



Teacher Support Materials 2009

Maths GCE

Paper Reference MS/SS/1B

Copyright © 2009 AQA and its licensors. All rights reserved.

Permission to reproduce all copyrighted material has been applied for. In some cases, efforts to contact copyright holders have been unsuccessful and AQA will be happy to rectify any omissions if notified.

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.

Question 1

- 1 A large bookcase contains two types of book: hardback and paperback. The number of books of each type in each of four subject categories is shown in the table.

		Subject category				Total
		Crime	Romance	Science fiction	Thriller	
Type	Hardback	8	16	18	18	60
	Paperback	16	40	14	30	100
	Total	24	56	32	48	160

- (a) A book is selected at random from the bookcase. Calculate the probability that the book is:
- (i) a paperback; (1 mark)
 - (ii) not science fiction; (2 marks)
 - (iii) science fiction or a hardback; (2 marks)
 - (iv) a thriller, given that it is a paperback. (2 marks)
- (b) Three books are selected at random, without replacement, from the bookcase. Calculate, to three decimal places, the probability that one is crime, one is romance and one is science fiction. (4 marks)

Student Response

	a					
1	i $P(C \text{ Paperback}) = \frac{100}{160} = \frac{5}{8}$					Leave blank
						31
ii	$P(C \text{ Science fiction}) = \frac{32}{160} = \frac{4}{5}$					2
iii	$P(C \text{ Science fiction or Hardback}) = \frac{74}{160} = \frac{37}{80}$					2
iv	$P(C \text{ Thriller Paperback}) = \frac{30}{48} = \frac{5}{8}$					0
b	$\frac{32}{160} \times \frac{24}{159} \times \frac{56}{158} = 0.011$					M2
						①

Commentary

As was usually the case, the candidate has scored full marks in parts (a)(i), (ii) & (iii), although in the latter part an incorrect answer of $\frac{92}{160}$ was regularly seen. Part (a)(iv) was often more of a challenge and here the candidate has found P(Paperback | Thriller). The candidate's answer to part (b) was by far the most common. The three correct subject totals have been divided by 160, 159 & 158 respectively (those who used 160^3 lost a further mark) but no account has been taken of the $3! = 6$ permutations.

Mark scheme

Q	Solution	Marks	Total	Comments
1(a)				In (a), ratios (eg 100:160) are only penalised by 1 mark at first correct answer
(i)	$P(P) = \frac{100}{160} = \frac{50}{80} = \frac{25}{40} = \frac{10}{16} = \frac{5}{8} = 0.625$	B1	1	CAO
(ii)	$P(S') = 1 - \frac{32}{160}$ or $P(S) = \frac{32}{160} = \frac{128}{160} = \frac{64}{80} = \frac{32}{40} = \frac{16}{20} = \frac{8}{10} = \frac{4}{5} = 0.8$	M1		Or equivalent Ignore labels of S' & S Can be implied by correct answer
(iii)	$P(S \text{ or } H) = P(S \cup H) = \frac{60+32-18}{160} \text{ or } \frac{60+14}{160} \text{ or } \frac{32+8+16+18}{160} = \frac{74}{160} = \frac{37}{80} = 0.462 \text{ to } 0.463$	M1		Or equivalent Can be implied by correct answer
(iv)	$P(T P) = \frac{30}{160}$ (i) $= \frac{3}{100} = \frac{3}{10} = 0.3$	M1	2	CAO/AFWW (0.4625) Or equivalent Can be implied by correct answer But watch for $\frac{18}{160}$ or $\frac{48}{160}$
(b)	$P(1C \text{ & } 1R \text{ & } 1S) = \frac{24}{160} \times \frac{56}{159} \times \frac{32}{158} \times 6 \times 0.15 \times 0.35220 \times 0.20253 = 0.064 \text{ to } 0.0644$	M1 M1 M1 A1	2 2 2 4	Multiplication of any 3 different given subject totals Multiplication of 160, 159 & 158 Accept 3dp accuracy Award for $3 \leq \text{multiplier} \leq 6$ AWFW (0.0642) Do not accept a fraction as answer A correct answer can imply 4 marks
	Special Case: (Any given subject total) + 160 seen anywhere in (b)	(M1)	4	Can award if no marks scored in (b) Accept a decimal equivalent
		Total	11	

Question 2

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

Hermione, who is studying reptiles, measures the length, x cm, and the weight, y grams, of a sample of 11 adult snakes of the same type. Her results are shown in the table.

