisson, $\lambda = 9$

X is the random variable for the distribution of barn owls per 1000 km^2.

$X \sim Po(9)$ Using tables.

$P(X \geq 10) = 1 - P(X \leq 9) = 1 - 0.5874 = \underline{0.4126}$ ✓

b) Y is the random variable for the distribution of barn owls per 200 km^2. $(200\text{km}^2 / 1000\text{km}^2 = \tfrac{1}{5})$

$Y \sim P_o(9/5) \sim P(1.8)$

$P(X=x) = \dfrac{e^{-\lambda}\lambda^{x}}{x!}$

$P(X=2) = (e^{-1.8}\,1.8^2)/2! = \underline{0.2678}$ ✓

c) For 50000 km^2, $\lambda \to 9 \times 50$, ie. $P_o(450)$

Normal approximation for random variable Q:

$Q \approx N(\mu, \sigma^2) \approx N(450, 450)$

$P(Q \geq 470) = 1 - P(Q \leq 469)$
$= 1 - P(Q < 469.5)$ continuity correction

1. c) (continued)

$$P(Q \geq 470) = 1 - P\left(Z < \frac{469.5 - 450}{\sqrt{450}}\right)$$

$$= 1 - P(Z < 0.9192)$$

$$= 1 - P(Z < 0.92)$$

$$= 1 - 0.8212 \quad \text{(using tables)}$$

$$\underline{= 0.1788} \quad \checkmark$$

constant probability

Leave blank

2. The proportion of houses in Radville which are unable to receive digital radio is 25%. In a survey of a random sample of 30 houses taken from Radville, the number, X, of houses which are unable to receive digital radio is recorded.

(a) Find $P(5 \leqslant X < 11)$ fixed number of trials

(3)

A radio company claims that a new transmitter set up in Radville will reduce the proportion of houses which are unable to receive digital radio. After the new transmitter has been set up, a random sample of 15 houses is taken, of which 1 house is unable to receive digital radio.

(b) Test, at the 10% level of significance, the radio company's claim. State your hypotheses clearly.

(5)

a) X is the random variable representing the number of houses which are unable to receive digital radio.

$X \sim B(n, p) \sim (30, 0.25)$

$P(5 \leq X < 11) = P(X \leq 10) - P(X \leq 4)$

$= 0.8943 - 0.0979 = \underline{0.7964}$ ✓

b) $Y \sim B(15, 0.25)$

$H_0 : p = 0.25$ At 10% level of significance (α)

Reject if $P(Y \leq y) \leq \alpha$

$H_1 : p < 0.25$ So one-tailed.

$P(Y \leq 1) = \underline{0.0802}$ (using tables) ✓

As $0.0802 < 0.1$, H_0 is rejected. ✓

There is sufficient evidence to say that the new transmitter is better, their claim is true. ✓

Leave blank

3. A random variable X has probability density function given by

$$f(x) = \begin{cases} kx^2 & 0 \leqslant x \leqslant 2 \\ k\left(1 - \dfrac{x}{6}\right) & 2 < x \leqslant 6 \\ 0 & \text{otherwise} \end{cases}$$

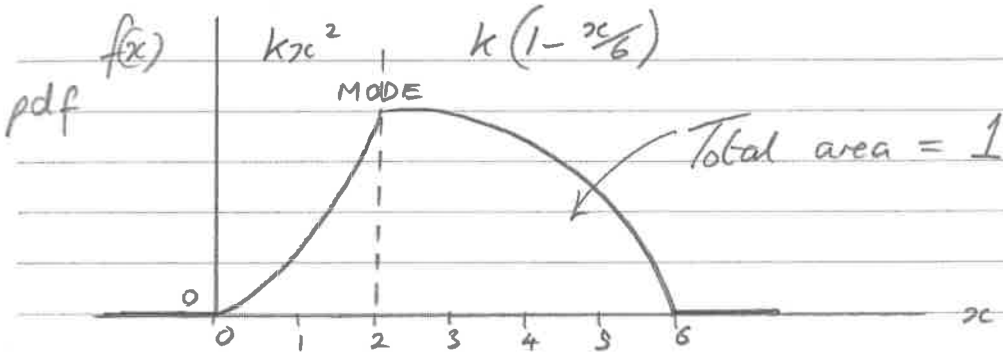
where k is a constant.

