General Certificate of Education
January 2008
Advanced Subsidiary Examination

MATHEMATICS **MS/SS1B**
Unit Statistics 1B

STATISTICS
Unit Statistics 1B

Tuesday 22 January 2008 1.30 pm to 3.00 pm

For this paper you must have:
- an 8-page answer book
- the blue AQA booklet of formulae and statistical tables
- an insert for use in Question 4 (enclosed).

You may use a graphics calculator.

Time allowed: 1 hour 30 minutes

Instructions
- Use blue or black ink or ball-point pen. Pencil should only be used for drawing.
- Write the information required on the front of your answer book. The *Examining Body* for this paper is AQA. The *Paper Reference* is MS/SS1B.
- Answer **all** questions.
- Show all necessary working; otherwise marks for method may be lost.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.
- Fill in the boxes at the top of the insert.

Information
- The maximum mark for this paper is 75.
- The marks for questions are shown in brackets.
- Unit Statistics 1B has a **written paper only**.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.

Answer **all** questions.

1 In large-scale tree-felling operations, a machine cuts down trees, strips off the branches and then cuts the trunks into logs of length X metres for transporting to a sawmill.

It may be assumed that values of X are normally distributed with mean μ and standard deviation 0.16, where μ can be set to a specific value.

(a) Given that μ is set to 3.3, determine:

(i) $\mathrm{P}(X < 3.5)$; *(3 marks)*

(ii) $\mathrm{P}(X > 3.0)$; *(3 marks)*

(iii) $\mathrm{P}(3.0 < X < 3.5)$. *(2 marks)*

(b) The sawmill now requires a batch of logs such that there is a probability of 0.025 that any given log will have a length less than 3.1 metres.

Determine, to two decimal places, the new value of μ. *(4 marks)*

2 The head and body length, x millimetres, and tail length, y millimetres, of each of a sample of 20 adult dormice were measured. The following statistics are derived from the results.

$$S_{xx} = 1280.55 \qquad S_{yy} = 281.8 \qquad S_{xy} = 416.3$$

(a) Calculate the value of the product moment correlation coefficient between x and y. *(2 marks)*

(b) Interpret your value in the context of this question. *(2 marks)*

(c) Write down the value of the product moment correlation coefficient if the measurements had been recorded in centimetres. *(1 mark)*

(d) Give a reason why it is not generally advisable to calculate the value of the product moment correlation coefficient without first viewing a scatter diagram of the data. Illustrate your answer with a sketch. *(2 marks)*

3 The height, in metres, of adult male African elephants may be assumed to be normally distributed with mean μ and standard deviation 0.20.

The heights of a sample of 12 such elephants were measured with the following results, in metres.

3.37 3.45 2.93 3.42 3.49 3.67 2.96 3.57 3.36 2.89 3.22 2.91

(a) Stating a necessary assumption, construct a 98% confidence interval for μ. *(6 marks)*

(b) The mean height of adult male **Asian** elephants is known to be 2.90 metres.

Using your confidence interval, state, with a reason, what can be concluded about the mean heights of adult males in these two types of elephant. *(2 marks)*

4 [**Figure 1**, printed on the insert, is provided for use in this question.]

Roseen is a self-employed decorator who wishes to estimate the times that it will take her to decorate bedrooms based upon their floor areas. She records the floor area, $x\,\text{m}^2$, and the decorating time, y hours, for each of 10 bedrooms she has recently decorated.

x	11.0	22.0	7.5	21.0	13.0	16.5	14.0	16.0	18.5	20.5
y	15.0	35.0	16.0	23.5	24.0	17.5	14.5	27.5	22.5	34.5

(a) On **Figure 1**, plot a scatter diagram of these data. *(2 marks)*

(b) Calculate the equation of the least squares regression line of y on x. *(4 marks)*

(c) Draw your regression line on **Figure 1**. *(2 marks)*

(d) (i) Use your regression equation to estimate the time that Roseen will take to decorate a bedroom with a floor area of $15\,\text{m}^2$. *(2 marks)*

(ii) Making reference to **Figure 1**, comment on the likely reliability of your estimate in part (d)(i). *(2 marks)*

5 A health club has a number of facilities which include a gym and a sauna. Andrew and his wife, Heidi, visit the health club together on Tuesday evenings.

