Paper Reference(s)

6684/01

Edexcel GCE

Statistics S2

Bronze Level B3

Time: 1 hour 30 minutes

Materials required for examination papers	**Items included with question**
Mathematical Formulae (Green)	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, differentiation and integration, or have retrievable mathematical formulas stored in them.

Instructions to Candidates

Write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S2), the paper reference (6684), your surname, initials and signature.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
There are 8 questions in this question paper. The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner. Answers without working may gain no credit.

Suggested grade boundaries for this paper:

A*	A	B	C	D	E
73	68	61	54	47	41

1. A bag contains a large number of counters of which 15% are coloured red. A random sample of 30 counters is selected and the number of red counters is recorded.

(a) Find the probability of no more than 6 red counters in this sample.

(2)

A second random sample of 30 counters is selected and the number of red counters is recorded.

(b) Using a Poisson approximation, estimate the probability that the total number of red counters in the combined sample of size 60 is less than 13.

(3)

---

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

Find the probability that, in 9 games, Bhim loses

(a) exactly 3 of the games,

(3)

(b) fewer than half of the games.

(2)

Bhim attends coaching sessions for 2 months. After completing the coaching, the probability that he loses each game, independently of all others, is 0.05.

Bhim and Joe agree to play a further 60 games.

(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)

---

3. The random variable X has a continuous uniform distribution on $[a, b]$ where a and b are positive numbers.

Given that $\mathrm{E}(X) = 23$ and $\mathrm{Var}(X) = 75$,

(a) find the value of a and the value of b.

(6)

Given that $\mathrm{P}(X > c) = 0.32$,

(b) find $\mathrm{P}(23 < X < c)$.

(2)

---

4. The continuous random variable Y has cumulative distribution function F(y) given by

$$\mathrm{F}(y) = \begin{cases} 0 & y < 1 \\ k(y^4 + y^2 - 2) & 1 \le y \le 2 \\ 1 & y > 2 \end{cases}$$

(a) Show that $k = \dfrac{1}{18}$

(2)

(b) Find P(Y > 1.5).

(2)

(c) Specify fully the probability density function f(y).

(3)

---

5. The probability of an electrical component being defective is 0.075.
The component is supplied in boxes of 120.

(a) Using a suitable approximation, estimate the probability that there are more than 3 defective components in a box.

(5)

A retailer buys 2 boxes of components.

(b) Estimate the probability that there are at least 4 defective components in each box.

(2)

---

6. The three independent random variables A, B and C each has a continuous uniform distribution over the interval [0, 5].

(*a*) Find $\mathrm{P}(A > 3)$.
(1)

(*b*) Find the probability that A, B and C are all greater than 3.
(2)

The random variable Y represents the maximum value of A, B and C.

The cumulative distribution function of Y is

$$\mathrm{F}(y) = \begin{cases} 0, & y < 0 \\ \dfrac{y^3}{125}, & 0 \le y \le 5 \\ 1, & y > 5 \end{cases}$$

(*c*) Find the probability density function of Y.
(2)

(*d*) Sketch the probability density function of Y.
(2)

(*e*) Write down the mode of Y.
(1)

(*f*) Find $\mathrm{E}(Y)$.
(3)

(*g*) Find $\mathrm{P}(Y > 3)$.
(2)

---

7. (*a*) (i) Write down two conditions for $X \sim \text{Bin}(n, p)$ to be approximated by a normal distribution $Y \sim \text{N}(\mu, \sigma^2)$.

(2)

(ii) Write down the mean and variance of this normal approximation in terms of n and p.

(2)

A factory manufactures 2000 DVDs every day. It is known that 3% of DVDs are faulty.

(*b*) Using a normal approximation, estimate the probability that at least 40 faulty DVDs are produced in one day.

(5)

The quality control system in the factory identifies and destroys every faulty DVD at the end of the manufacturing process. It costs £0.70 to manufacture a DVD and the factory sells non-faulty DVDs for £11.