Snake	A	B	C	D	E	F	G	H	I	J	K
x	46	39	54	79	47	58	73	35	43	51	36
y	55	48	58	88	61	55	82	51	50	66	57

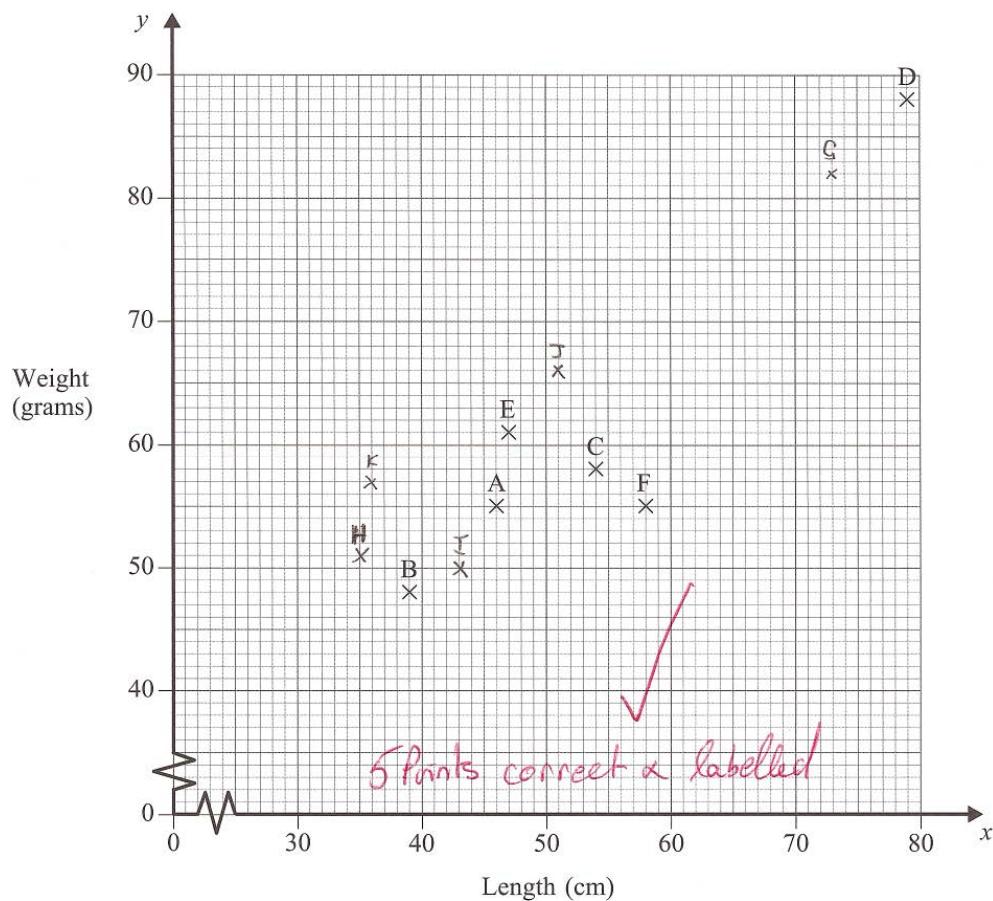
- (a) Calculate the value of the product moment correlation coefficient, r , between x and y . (3 marks)
- (b) Interpret your value in context. (2 marks)
- (c) On Figure 1, complete the scatter diagram for these data. (2 marks)
- (d) Subsequently it is found that, of the 11 adult snakes, 9 are male and 2 are female.
- (i) Given that female adult snakes are generally larger than male adult snakes, identify the 2 snakes which are most likely to be female. (1 mark)
- (ii) Hence, without further calculation, estimate the value of r for the 9 male snakes and revise, as necessary, your interpretation in part (b). (2 marks)

Student Response

number		
2a) $r = 0.89$	Not 3sf	Leave blank
b) There is a strong positive correlation between length of snakes and their weights		B2
c) [Figure 1]	✓	G 2
d) i) Female snakes - Snakes G and D	✓	B1
ii) $r \approx 0.2$	✗	B0
<u>revised part b)</u> For male snakes, there is a weak positive correlation between length and weight (dep)		
B0 (7)		

Figure 1 (for use in Question 2)

Lengths and Weights of Snakes



Commentary

Whilst many candidates scored full marks for this question, the above solution illustrates some common errors that lost marks. Given Instruction 5 on the Front Page of the Question Paper, the value of r was required to at least three decimal places. Part (b) required a reference to the strength and the sign of the correlation in context; all referenced here. The 5 points are accurately plotted and labelled (candidates were penalised for omitting the latter). The candidate has identified the two most likely female snakes but the estimated value of r for the remaining 9 male snakes is outside the acceptable range of 0.25 to 0.75, actual value is 0.488. Despite a correct revised interpretation, the final mark is not available as it is dependent upon the estimated value of r being within the acceptable range.