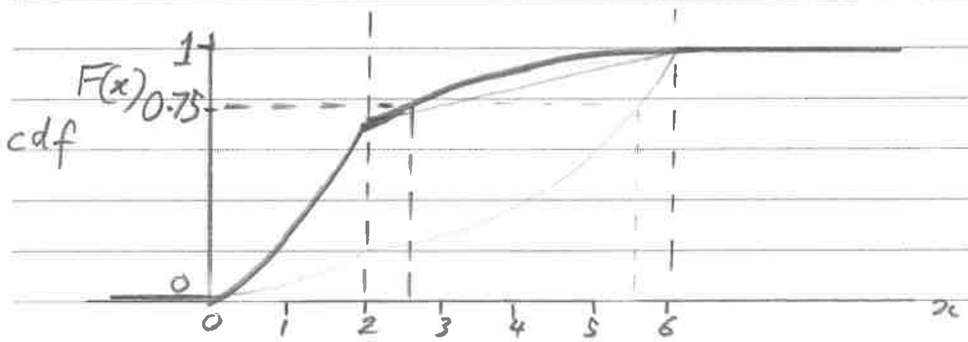
(a) Show that $k = \frac{1}{4}$ **(4)**

(b) Write down the mode of X. **(1)**

(c) Specify fully the cumulative distribution function F(x). **(5)**

(d) Find the upper quartile of X. **(4)**





a) $\int_1^6 f(x) \; dx = 1$ Area under curve = 1

$= \int_0^2 kx^2 \, dx + \int_2^6 k\left(1 - x/6\right) dx$

$= k\int_0^2 x^2 \, dx + k\int_2^6 1 - x/6 \; dx$

P 4 4 8 4 6 A 0 6 2 4

3. c)

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{x^3}{12} & , \quad 0 \leq x \leq 2 \\ \frac{x}{4} - \frac{x^2}{48} + \frac{1}{4} & , \quad 2 < x \leq 6 \\ 1 & , \quad x > 6 \end{cases}$$ ✓

d) Upper quartile:

$F(x) = 0.75$ (See graph)

$$\therefore \frac{x}{4} - \frac{x^2}{48} + \frac{1}{4} = \frac{3}{4}$$

$$x - \frac{x^2}{12} + 1 = 3$$

$$12x - x^2 + 12 - 36 = 0$$

$$x^2 - 12x + 24 = 0$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$x = \frac{-(-12) \pm \sqrt{(-12)^2 - 4(1)(24)}}{2(1)}$$

$\underline{x = 6 - 2\sqrt{3}}$ ✓ (reject negative)

Leave blank

Question 3 continued

a) $$1 = k\left[\frac{x^3}{3}\right]_0^2 + k\left[x - \frac{x^2}{12}\right]_2^6$$

$$1 = k\left(\frac{2^3}{3}\right) + k\left[\left(6-\frac{6^2}{12}\right)-\left(2-\frac{2^2}{12}\right)\right]$$

$$1 = k\left(\frac{8}{3} + \frac{9}{3} - \frac{5}{3}\right) = k\,\frac{12}{3}$$

$$\Rightarrow \underline{k = \tfrac{1}{4}}$$

✓

b) Mode = most common = highest point on pdf curve ✓
$\therefore \underline{x = 2}$ (see graph)

c) $0 \leq x \leq 2$

$$F(x) = \int_0^{x_0} kx^2\,dx = \frac{kx^3}{3}$$ Using x_0 as a temporary variable.

$2 < x \leq 6$

$$F(x) = F(2) + \int_2^{x_0} k\left(1-\frac{x}{6}\right)dx$$

$$= \frac{k(2)^3}{3} + k\left[x - \frac{x^2}{12}\right]_2^{x_0}$$

$$= k\frac{8}{3} + k\left(x_0 - \frac{x_0^2}{12} - 2 - -\frac{2^2}{12}\right)$$

$$\underline{= \frac{x}{4} - \frac{x^2}{48} + \frac{1}{4}}$$ (Using $k=\frac{1}{4}$) ✓

Again, using x_0 as a temporary variable.