On any visit, Andrew uses either the gym or the sauna or both, but no other facilities. The probability that he uses the gym, $P(G)$, is 0.70. The probability that he uses the sauna, $P(S)$, is 0.55. The probability that he uses both the gym and the sauna is 0.25.

(a) Calculate the probability that, on a particular visit:

(i) he does not use the gym; *(1 mark)*

(ii) he uses the gym but not the sauna; *(2 marks)*

(iii) he uses either the gym or the sauna but not both. *(2 marks)*

(b) Assuming that Andrew's decision on what facility to use is independent from visit to visit, calculate the probability that, during a month in which there are exactly four Tuesdays, he does not use the gym. *(2 marks)*

(c) The probability that Heidi uses the gym when Andrew uses the gym is 0.6, but is only 0.1 when he does not use the gym.

Calculate the probability that, on a particular visit, Heidi uses the gym. *(3 marks)*

(d) On any visit, Heidi uses **exactly one** of the club's facilities.

The probability that she uses the sauna is 0.35.

Calculate the probability that, on a particular visit, she uses a facility other than the gym or the sauna. *(2 marks)*

6 For each of the Premiership football seasons 2004/05 and 2005/06, a count is made of the number of goals scored in each of the 380 matches. The results are shown in the table.

Number of goals scored in a match	**Number of matches**	
	2004/05	**2005/06**
0	30	32
1	79	82
2	99	95
3	68	78
4	60	48
5	24	30
6	11	9
7	6	6
8	2	0
9	1	0
Total	**380**	**380**

(a) For the number of goals scored in a match during the **2004/05** season:

(i) determine the median and the interquartile range; *(4 marks)*

(ii) calculate the mean and the standard deviation. *(4 marks)*

(b) Two statistics students, Jole and Katie, independently analyse the data on the number of goals scored in a match during the **2005/06** season.

- Jole determines correctly that the median is 2 and that the interquartile range is also 2.
- Katie calculates correctly, to two decimal places, that the mean is 2.48 and that the standard deviation is 1.59 .

(i) Use your answers from part (a), together with Jole's and Katie's results, to compare briefly the two seasons with regard to the average and the spread of the number of goals scored in a match. *(2 marks)*

(ii) Jole claims that Katie's results must be wrong as 95% of values always lie within 2 standard deviations of the mean and $(2.48 - 2 \times 1.59) < 0$ which is nonsense.

Explain why Jole's claim is incorrect. (You are not expected to confirm Katie's results.) *(2 marks)*

7 A travel agency in Tunisia offers customers a 3-day tour into the Sahara desert by either coach or minibus.

(a) The agency accepts bookings from 50 customers for seats on the coach. The probability that a customer, who has booked a seat on the coach, will **not** turn up to claim the seat is 0.08, and may be assumed to be independent of the behaviour of other customers.

Determine the probability that, of the customers who have booked a seat on the coach:

(i) two or more will **not** turn up;

(ii) three or more will **not** turn up. *(4 marks)*

(b) The agency accepts bookings from 15 customers for seats on the minibus. The probability that a customer, who has booked a seat on the minibus, will **not** turn up to claim the seat is 0.025, and may be assumed to be independent of the behaviour of other customers.

Calculate the probability that, of the customers who have booked a seat on the minibus:

(i) all will turn up;

(ii) one or more will **not** turn up. *(4 marks)*

(c) The coach has 48 seats and the minibus has 14 seats. If 14 or fewer customers who have booked seats on the minibus turn up, they will be allocated a seat on the minibus. If all 15 customers who have booked seats on the minibus turn up, one will be allocated a seat on the coach. This will leave only 47 seats available for the 50 customers who have booked seats on the coach.

Use your results from parts (a) and (b) to calculate the probability that there will be seats available on the coach for all those who turn up having booked such seats. *(4 marks)*

END OF QUESTIONS

Surname		Other Names	
Centre Number			Candidate Number				
Candidate Signature									

General Certificate of Education
January 2008
Advanced Subsidiary Examination

ASSESSMENT and QUALIFICATIONS ALLIANCE

MATHEMATICS **MS/SS1B**
Unit Statistics 1B

STATISTICS
Unit Statistics 1B

Insert

Insert for use in **Question 4**.