Mark Scheme

Q	Solution	Marks	Total	Comments
2(a)	$r = 0.893$ to 0.8933 $r = 0.89$ to 0.896 $r = 0.8$ to 0.95 or Attempt at $\sum x$ $\sum x^2$ $\sum y$ $\sum y^2$ & $\sum xy$ or Attempt at S_{xx} S_{yy} & S_{xy} Attempt at correct corresponding formula for r $r = 0.893$ to 0.8933	B3 (B2) (B1) (M1) (m1) (A1)		AWFW (0.89319) AWFW AWFW 561 30667 671 42613 & 35882 (all 5 attempted) 2056 1682 & 1661 (all 3 attempted) AWFW
(b)	Fairly strong / strong / very strong positive (linear) correlation / relationship / association / link (but not trend) between length and weight of adult snakes	B1dep	3	Or equivalent; must qualify strength and indicate positive Dependant on $0.8 \leq r \leq 0.95$ B0 for some/average/medium/etc
(c)	Figure 1: 5 correct labelled points 4 or 3 correct labelled points	B2 (B1)	2	Deduct 1 mark if points not labelled
(d)(i)	D and G	B1	1	Both CAO
(ii)	$r = 0.25$ to 0.75 Fairly weak / weak / some / moderate positive (linear) correlation / relationship / association / link Do not accept comparison with value in (a) or statement in (b)	B1 B1dep	2	AWFW (0.48790) No penalty for calculation Accept a range only if whole of it falls within 0.25 to 0.75 Or equivalent; must qualify strength and indicate positive Dependant on $0.25 \leq r \leq 0.75$ B0 for very weak/little/ slightest/hardly any/fair/average/medium/anything involving strong/etc
		Total	10	

Question 3

3 The weight, X grams, of talcum powder in a tin may be modelled by a normal distribution with mean 253 and standard deviation σ .

(a) Given that $\sigma = 5$, determine:

(i) $P(X < 250)$; *(3 marks)*

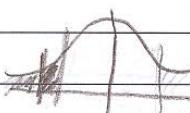
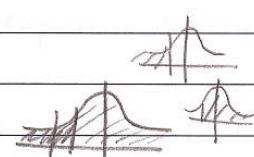
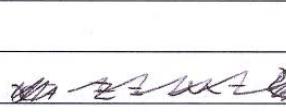
(ii) $P(245 < X < 250)$; *(2 marks)*

(iii) $P(X = 245)$. *(1 mark)*

(b) Assuming that the value of the mean remains unchanged, determine the value of σ necessary to ensure that 98% of tins contain more than 245 grams of talcum powder.

(4 marks)

Student Response

3(a)	$X \sim N(253, 5^2)$	Leave blank
a(i)	$P(X < 250)$	
	$\frac{z = 250 - 253}{5} = -0.6$ ✓	
	$P(X < 250) = 1 - 0.72575$ $= 0.27425$ ✓	3
(ii)	$P(245 < X < 250)$	
	$z = \frac{245 - 253}{5} = -1.6$ ✓	
	$= 0.9452$ ✓	M1
	$P(245 < X < 250) = 0.9452 - 0.72575$ $= 0.22845$ ✗	AO
(iii)	$P(X = 245) = 0$ ✓	B1
b)	mean = 253 for 98% $z = 2.3263$ ✗	 $z = \frac{x - \mu}{\sigma}$
	$2.3263 = \frac{245 - 253}{\sigma}$ sign	BO
	$\sigma = 3.44$	M1
		AI
		AO
		(7)

Commentary

The normal distribution sketches have enabled the candidate to identify the appropriate areas (< 0.5 or > 0.5) and only an arithmetic slip has prevented full marks in part (a)(ii). Many candidates gave an answer of 0.72575 to part (a)(i) and one of 0.67095 to part (a)(ii). Such answers only scored a total of 1 or 2 marks. About 50% of candidates attempted calculations in part (a)(iii) for no reward whatsoever. The above answer to part (b) shows the two most common errors; an incorrect z-value coupled with the mysterious loss of a negative sign!