Leave blank

4. The continuous random variable L represents the error, in metres, made when a machine cuts poles to a target length. The distribution of L is a continuous uniform distribution over the interval [0, 0.5]

(a) Find $\mathrm{P}(L < 0.4)$.
(1)

(b) Write down $\mathrm{E}(L)$.
(1)

(c) Calculate $\mathrm{Var}(L)$.
(2)

A random sample of 30 poles cut by this machine is taken.

(d) Find the probability that fewer than 4 poles have an error of more than 0.4 metres from the target length.
(3)

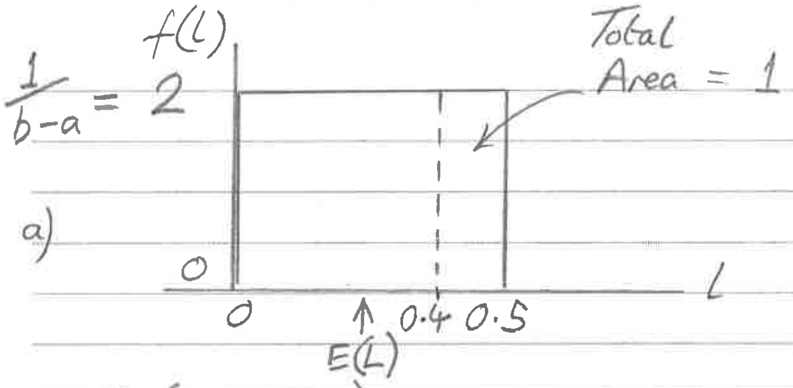
When a new machine cuts poles to a target length, the error, X metres, is modelled by the cumulative distribution function $\mathrm{F}(x)$ where

$$\mathrm{F}(x) = \begin{cases} 0 & x < 0 \\ 4x - 4x^2 & 0 \leqslant x \leqslant 0.5 \\ 1 & \text{otherwise} \end{cases}$$

(e) Using this model, find $\mathrm{P}(X > 0.4)$
(2)

A random sample of 100 poles cut by this new machine is taken.

(f) Using a suitable approximation, find the probability that at least 8 of these poles have an error of more than 0.4 metres.
(3)



$P(L < 0.4) = 0.4 \times 2 = \underline{0.8}$

(area of rectangle)

P 4 4 8 4 6 A 0 1 0 2 4

4. f) From question: $n = 100$, $p = 0.04$

- n is too great to use Binomial tables
- As n is large and p is small, use a Poisson approximation.

- Y is the random variable for the distribution of poles with an error greater than 0.4 metres.

$$\lambda \simeq np$$
$$\lambda \simeq 100 \times 0.04 = 4$$

$$\therefore Y \sim Po(4)$$

$$P(Y \geq 8) = 1 - P(X \leq 7)$$
$$= 1 - 0.9489 \quad \text{(using Poisson tables)}$$
$$= \underline{0.0511}$$

Leave blank

Question 4 continued

b) $E(L) = \text{midpoint} = \dfrac{a+b}{2} = 0.25$

c) $Var(L) = \dfrac{(b-a)^2}{12}$

$= \dfrac{(0.5-0)^2}{12} = \dfrac{1}{48}$

d) If $P(L<0.4) = 0.8$

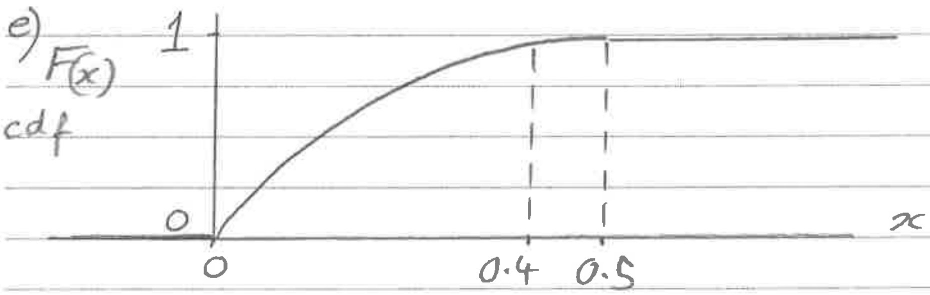
Then $P(L>0.4) = 1-0.8 = 0.2$

Using a binomial approximation:

$Y \sim B(30, 0.2)$

$P(Y<4) = P(Y \leq 3) = 0.1227$ (using tables)

e)



$P(X>0.4) = 1 - P(X<0.4)$

$= 1 - F(0.4)$

$= 1 - \left(4(0.4) - 4(0.4)^2\right)$

$= 1 - (1.6 - 0.64)$

$= 0.04$

Leave blank

5. *Liftsforall* claims that the lift they maintain in a block of flats breaks down at random at a mean rate of 4 times per month. To test this, the number of times the lift breaks down in a month is recorded.