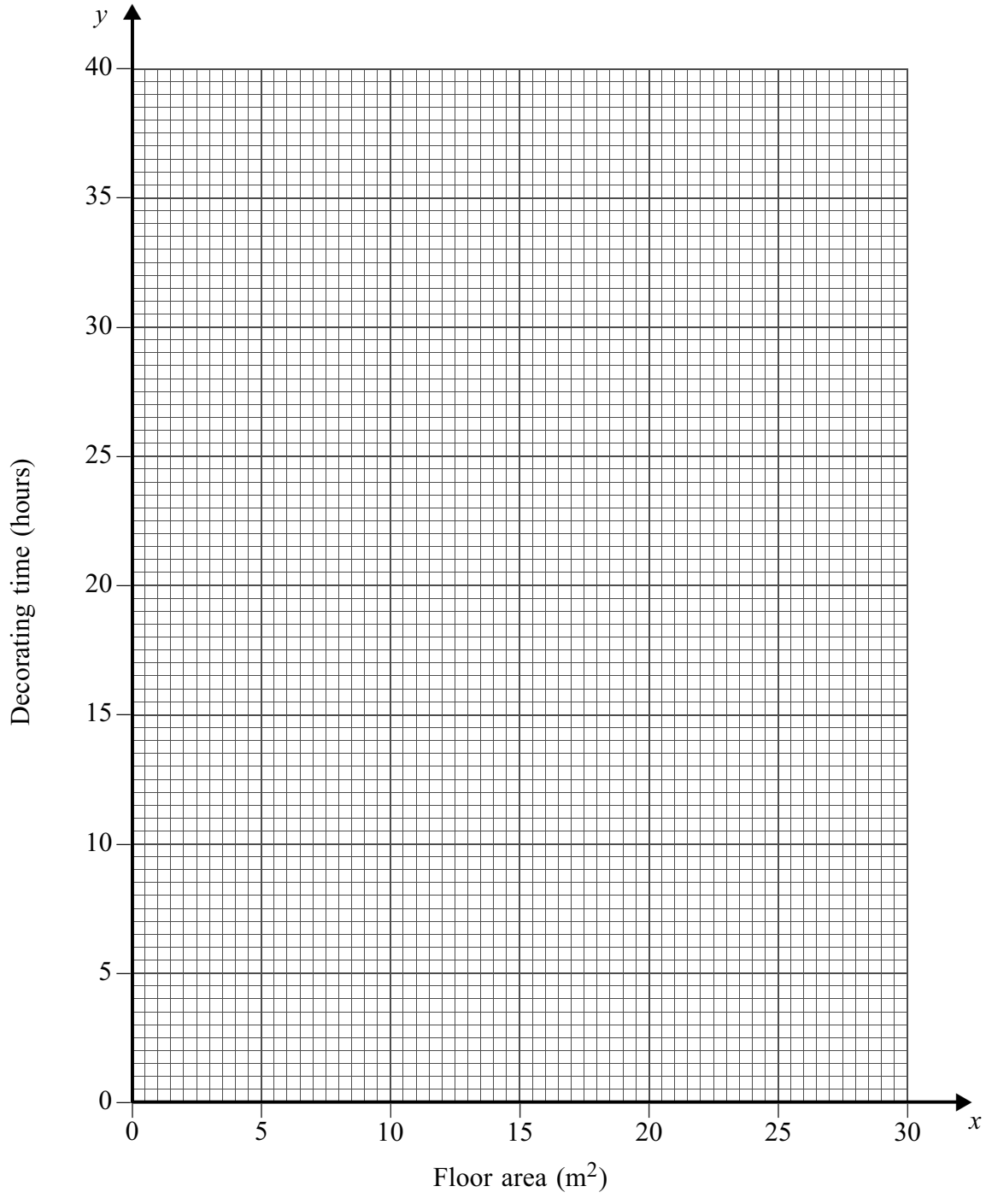
Fill in the boxes at the top of this page.

Fasten this insert securely to your answer book.