Mark Scheme

Q	Solution	Marks	Total	Comments
3(a)	$X \sim N(253, 5^2)$			
(i)	$P(X < 250) = P\left(Z < \frac{250 - 253}{5}\right) =$ $P(Z < -0.6) = 1 - P(Z < 0.6)$ $= 1 - 0.72575$ $= 0.274$ to 0.275	M1 m1 A1	3	Standardising (249.5, 250 or 250.5) with 253 and $(\sqrt{5}, 5$ or $5^2)$ and/or $(253 - x)$ Area change; may be implied AWFW (0.27425) $(1 - \text{answer}) \Rightarrow M1 \text{ max}$
(ii)	$P(245 < X < 250) = [C's(a)(i)] - P(X < 245)$ $= (i) - P(Z < -1.6) = 0.27425 - 0.0548$ $= 0.219$ to $0.22(0)$	M1 A1	2	Or equivalent; must be clear correct method if answer incorrect and answer > 0 AWFW (0.21945) M1 A0 for $[1 - (i)] - 0.0548 = 0.67095$ M0 A0 for $0.9452 - [(i)] = 0.67095$ M1 A1 for $0.9452 - [1 - (i)] = 0.21945$
(iii)	$P(X = 245) = 0$ or zero or impossible	B1	1	Ignore any working B0 for 'for impossible to calculate'
(b)	$98\% (0.98) \Rightarrow z = -2.05$ to -2.06 $z = \frac{245 - 253}{\sigma}$ $= -2.0537$ $\sigma = 3.88$ to $3.9(0)$	B1 M1 A1 A1		AWFW; ignore sign (-2.0537) Standardising 245 with 253 and σ ; allow $(253 - 245)$ Only allow: ± 2.05 to ± 2.06 ± 2.32 to ± 2.33 AWFW (3.8954)
Note:			4	Or equivalent inconsistent signs
	$\frac{245 - 253}{\sigma} = 2.0537 \Rightarrow \sigma = 3.8954$ $\Rightarrow B1 M1 A1 A0$		Total 10	

Question 4

- 4 As part of an investigation, a chlorine block is immersed in a large tank of water held at a constant temperature. The block slowly dissolves, and its weight, y grams, is noted x days after immersion. The results are shown in the table.

x days	5	10	15	20	30	40	50	60	75
y grams	47	44	42	38	35	27	23	16	9

- (a) Calculate the equation of the least squares regression line of y on x . (4 marks)
- (b) Hence estimate, to the nearest gram, the initial weight of the block. (1 mark)
- (c) A company which markets the chlorine blocks claims that a block will usually dissolve completely after about 13 weeks. Comment, with justification, on this claim. (3 marks)

Student Response

4)	Leave blank
a) $y = a + bx$ $y = 49.798 - 0.548x$	4
b) $y = 49.798 - (0.548 \times 0)$ $y = 50$ grams	B1
c) $13 \times 7 = 91$ $y = 49.798 - (0.548 \times 91)$ $y = -0.07$	3
<p>The above calculation shows the comment is justified, the block is completely dissolved around 91 days (13 weeks).</p> 8	

Commentary

Undoubtedly, the best answered question on the paper with a mode of 8 marks. The above illustrates a typical concise solution. A correct equation (to appropriate accuracy – not -0.55) followed by the (rather unnecessary) substitution of $x = 0$ with an answer as requested to the nearest gram. In part (c), the candidate has shown the necessary substitution of 91 to give $y \approx 0$ followed by a sensible conclusion. At this final step, a minority of candidates felt that any (minute) weight disproved the claim! Those candidates who substituted $y = 0$ to give $x \approx 91$ were equally rewarded. However comments only referencing 'extrapolation' did not gain full marks.