(*c*) Find the expected profit made by the factory per day.

(3)

---

8. The continuous random variable X has probability density function f(x) given by

$$\text{f}(x) = \begin{cases} 2(x-2) & 2 \leq x \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

(*a*) Sketch f(x) for all values of x.

(3)

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

(1)

Find

(*c*) E(X),

(3)

(*d*) the median of X.

(4)

(*e*) Comment on the skewness of this distribution. Give a reason for your answer.

(2)

TOTAL FOR PAPER: 75 MARKS

END

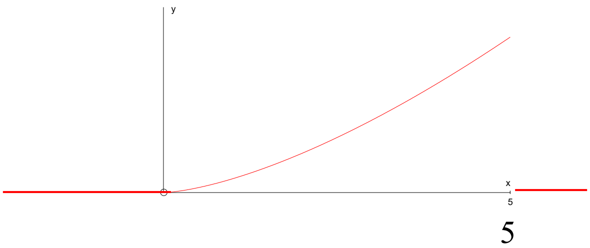
Question Number	Scheme	Marks
1. (a)	$[X \sim B(30, 0.15)]$	
	$P(X \leq 6)$, = 0.8474 awrt 0.847	M1, A1 (2)
(b)	$Y \sim B(60, 0.15) \approx Po(9)$ for using Po(9)	B1
	$P(Y \leq 12)$, = 0.8758 awrt 0.876	M1, A1 (3)
		(5 marks)

2. (a)	Let X be the random variable the number of games Bhim loses. X ~B(9, 0.2)	B1
	$P(X \leq 3) - P(X \leq 2) = 0.9144 - 0.7382$ or $(0.2)^3(0.8)^6\frac{9!}{3!6!}$	M1
	= 0.1762 = 0.1762 awrt 0.176	A1 (3)
(b)	$P(X \leq 4) = 0.9804$ awrt 0.98	M1A1 (2)
(c)	Mean = 3 variance = 2.85, $\frac{57}{20}$	B1 B1 (2)
(d)	Po(3) poisson	M1
	$P(X > 4) = 1 - P(X < 4)$	M1
	$= 1 - 0.8153$ $= 0.1847$	A1 (3)
		[10]

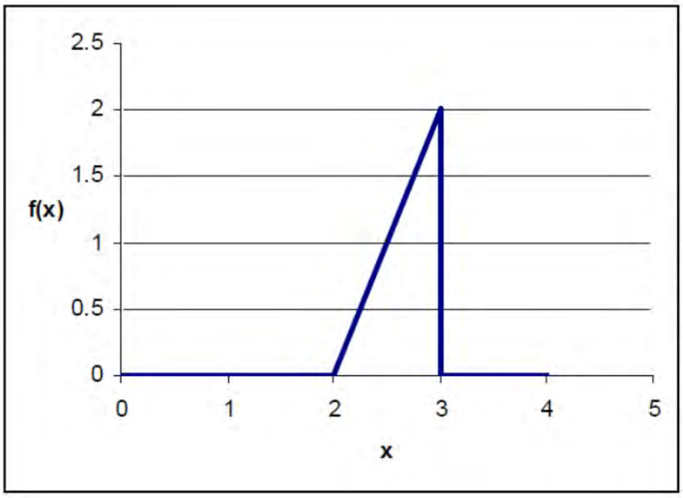
Question Number	Scheme	Marks
3.		
(a)	$\frac{1}{2}(a+b) = 23$ and $\frac{1}{12}(b-a)^2 = 75$	B1B1
	$a+b = 46$ and $b-a = \sqrt{12 \times 75}\ (=30)$	M1
	Adding gives $2b = 76$	M1
	$\underline{\boldsymbol{b = 38}}$ and $\underline{\boldsymbol{a = 8}}$	A1 A1
		(6)
(b)	$P(23 < X < c) = 0.5 - 0.32$ or $c = 28.4$ and prob $= \frac{5.4}{30}$	M1
	$= \underline{\mathbf{0.18}}$	A1
		(2) **[8]**