Turn over for Figure 1

Figure 1 (for use in Question 4)

Floor Areas and Decorating Times



Version 1.0: 01.08

ASSESSMENT and QUALIFICATIONS ALLIANCE

General Certificate of Education

Mathematics 6360
Statistics 6380

MS/SS1B Statistics 1B

Mark Scheme

2008 examination - January series

Mark schemes are prepared by the Principal Examiner and considered, together with the relevant questions, by a panel of subject teachers. This mark scheme includes any amendments made at the standardisation meeting attended by all examiners and is the scheme which was used by them in this examination. The standardisation meeting ensures that the mark scheme covers the candidates' responses to questions and that every examiner understands and applies it in the same correct way. As preparation for the standardisation meeting each examiner analyses a number of candidates' scripts: alternative answers not already covered by the mark scheme are discussed at the meeting and legislated for. If, after this meeting, examiners encounter unusual answers which have not been discussed at the meeting they are required to refer these to the Principal Examiner.

It must be stressed that a mark scheme is a working document, in many cases further developed and expanded on the basis of candidates' reactions to a particular paper. Assumptions about future mark schemes on the basis of one year's document should be avoided; whilst the guiding principles of assessment remain constant, details will change, depending on the content of a particular examination paper.

Set and published by the Assessment and Qualifications Alliance.

The Assessment and Qualifications Alliance (AQA) is a company limited by guarantee registered in England and Wales (company number 3644723) and a registered charity (registered charity number 1073334).
Registered address: AQA, Devas Street, Manchester M15 6EX *Dr Michael Cresswell Director General*

Key to mark scheme and abbreviations used in marking

M	mark is for method		
m or dM	mark is dependent on one or more M marks and is for method		
A	mark is dependent on M or m marks and is for accuracy		
B	mark is independent of M or m marks and is for method and accuracy		
E	mark is for explanation		
√or ft or F	follow through from previous incorrect result	MC	mis-copy
CAO	correct answer only	MR	mis-read
CSO	correct solution only	RA	required accuracy
AWFW	anything which falls within	FW	further work
AWRT	anything which rounds to	ISW	ignore subsequent work
ACF	any correct form	FIW	from incorrect work
AG	answer given	BOD	given benefit of doubt
SC	special case	WR	work replaced by candidate
OE	or equivalent	FB	formulae book
A2,1	2 or 1 (or 0) accuracy marks	NOS	not on scheme
–*x* EE	deduct *x* marks for each error	G	graph
NMS	no method shown	c	candidate
PI	possibly implied	sf	significant figure(s)
SCA	substantially correct approach	dp	decimal place(s)

No Method Shown

Where the question specifically requires a particular method to be used, we must usually see evidence of use of this method for any marks to be awarded. However, there are situations in some units where part marks would be appropriate, particularly when similar techniques are involved. Your Principal Examiner will alert you to these and details will be provided on the mark scheme.

Where the answer can be reasonably obtained without showing working and it is very unlikely that the correct answer can be obtained by using an incorrect method, we must award **full marks**. However, the obvious penalty to candidates showing no working is that incorrect answers, however close, earn **no marks**.

Where a question asks the candidate to state or write down a result, no method need be shown for full marks.

Where the permitted calculator has functions which reasonably allow the solution of the question directly, the correct answer without working earns **full marks**, unless it is given to less than the degree of accuracy accepted in the mark scheme, when it gains **no marks**.

Otherwise we require evidence of a correct method for any marks to be awarded.