Mark Scheme

Q	Solution	Marks	Total	Comments
4(a)	$b \text{ (gradient)} = -0.5485 \text{ to } -0.5475$ $b \text{ (gradient)} = -0.55 \text{ to } -0.54$ Omission of -ve sign $a \text{ (intercept)} = 49.7 \text{ to } 49.9$ $a \text{ (intercept)} = 49 \text{ to } 50$ or Attempt at $\sum x \quad \sum x^2 \quad \sum y \quad \& \quad \sum xy \quad (\sum y^2)$ or Attempt at $S_{xx} \quad \& \quad S_{xy}$ Attempt at correct formula for b (gradient) $b \text{ (gradient)} = -0.5485 \text{ to } -0.5475$ $a \text{ (intercept)} = 49.7 \text{ to } 49.9$ Accept a & b interchanged only if identified correctly by a clearly shown equation (stated answers are not sufficient) in (b) or (c)	B2 (B1) (B0) B2 (B1) (M1) (m1) (A1) (A1)	4	AWFW AWFW AWFW AWFW 305 14975 281 & 6980 (10173) (all 4 attempted) 4638.89 & -2542.78 (both attempted) AWFW AWFW If a and b not identified anywhere in question, then: -0.5485 to -0.5475 \Rightarrow B1 49.7 to 49.9 \Rightarrow B1
(b)	C's value of intercept from (a) providing > 47 or Value 50 stated even if (a) incorrect or not attempted	B1F	1	Accept value rounded to nearest integer (50)
(c)	13 weeks \Rightarrow 91 days or $y = -1.1 \text{ to } + 1.1$ $y = 0 \Rightarrow x = 89 \text{ to } 93$ \Rightarrow 13 weeks (approximately)	B1 B1 (B1) (B1)		Stated or used Accept a descriptive answer that includes 91 and a value in range AWFW (-0.08254) AWFW (90.84942) Accept a descriptive answer that includes a value in range and 13 Stated
	Note: B1 B1 or (B1) (B1) are available even if (a) not attempted or Thus claim appears justified or Thus tablet likely to have dissolved or Extrapolation required so cannot comment	B1 dep	3	Or equivalent; ignore reasoning unless contradictory Dependent upon 2 nd B1 in (c) or 2 nd (B1) in (c) Not dependent
	If (B1) for extrapolation maximum mark is 2; other mark available is for 91		Total 8	

Question 5

- 5 A survey of all the households on an estate is undertaken to provide information on the number of children per household.

The results, for the 99 households with children, are shown in the table.

Number of children (x)	1	2	3	4	5	6	7
Number of households (f)	14	35	25	13	9	2	1

- (a) For these 99 households, calculate values for:
- (i) the median and the interquartile range; *(3 marks)*
 - (ii) the mean and the standard deviation. *(3 marks)*
- (b) In fact, 163 households were surveyed, of which 64 contained no children.
- (i) For all 163 households, calculate the value for the mean number of children per household. *(2 marks)*
 - (ii) State whether the value for the standard deviation, when calculated for all 163 households, will be smaller than, the same as, or greater than that calculated in part (a)(ii). *(1 mark)*
 - (iii) It is claimed that, for all 163 households on the estate, the number of children per household may be modelled approximately by a normal distribution.
Comment, with justification, on this claim. Your comment should refer to a fact other than the discrete nature of the data. *(2 marks)*

Student Response

Sai median = $\left(\frac{n+1}{2}\right)^{\text{th}}$ piece of data = $\frac{99+1}{2} = 50^{\text{th}}$ ✓

- 3 children ✓

iQR = UQ - LQ

$$\text{UQ} = \frac{3(99+1)}{4} = 75^{\text{th}} \quad - \quad \text{LQ} = \frac{99+1}{4} = 25^{\text{th}}$$

$$= 4 \text{ children} \quad - \quad = 2 \text{ children}$$

$$= 2 \text{ children}$$

Leave blank

B1

Sai

$$\text{mean} = \frac{\sum x}{n} = \frac{275}{99} = 2.78 \text{ (3sf)}$$

$$sd = \sqrt{\frac{\sum x^2}{n-1}} = \sqrt{\frac{1522}{98}} = 1.31 \text{ (3sf)}$$

B1

B2

M1

A1

Sbi mean = 275 ✓ = 1.69 (3sf)

IG3

Sbi It will be larger ✓

Sbi It can not be modelled by a normal distribution as it is not a random sample, people will decide how many children they want to have.