Question Number	Scheme	Marks
4.		
(a)	$K(2^4 + 2^2 - 2) = 1$	M1
	$K = \frac{1}{18}$	A1
		(2)
(b)	$1 - F(1.5) = 1 - \frac{1}{18}(1.5^4 + 1.5^2 - 2)$	M1
	$= 0.705$ or $\frac{203}{288}$	A1
		(2)
(c)		M1 A1
	$f(y) = \begin{cases} \frac{1}{9}(2y^3 + y) & 1 \le y \le 2 \\ 0 & \textit{otherwise} \end{cases}$	B1
		(3)

Question Number	Scheme	Marks
5. (a)	$X \sim B(120, 0.075)$	B1
	Approximated by Po(9)	M1A1
	$P(X > 3) = 1 - P(X \le 3)$	M1
	$= 1 - 0.0212$	
	$= 0.9788$ awrt 0.979	A1 (5)
	P(At least 4 defective components in each box)	
(b)	$= P(X > 3) \times P(X > 3)$	M1
	$= 0.9788^2$	
	$= 0.95804944$ awrt 0.958	A1 (2)
		(7 marks)

Question Number	Scheme	Marks
6. (a)	$P(A > 3) = \frac{2}{5} = 0.4$	B1 (1)
(b)	$(0.4)^3$,$= 0.064$ or $\dfrac{8}{125}$	M1, A1 (2)
(c)	$f(y) = \dfrac{d}{dy}(F(y)) = \begin{cases} \dfrac{3y^2}{125} & 0 \le y \le 5 \\ 0 & otherwise \end{cases}$	M1 A1 (2)
(d)	 Shape of curve and start at (0,0)	B1
	Point (5, 0) labelled and curve between 0 and 5 and pdf ≥ 0	B1 (2)
(e)	Mode = 5	
(f)	$E(Y) = \int_0^5 \left(\dfrac{3y^3}{125}\right) dy = \left[\dfrac{3y^4}{500}\right]_0^5 = \dfrac{15}{4}$ or 3.75	B1 (1)
		M1 M1 A1 (3)
(g)	$P(Y > 3) = \begin{cases} \int_3^5 \frac{3y^2}{125} dy \\ \text{or } 1 - F(3) \end{cases} = 1 - \dfrac{27}{125} = \dfrac{98}{125} = 0.784$	
		M1 A1 (2)
		(13 marks)

Question Number	Scheme		Marks
7(a) i	If $X \sim B(n,p)$ and		
	n is large or $n > 10$ **or** $np > 5$ **or** $nq > 5$		B1
	p is close to 0.5 **or** $nq > 5$ <u>and</u> $np > 5$		B1
	then X can be approximated by $N(np, np(1-p))$		
			(2)
ii	mean = np		B1
	variance = $np(1-p)$	must be in terms of p	B1
			(2)
(b)	$X \sim N(60, 58.2)$ or $X \sim N(60, 7.63^2)$	60, 58.2	B1, B1
	$P(X \geq 40) = P(X > 39.5)$	using 39.5 or 40.5	M1
	$= 1 - P\left(z < \pm\left(\dfrac{39.5-60}{\sqrt{58.2}}\right)\right)$	standardising 39.5 or 40 or 40.5 and their μ and σ	M1
	$= 1 - P(z < -2.68715\ldots)$		
	$= 0.9965$	allow answers in range 0.996 – 0.997	A1dep on both M
			(5)
(c)	$E(X) = 60$	may be implied or ft from part (b)	B1ft
	Expected profit = $(2000 - 60) \times 11 - 2000 \times 0.70$		M1
	$= £19\,940.$		A1
			(3)
			Total 12