MS/SS1B

Q	Solution	Marks	Total	Comments
1(a)(i)	$P(X < 3.5) = P\left(Z < \frac{3.5-3.3}{0.16}\right) =$	M1		Standardising (3.45, 3.5 or 3.55) with 3.3 & ($\sqrt{0.16}$, 0.16 or 0.16^2) and/or $(3.3 - x)$
	$P(Z < 1.25) =$	A1		CAO; ignore sign
	0.894 to 0.895	A1	3	AWFW (0.89435)
(ii)	$P(X > 3.0) = P\left(Z > \frac{3.0-3.3}{0.16}\right) =$	M1		Standardising (2.95, 3 or 3.05) with 3.3 & ($\sqrt{0.16}$, 0.16 or 0.16^2) and/or $(3.3 - x)$
	$P(Z > -1.875) = P(Z < 1.875) =$	m1		Correct area change
	0.969 to 0.97(0)	A1	3	AWFW (0.96960)
(iii)	$P(3.0 < X < 3.5) =$ (i) – [1 – (ii)] =	M1		OE
	0.863 to 0.865	A1	2	AWFW: CSO (0.86395)
(b)	$0.025 \Rightarrow z = 1.96$	B1		CAO; ignore sign
	$z = \frac{3.1-\mu}{0.16}$	M1		Standardising 3.1 with μ and 0.16; allow $(\mu - 3.1)$
	$= -1.96$	m1		Equating z-term to z-value; not using 0.025, 0.975, $\lvert 1 - z \rvert$ or $\Phi(0.025) = 0.507$ to 0.512
	Hence $\mu = 3.4(0)$ to 3.42	A1	4	AWFW; CSO (3.4136)
	Total		**12**	

MS/SS1B (cont)

Q	Solution	Marks	Total	Comments
2(a)	$r = \frac{416.3}{\sqrt{1280.55 \times 281.8}}$ =	M1		Allow no $\sqrt{}$
	0.69 to 0.7(0)	A1	2	AWFW (0.693) (0.00115)
(b)	(Quite or fairly) **weak / some / moderate** (quite or fairly) **strong positive correlation** (relationship / association)	A1		OE; must qualify strength and indicate positive A0 for poor / reasonable / average / medium / good A0 for very weak / very strong etc
	between **head & body length** and **tail length** *Ignore subsequent alternative comments only if A1 B1 already scored*	B1	2	Context; accept 'body and tail' or even 'head and tail'
	OR **Some** evidence that mice with large head & body lengths also have long tails	(A1) (B1)		OE; must qualify strength and indicate positive in context
(c)	0.69 to 0.7(0) **OR** Answer to (a)	B1$\surd$	1	$0 < r < 1$
(d)	Existence of: **Non-linear** relationship **Outliers** **More than one** relationship	B1		Any one; OE Not reasons identifiable from context (eg spurious)
	Sensible related sketch SC: Check on calculation $\Rightarrow$ B1 B0	B1	2	
	Total		**7**	

MS/SS1B (cont)

Q	Solution	Marks	Total	Comments
3(a)	12 elephants are a **random sample** **OR** are **selected independently**	B1		OE; eg representative
	Mean $\bar{x} = \frac{39.24}{12} = 3.27$	B1		CAO
	98% $\Rightarrow$ z = 2.32 to 2.33	B1		AWFW (2.3263)
	CI for μ is $\bar{x} \pm z \times \frac{\sigma}{\sqrt{n}}$	M1		Used; must have $\sqrt{n}$ with $n > 1$
	Thus $3.27 \pm 2.3263 \times \frac{0.20}{\sqrt{12}}$	A1√		√ on $\bar{x}$ and z only
	Hence 3.27 ± 0.1343			
	Hence $3.27 \pm$ (0.13 to 0.14) **OR** (3.13 to 3.14, 3.4(0) to 3.41)	A1	6	AWFW
(b)	Value of 2.90 is **below / outside** CI	B1√		√ on (a); OE
	Suggests **mean height** of adult male **Asian elephants** is **less than / different from that / mean height** of adult male **African elephants**	B1√ dep	2	√ on (a); OE
	Total		**8**	

MS/SS1B (cont)