O

(9)

Commentary

Unlike here, it was not unusual to simply see stated correct answers to parts (a)(i) & (ii). Where such stated answers were incorrect, often 3 or even 6 marks were lost. Many candidates, as here, also answered parts (b)(i) & (ii) correctly; some even obtaining 1.69 after making a total mess of part (a)(ii). Almost all answers to part (a)(iii), as here, were incorrect though usually claiming 'correct' due to either a large sample or the CLT. This revealed a marked lack of knowledge of the latter. **A large (>30 say) sample enables one to assume that the *sample mean* is approximately normally distributed; not the sample and certainly not the population!**

Mark Scheme

Q	Solution	Marks	Total	Comments
5(a) (i)	Median (50) = 3 If not identified, then assume order is median then IQR $IQR(75 - 25) = 4 - 2 = 2$	B1 B2		CAO Do not award marks if correct answers are based on shown incorrect method; eg accept use of 99/2, etc but not 276/2, etc CAO; but 25 th value \Rightarrow IQR = 2 \Rightarrow B0
	Special Cases: Identification that LQ = 2 and UQ = 4 Statement of ≥ 4 cumulative frequencies $F: 14 \ 49 \ 74 \ 87 \ 96 \ 98 \ 99$	(B1) (M1)	3	Both CAO Can award if no marks scored in (i) even if then applied to continuous data
(ii)	Mean = $\frac{\sum fx}{\sum f} = \frac{275}{99} = 2.77$ to 2.78 If not identified, assume order is \bar{x} then s $SD (\sum fx^2 = 933) = 1.3(0)$ to 1.32	B1 B2		AWFW (2.778) Treat rounding to integers as ISW AWFW (1.307 & 1.314)
	Special Case: Evidence of $\frac{\sum fx}{99}$	(M1)	3	Can award if no marks scored in (ii)
(b)(i)	Mean ₁₆₃ = $\frac{99 \times \text{Mean}_{99}}{163}$ or $\frac{\sum fx \text{ from (a)(ii)}}{163}$ $= 1.68$ to 1.69	M1		Or equivalent; may be implied by an answer within range
(ii)	Increase	B1	2	AWFW (1.687) CAO; or equivalent (1.696) Ignore any working (1.702)
(iii)	Data is (positively/negatively) skewed / not symmetric / bimodal / not bell-shaped from frequency distribution / given table or [C's mean in (b)(i)] - 2 \times [C's SD in (a)(ii)] < 0 or [C's mean in (b)(i)] - 2 \times [1.69 to 1.71] < 0 Thus claim appears not valid	B1 B1 dep	1 2	Or equivalent (-1.75 to -0.90) Or equivalent Dependent upon previous B1
	Total		11	

Question 6

- 6 (a) The time taken, in minutes, by *Domesat* to install a domestic satellite system may be modelled by a normal distribution with unknown mean, μ , and standard deviation 15.

The times taken, in minutes, for a random sample of 10 installations are as follows.

47 39 25 51 47 36 63 41 78 43

Construct a 98% confidence interval for μ . (5 marks)

- (b) The time taken, Y minutes, by *Teleair* to erect a TV aerial and then connect it to a TV is known to have a mean of 108 and a standard deviation of 28.

Estimate the probability that the mean of a random sample of 40 observations of Y is more than 120. (4 marks)

- (c) Indicate, with a reason, where, if at all, in this question you made use of the Central Limit Theorem. (2 marks)

Student Response

6) a)	$\bar{x} = 47$ ✓	Leave blank	B1
	$98\% \Rightarrow z = 2.0537$ X		B0
	$CI = \bar{x} \pm z \times \frac{\sigma}{\sqrt{n}}$		M1
	$\therefore 98\% CI = 47 \pm 2.0537 \times \frac{15}{\sqrt{40}}$ ✓		A1
	$= (37.26, 56.74)$ X		A0
6) b)	$X \sim N(108, 28^2)$		
	Variance = $\frac{\sigma^2}{n}$		
	$= \frac{784}{40}$		
	$= 19.6$ ✓		
	$\therefore P(Y > 120) = P(z > \frac{120 - 108}{\sqrt{19.6}})$ ✓		
	$= P(z > 2.71)$		
	$= 1 - P(z < 2.71)$		
	$= 1 - 0.99664$		
	$= 0.00336$ ✓		
6) c)	In part (b); the sample size was sufficiently large (i.e. > 30) for the Central Limit Theorem to apply. X		
		(8)	

Commentary

Very few candidates scored full marks here; the above illustrates typical mistakes. Perhaps as a result of Question 3(b), the candidate has used an incorrect z-value in part (a) and so lost 2 of the 5 marks. Part (b) involving standard error is completely correct; illustrating the noticeable improvement. Standardising using $\sigma = 28$ lost all 4 marks. In answering part (c), the candidate has identified (b) correctly but, in common with almost all candidates, has not given the correct reason **that the distribution of Y was unknown**.