Question Number	Scheme	Marks
8.		
(a)	 Max height of 2 labelled and goes through(2,0) shape must be between 2 and 3 and no other lines drawn (accept patios drawn) / correct shape	**B1** **B1 B1** **(3)**
(b)	3	**B1** **(1)**
(c)	$\int_2^3 2x(x-2)\mathrm{d}x = \left[\frac{2x^3}{3} - 2x^2\right]_2^3$ $= 2\frac{2}{3}$	**M1A1** **A1** **(3)**
(d)	$\int_2^m 2(x-2)\mathrm{d}x = 0.5$ $\left[x^2 - 4x\right]_2^m = 0.5$ $m^2 - 4m + 4 = 0.5$ $m^2 - 4m + 3.5 = 0$ $m = \frac{4 \pm \sqrt{2}}{2}$ $m = 2.71$	**M1** **A1** **M1** **(4)**
(e)	Negative skew. mean < median < mode .	**B1** **B1dep** **(2)**

Examiner reports

Question 1

The majority of candidates achieved full marks on this question with the most common errors caused by difficulties in identifying, interpreting and/or working with the inequalities. In part (b) whilst a few candidates wrote down $P(X \leq 13)$ they were unable to find this probability correctly. The most common error was to use $1 - P(X \leq 12)$

Question 2

This question was well answered by the majority of candidates with many scoring full marks. There were, of course, candidates who failed to score full marks. This was usually the result of inaccurate details, rather than lack of knowledge. In particular, manipulation of inequalities requires concentration and attention to detail. In part (a) the most common error seen was using $P(X = 3) = P(X \leq 4) - P(X \leq 3)$.

Parts (b) and (c) were usually correct. The most common error was to find $P(X \leq 3)$ rather than $P(X \leq 4)$ in part (b). A minority of candidates used the Normal as their approximation in part (d). The simple rule "*n* is large, *p* is small: use Poisson" clearly applies in this case.

Question 3

Part (a) frequently involved complicated quadratics and substitution without realising that the quadratic could easily be eliminated immediately from the variance equation as $b-a>0$. Even those that managed to set up two correct linear equations continued with a substitution rather than simply adding the two equations.

Part (b) was done well and most candidates took the route of calculating a value for c and then calculating the required probability. Some candidates made arithmetic errors and some correctly calculated $c = 28.4$ but then failed to do anything with it. Others forgot that 23 was the mean and so $P(X < 23) = 0.5$ had to be worked out.

Although the majority of candidates succeeded in finding the correct answers here, few achieved neat and mathematically fluent solutions.

Question 4

The majority of candidates were able to attempt this question with a high degree of success

(a) Many candidates had a number of attempts at this part before getting a solution. In some cases, responses showed a lack of understanding between the p.d.f. and the c.d.f. This occurred when the candidate differentiated the given function then proceeded to integrate it. The most common error was to interpret the given function as the p.d.f., integrate it and put the answer equal to 1. A small number of candidates took the value of $k = 1/18$ and used it to work backwards.

(b) The most common errors were to find F(1.5) or integrate the given function.

(c) There were many correct solutions with a minority of candidates being unsuccessful. Marks were mainly lost through, having differentiated correctly to find the function, not specifying the p.d.f. fully. A few candidates tried integrating to find the p.d.f.

Question 5

In part (a) a significant minority of candidates used a Normal approximation to the Binomial, despite the fact that the Binomial parameters, 120 and 0.075, can only be described as large and small respectively. Those candidates who made their method clear were able to score some marks but those who simply wrote their answers down gained few marks if any. However, a large number of scripts were awarded full marks.

There were some interesting features to the overall response to part (b). Some candidates used an elaborate method that consisted of defining a new Binomial distribution, B(2, 0.9788), and then using the formula for Binomial probability to complete their solution. This was almost always successful, but contrasts (in the amount of work involved) with those candidates who used the multiplication rule for independent events and therefore simply wrote 0.9788×0.9788 or 0.97882^2.