Q	Solution	Marks	Total	Comments
4(a)	≥ 8 points plotted accurately (≥ 6 points plotted accurately)	B2 (B1)	2	
(b)	*b* (gradient) = 1.19 to 1.2(0) (*b* (gradient) = 1.1 to 1.3)	B2 (B1)		AWFW (1.19066)
	a (intercept) = 3.8 to 4(.0) (*a* (intercept) = 2.2 to 5.4)	B2 (B1)	4	AWFW (3.94949)
	Attempt at $\sum x$, $\sum x^2$, $\sum y$ and $\sum xy$ **OR** Attempt at S_{xx} and S_{xy}	(M1)		160, 2758, 230 and 3915.75 198 and 235.75
	Attempt at correct formula for *b* (gradient) *b* (gradient) = 1.19 to 1.2(0) *a* (intercept) = 3.8 to 4(.0)	(m1) (A1) (A1)		 AWFW AWFW
	Accept *a* and *b* interchanged only if then identified correctly later in question			
(c)	Line plotted accurately (Evidence of correct method for ≥ 2 points)	B2 (M1)	2	At least from $x \approx 7.5$ to 22.0 $x = 10 \Rightarrow y = 15.5$ to 16.5 $x = 20 \Rightarrow y = 27.0$ to 28.5
(d)(i)	When $x = 15$:			
	$y = 21.5$ to 22(.0) ($y = 18.5$ to 25(.0))	B2 (B1)	2	AWFW (21.8) AWFW
	If B0, then use of c's equation with $x = 15$	(M1)		
(ii)	**Points** are quite **widely scattered** about line	B1		When $x = 14$ then $y = 14.5$ When $x = 16$ then $y = 27.5$
	Hence **not** very **reliable**	B1 dep	2	B0 B0 for 'interpolation so reliable'
	Total		**12**	

MS/SS1B (cont)

Q	Solution	Marks	Total	Comments
5(a)(i)	$P(G') = 1 - 0.70 = 0.3(0)$	B1	1	CAO; OE
(ii)	$P(G \cap S') =$			
	$0.70 - (0.25 \text{ or } 0.55 \text{ or } 0.45)$	M1		Can be implied only if answer is correct
	or $1 - 0.55$			
	$= 0.45$	A1	2	CAO; OE
(iii)	P(1 only) =			
	$0.70 + 0.55 - (2 \times 0.25)$			Can be implied only if answer is correct;
	or $1 - 0.25$	M1		allow no $(\times 2)$ but not by implication
	or $0.45 + 0.30$			from answer
	$= 0.75$	A1	2	CAO; OE
(b)	$P(G' \cap G' \cap G' \cap G') = [(a)(i)]^4$	M1		Can be implied by correct answer Ignore multiplier of 4
	$= 0.0081$	A1	2	CAO; OE
(c)	$P(H_G) = P(A_G \cap H_G) + P(A_{G'} \cap H_G) =$			
	(0.70×0.60) or 0.42	M1		Can be implied by correct answer Ignore additional terms
	(0.30×0.10) or 0.03	M1		Can be implied by correct answer
	$= 0.42 + 0.03 = 0.45$	A1	3	CAO; OE
(d)	$P(H_O) = 1 - [0.35 + (c)]$	M1		Can be implied by correct answer
	$= 0.2(0)$	A1	2	CAO; OE
	Total		**12**	

MS/SS1B (cont)

Q	Solution	Marks	Total	Comments
6 (a)(i)	x: 0 1 2 3 4 5 6 7 8 9 F: 30 109 208 276 336 360 371 377 379 380			
	Median ($\approx 190.5^{th}$) = 2	B2		CAO; B0 if shown method incorrect
	Interquartile range ($\approx 285.75^{th} - \approx 95.25^{th}$)			
	= 4 – 1 = 3	B2	4	CAO; B0 if shown method incorrect B1 for identification of 4 and 1
	If neither is correct but F attempted	(M1)		Allow for median = $1+\frac{x}{99}$
	and matched correctly with ≥ 5 x-values	(A1)		
(ii)	Mean $(\bar{x})$ = 2.56 to 2.57	B2		AWFW (2.56316)
	(2.5 to 2.6)	(B1)		AWFW $\sum fx = 974$ and $\sum fx^2 = 3546$
	Standard Deviation (s_n, s_{n-1}) =			
	1.66 to 1.67	B2	4	AWFW (1.66187)
	(1.6 to 1.7)	(B1)		AWFW (1.66406)
	If neither is correct but $\sum fx$ attempted	(M1)		
	and result divided by 380	(M1)		
(b)(i)	Average: Same/similar/greater in 2004/05	B1 dep		OE; dep on 2 and 2.5 to 2.6
	Spread: Similar/greater in 2004/05	B1 dep	2	OE; dep on 3 and 1.6 to 1.7
(ii)	Rule applies to data that is (approximately) symmetric/normal/bell-shaped	B1		OE
	Data for 2005/06 is (positively) skewed/ not symmetric/not normal/not bell-shaped	B1	2	OE
	Total		**12**	