(a) Using a 5% level of significance, find the critical region for a two-tailed test of the null hypothesis that 'the mean rate at which the lift breaks down is 4 times per month'. The probability of rejection in each of the tails should be as close to 2.5% as possible.
(3)

Over a randomly selected 1 month period the lift broke down 3 times.

(b) Test, at the 5% level of significance, whether *Liftsforall*'s claim is correct. State your hypotheses clearly.
(2)

(c) State the actual significance level of this test.
(1)

The residents in the block of flats have a maintenance contract with *Liftsforall*. The residents pay *Liftsforall* £500 for every quarter (3 months) in which there are at most 3 breakdowns. If there are 4 or more breakdowns in a quarter then the residents do not pay for that quarter.

Liftsforall installs a new lift in the block of flats.

Given that the new lift breaks down at a mean rate of 2 times per month,

(d) find the probability that the residents do not pay more than £500 to *Liftsforall* in the next year.
(6)

a) Let X be the random variable for the number of breakdowns in 1 month.

Mean rate of 4 breakdowns/month $\Rightarrow$ Poisson, $\lambda=4$

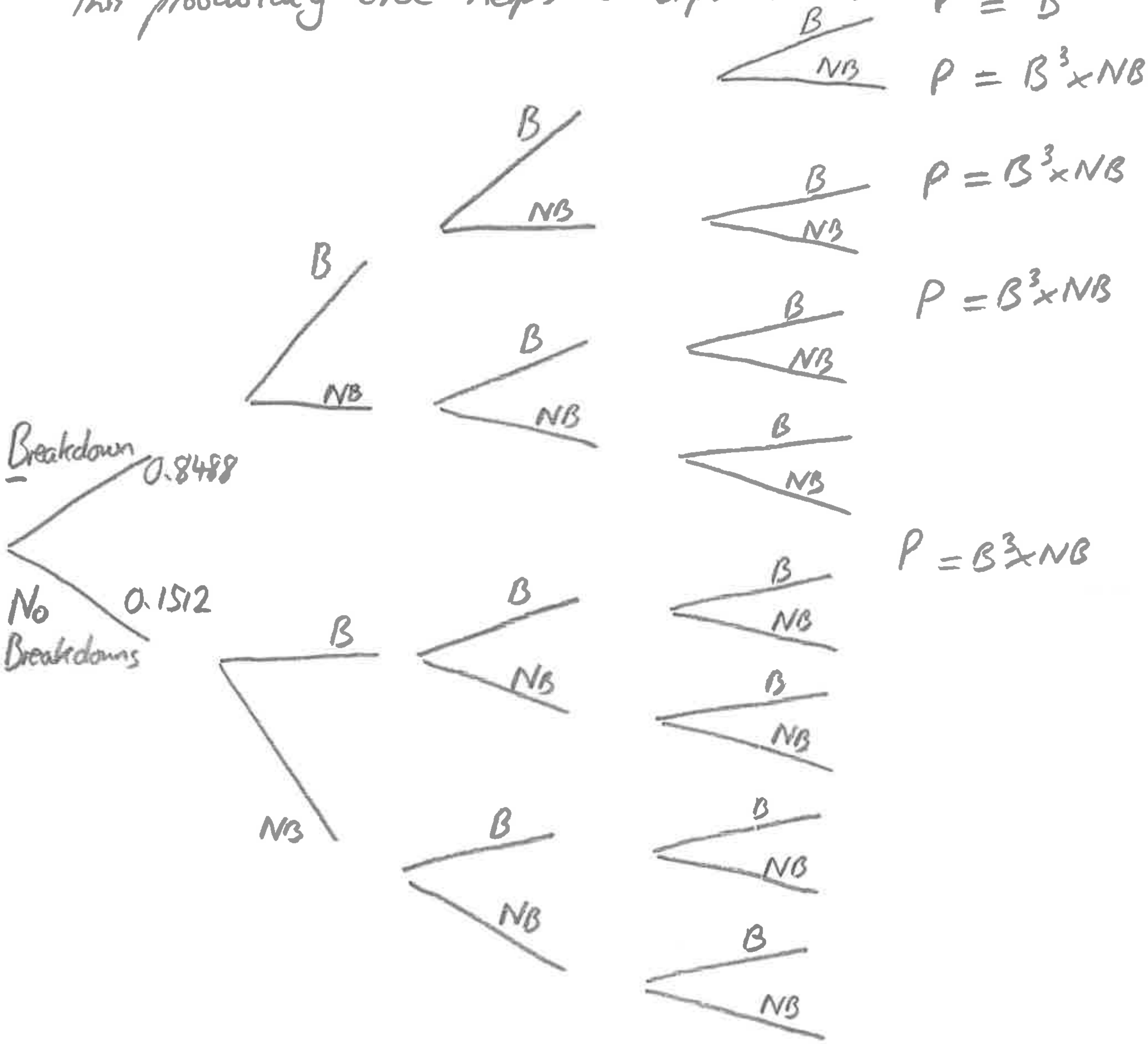
$X \sim Po(4)$

$H_0 : \lambda = 4$

$H_1 : \lambda \neq 4$ $\Rightarrow$ 2-tailed test

5. d) The probability that the residents do not pay more than £500 to Liftsforall means either 4 breakdowns or 3 in a year.

This probability tree helps to explain it:



$P(\text{3 or 4 breakdowns in 1 year})$