The most curious aspect of part (b) was the frequency with which candidates retraced their steps by repeating their work for part (a), before proceeding with part (b). Intriguingly, it was not uncommon for candidates to have been incorrect in part (a), but to go on and write a different solution to part (a), correct this time, as part of (b). Unfortunately, correct answers only earn marks when they occur in the parts to which they relate. If only these candidates had reflected on the discrepancy between their two attempts they could have corrected part (a) and earned full marks.

Question 6

Many candidates attained full or nearly full marks for this question.

In part (a) many candidates were unable to correctly state $P(A > 3) = \frac{2}{5}$.

In part (b) some candidates multiplied their answer to (a) by 3 rather than finding (a) cubed.

Part (c) was well done and apart from those who made errors in the differentiation candidates gained full marks for this part.

The most common error in question 6 was drawing the sketch graph incorrectly in part (d). A straight line was often seen, either sloping from 0 to 5 or parallel to the x-axis.

In part (e) a few candidates attempted to calculate the mode rather than reading it straight from the graph. Not only did this waste time the result was usually incorrect.

In part (f) candidates confused $f(x)$ with $F(x)$ and/or used the formula for $E(Y)$ incorrectly. Errors in the simple integration were often seen.

In part (g) few candidates chose to use $P(X > 3) = 1 - F(3)$ and instead used the method involving integration where too often the incorrect limits were used.

Question 7

The quality of answers to this question was better than to similar questions in previous years. Most used the discriminant to answer part (a) and, apart from occasional slips with signs, were able to establish the inequality correctly. A few realised that the discriminant had to be used but tried to apply it to $k\,2 - 4k - 12$. In part (b) the majority were able to find the critical values of – 2 and 6 but many then failed to find the correct inequalities with $x > -2$ and $x > 6$ being a common incorrect answer. Some candidates still thought that the correct regions could be written as $6 < k < -2$ but there were many fully correct solutions seen often accompanied by correct sketches.

Question 8

The majority of candidates attempted this question.

(a) Most sketches were clearly labelled with a few omitting the value on the y-axis. Candidates should draw their sketch in the space in the question book, not on graph paper.

(b) A few gave the mode as 2 or 1.

(c) Most were able to find E(X) with only occasional errors in using xf(x) = $2x – 4$ or in substituting the limits.

(d) Finding the median proved challenging for a sizeable minority. Although most wrote that F(m) = 0.5, finding F(m) proved difficult. There were many exemplary solutions but those candidates who struggled got x^2-4x, but then failed to use the limits correctly or made arithmetic mistakes. It was common for those who had no real understanding to put $2x – 4 = 0.5$ and solve to get $x = 2.25$.

(e) In many cases the answers to this part reflected confusion in understanding the concept of skewness. In many cases where responses were incorrect there was little or no evidence of using the results found, or positive skewness was stated but the reason related to negative skewness.

Statistics for S2 Practice Paper Bronze 3

				Mean average scored by candidates achieving grade:							
Qu	**Max Score**	**Modal score**	**Mean %**	**ALL**	**A***	**A**	**B**	**C**	**D**	**E**	**U**
1	5		86.8	4.34		4.76	4.55	4.28	3.89	3.41	2.25
2	10		87.2	8.72	9.67	9.45	9.02	8.45	7.71	6.74	4.81
3	8		93.1	7.45	7.92	7.83	7.52	6.91	6.64	6.21	4.99
4	7		82.1	5.75		6.58	5.84	4.94	3.82	2.87	1.35
5	7		83.7	5.86	6.54	6.30	5.68	5.00	4.24	3.62	1.72
6	13		77.0	10.01		12.06	10.63	9.31	7.50	5.44	2.93
7	12		76.8	9.22		10.57	9.73	8.91	8.06	6.81	4.26
8	13		72.3	9.40		11.40	9.78	7.26	5.46	5.14	2.32
	75		**81.0**	**60.75**		**68.95**	**62.75**	**55.06**	**47.32**	**40.24**	**24.63**