MS/SS1B (cont)

Q	Solution	Marks	Total	Comments
7(a)	Use of binomial in (a) or (b)	M1		Can be implied by answers
(i)	$P(X \geq x) = 1 - P(X \leq x - 1)$			Identified from an answer / 1 – answer
	OR	M1		Can be implied from a correct answer
	$= 1 - B(\Sigma x, 50, 0.08)$			Identified from an answer/expression
	$= 1 - 0.0827 = 0.915$ to $0.92(0)$	A1		AWFW (0.9173)
				≥ 1 correct ⇒ M1 M1
(ii)	$P(X \geq 3)$			
	$= 1 - 0.2260 = 0.77(0)$ to 0.775	A1	4	AWFW (0.7740)
(b)(i)	$P(Y = 0) = (1 - 0.025)^{15} = 0.975^{15}$	M1		Can be implied from correct answer
	$= 0.68(0)$ to 0.685	A1		AWFW (0.6840)
(ii)	$P(Y \geq 1) = 1 -$ (i)	M1		Can be implied from answer if ε (0, 1)
	$= 0.315$ to $0.32(0)$	A1√	4	√ on (i) if ε (0, 1) (0.3160)
(c)	Probability =			
	[(b)(ii) × (a)(i)] or (0.316×0.917)	M1		Ignore additional terms
	[(b)(i) × (a)(ii)] or (0.684×0.774)	M1		
	$= 0.2898 + 0.529$	A1		2 terms added with ≥ 1 correct
	$= 0.81$ to 0.83	A1	4	AWFW (0.8193)
	Total		**12**	
	TOTAL		**75**	

Scaled mark component grade boundaries - January 2008 exams

GCE

Component Code	Component Title	Maximum Scaled Mark	Scaled Mark Grade Boundaries A	B	C	D	E
MM2A/C	GCE MATHEMATICS UNIT M2A - COURSEWORK	25	20	18	15	13	10
MMA2/W	GCE MATHEMATICS UNIT M2A - WRITTEN	75	63	55	48	40	33
MM2B	GCE MATHEMATICS UNIT M2B	75	65	57	49	41	34
MPC1	GCE MATHEMATICS UNIT PC1	75	59	51	43	36	29
MPC2	GCE MATHEMATICS UNIT PC2	75	60	53	46	39	33
MPC3	GCE MATHEMATICS UNIT PC3	75	60	53	46	39	32
MPC4	GCE MATHEMATICS UNIT PC4	75	60	52	44	36	29
MS/SS1A/C	GCE MATHEMATICS UNIT S1A - COURSEWORK	25	20	18	15	13	10
MS/SS1A/W	GCE MATHEMATICS UNIT S1A - WRITTEN	75	59	51	44	36	30
MS2A/C	GCE MATHEMATICS UNIT S2A - COURSEWORK	25	*(no candidates were entered for this component)*				
MS2A/W	GCE MATHEMATICS UNIT S2A - WRITTEN	75	*(no candidates were entered for this component)*				
MS1B	GCE MATHEMATICS UNIT S1B	75	54	47	40	33	26
MS2B	GCE MATHEMATICS UNIT S2B	75	57	49	41	34	27
MED1	GCE MEDIA STUDIES UNIT 1	60	42	36	31	26	21
MED2	GCE MEDIA STUDIES UNIT 2	60	41	36	31	26	21
MED3	GCE MEDIA STUDIES UNIT 3	100	72	64	56	48	40
MED4	GCE MEDIA STUDIES UNIT 4	60	42	38	34	30	26
MED5	GCE MEDIA STUDIES UNIT 5	60	50	41	32	23	15
MED6	GCE MEDIA STUDIES UNIT 6	60	44	39	34	29	24
MUS1	GCE MUSIC UNIT 1	100	67	58	49	40	31
PA01	GCE PHYSICS A UNIT 1	50	36	32	28	24	20
PA02	GCE PHYSICS A UNIT 2	50	34	30	26	22	19