$= B^4 + B^3 \times NB \times 4$

$= 0.8488^4 + 0.8488^3 \times 0.1512 \times 4$

$= 0.8889\ldots \quad \underline{\simeq 0.889}$

Leave blank

Question 5 continued

Using Poisson tables, $\lambda = 4$:

$P(X \leq 0) = 0.0183$ *closest to 2.5%
$P(X \leq 1) = 0.0916$

$P(X \geq 8) = 1 - P(X \leq 7) = 1 - 0.9489 = 0.0511$
$P(X \geq 9) = 1 - P(X \leq 8) = 1 - 0.9786 = 0.0214$ *
*closest to 2.5%

Let c_1 and c_2 be the critical values.
$\therefore c_1 = 0$
$\therefore c_2 = 9$

$\therefore$ Critical region: $X \leq 0$ and $X \geq 9$

b) $H_0: \lambda = 4$
$H_1: \lambda \neq 4$ $\Rightarrow$ 2-tailed test
Significance level 5% $\Rightarrow$ 2.5% per tail.

As $X = 3$ is not in the critical region, there is sufficient evidence to reject H_0, there is evidence that Liftsforall's claim is true.

c) Significance level of test $= 0.0183 + 0.0214$
$= 0.0397$

d) $\lambda = 2$ (per month)

Let Y be the random variable for the number of breakdowns per quarter (3 months). $\Rightarrow \lambda = 6$

$Y \sim Po(6)$

$P(\lambda \geq 4) = 1 - P(Y \leq 3) = 1 - 0.1512$
$= 0.8488$

P 4 4 8 4 6 A 0 1 5 2 4

Turn over

Leave blank

6. A continuous random variable X has probability density function f(x) where

$$f(x) = \begin{cases} kx^n & 0 \leqslant x \leqslant 1 \\ 0 & \text{otherwise} \end{cases}$$

where k and n are positive integers.

(a) Find k in terms of n. **(3)**

(b) Find E(X) in terms of n. **(3)**

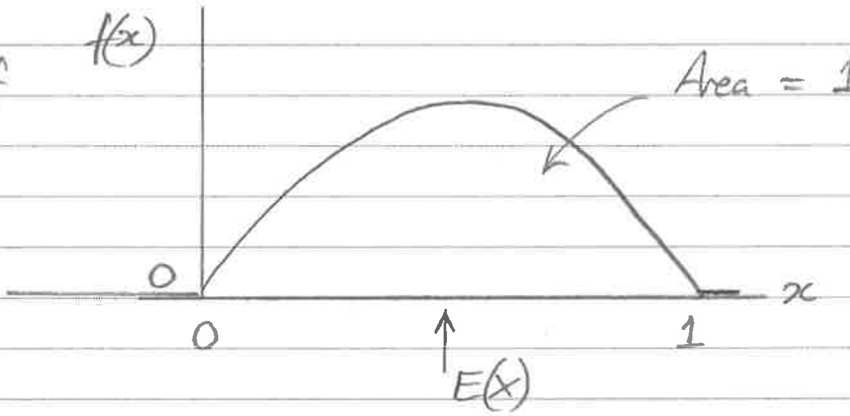
(c) Find E(X^2) in terms of n. **(2)**

Given that $n = 2$

(d) find Var(3X). **(3)**

pdf

a)



$$\int_0^1 f(x)\,dx = 1$$

$$\left[k\frac{x^{n+1}}{n+1}\right]_0^1 = 1$$

$$\frac{k(1)^{n+1}}{n+1} = 1$$

$$\underline{k = n+1}$$

✓

P 4 4 8 4 6 A 0 1 8 2 4

Leave blank

Question 6 continued

b) $E(X) = \int_{-\infty}^{\infty} x \; f(x) \; dx$

$= \int_0^1 x \, kx^n \; dx$

$= \int_0^1 k \, x^{n+1} \; dx$

$= \left[k \frac{x^{n+2}}{n+2} \right]_0^1 = \frac{k}{n+2} = \underline{\frac{n+1}{n+2}}$ ✓

c) $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) \; dx$

$= \int_0^1 x^2 kx^n \; dx = \left[\frac{kx^{n+3}}{n+3} \right] = \underline{\frac{n+1}{n+3}}$ ✓

d) $Var(X) = E(X^2) - (E(X))^2$

$Var(3X) = 3^2 \, Var(X)$

$= 9\left[\left(\frac{n+1}{n+3}\right) - \left(\frac{n+1}{n+2}\right)^2 \right]$

$= 9\left[\left(\frac{2+1}{2+3}\right) - \left(\frac{2+1}{2+2}\right)^2 \right]$

$= 9\left(\frac{3}{5} - \frac{9}{16}\right) = 9\left(\frac{48-45}{80}\right)$

$= \underline{\frac{27}{80}}$ ✓

Leave blank

7. A bag contains a large number of 10p, 20p and 50p coins in the ratio 1 : 2 : 2

A random sample of 3 coins is taken from the bag.

Find the sampling distribution of the median of these samples.

(7)

Ratio	1	: 2	: 2
X	10	20	50
$P(X=x)$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$

Possible samples			(m) Median
10, 10, 10			10
10, 10, 20	10, 20, 10	20, 10, 10	10
10, 10, 50	10, 50, 10	50, 10, 10	10
10, 20, 20	20, 10, 20	20, 20, 10	20
20, 20, 20			20
20, 20, 50	20, 50, 20	50, 20, 20	20
10, 20, 50	10, 50, 20	20, 10, 50	20
50, 10, 20	50, 20, 10	20, 50, 10	20
50, 50, 50			50
10, 50, 50	50, 10, 50	50, 50, 10	50
50, 50, 20	50, 20, 50	20, 50, 50	50

$P(m=10) = \left(\frac{1}{5}\right)^3 + 3\left(\frac{1}{5}\right)^2\left(\frac{2}{5}\right) + 3\left(\frac{1}{5}\right)^2\left(\frac{2}{5}\right)$

$= \frac{1 + 6 + 6}{125} = \frac{13}{125}$

$P(m=20) = 3\left(\frac{1}{5}\right)\left(\frac{2}{5}\right)^2 + \left(\frac{2}{5}\right)^3 + \left(\frac{1}{5}\right)\left(\frac{2}{5}\right)\left(\frac{2}{5}\right)6 +$

$\left(\frac{2}{5}\right)^2\left(\frac{2}{5}\right)3$

Leave blank

Question 7 continued

$$P(m=20) = \frac{12+8+24+24}{125} = \frac{68}{125}$$

$$P(m=50) = \left(\tfrac{2}{5}\right)^3 + \left(\tfrac{1}{5}\right)\left(\tfrac{2}{5}\right)^2 3 + \tfrac{2}{5}\left(\tfrac{2}{5}\right)^2 3$$

$$= \frac{8+12+24}{125} = \frac{44}{125}$$

m	10	20	50
$P(M=m)$	$\frac{13}{125}$	$\frac{68}{125}$	$\frac{44}{125}$