Mark Scheme

Q	Solution	Marks	Total	Comments
6(a)	Mean = $\frac{470}{10} = 47$ 98% (0.98) $\Rightarrow z = 2.32$ to 2.33 CI for μ is $\bar{x} \pm z \times \frac{\sigma}{\sqrt{n}}$ Thus $47 \pm 2.3263 \times \frac{15}{\sqrt{10}}$ Hence 47 ± 11.0 to 11.1 Or $(35.9$ to 36.0, 58.0 to 58.1)	B1 B1 M1 A1F A1		CAO AWFW (2.3263) Used Must have \sqrt{n} with $n > 1$ F on \bar{x} and z only CAO & AWRT (accept 11) AWRT (accept 36 & 58)
(b)	$Y \sim N(108, 28^2)$ Variance of $\bar{Y}_{40} = 28^2/40 = 19.6$ $\sqrt{\text{SD of } \bar{Y}_{40}} = 28\sqrt{40} = 4.425$ to 4.43 $P(\bar{Y}_{40} > 120) = P\left(Z > \frac{120 - 108}{28/\sqrt{40}}\right)$ $= P(Z > 2.71) = 1 - P(Z < 2.71)$ $= 1 - 0.99664 = 0.0033$ to 0.0034	B1 M1 m1 A1	5	CAO Stated or used AWFW Standardising 120 with 108 and $\sqrt{19.6}$ or (4.425 to 4.43) or equivalent; allow (108 – 120) Area change; may be implied AWFW (0.00336) (1 – answer) \Rightarrow B1 M1 max
(c)	Part (b) or Teleair times Distribution of Y not known Note: To score B1 B1 there must be both a clear indication of where in question and a valid reason	B1 B1	4	Or equivalent; ignore reasoning Or equivalent; must be clear reference to Y or population B0 for $n > 30$ Any reference to part (a) \Rightarrow B0 B0
			2	
		Total	11	

Question 7

7 Mr Alott and Miss Fewer work in a postal sorting office.

- (a) The number of letters per batch, R , sorted incorrectly by Mr Alott when sorting batches of 50 letters may be modelled by the distribution $B(50, 0.15)$.

Determine:

- (i) $P(R < 10)$; *(4 marks)*
- (ii) $P(5 \leq R \leq 10)$. *(4 marks)*

- (b) It is assumed that the probability that Miss Fewer sorts a letter incorrectly is 0.06, and that her sorting of a letter incorrectly is independent from letter to letter.

- (i) Calculate the probability that, when sorting a batch of 22 letters, Miss Fewer sorts exactly 2 letters incorrectly. *(3 marks)*
- (ii) Calculate the probability that, when sorting a batch of 35 letters, Miss Fewer sorts at least 1 letter incorrectly. *(2 marks)*
- (iii) Calculate the mean and the variance for the number of letters sorted **correctly** by Miss Fewer when she sorts a batch of 120 letters. *(2 marks)*
- (iv) Miss Fewer sorts a random sample of 20 batches, each containing 120 letters. The number of letters sorted **correctly** per batch has a mean of 112.8 and a variance of 56.86.

Comment on the assumptions that the probability that Miss Fewer sorts a letter incorrectly is 0.06, and that her sorting of a letter incorrectly is independent from letter to letter. *(3 marks)*

Student Response

<p>7)</p> <p>a) $X \sim B(50, 0.15)$</p> <p>i) $P(R < 10) = P(R \leq 9)$ $= 0.7911 = 0.791$ ✓</p> <p>ii) $P(5 \leq R \leq 10)$ $= P(R \leq 10) - P(\leq 5)$ $= 0.8801 - 0.2194$ ✓ $= 0.661$ ✗</p> <p>b) i) $\binom{22}{\text{extra}} \times 0.06^2 \times 0.94^{20}$ ✓ $X \sim B(22, 0.06)$ $= 0.241$ ✓</p> <p>ii) $X \sim B(35, 0.06)$ $\binom{35}{1} \times 0.06^1 \times 0.94^{34} = 0.256$ ✗ $1 - 0.256 = 0.744$ ✗</p> <p>iii) Mean = $np = 120 \times 0.94 = 112.8$ ✓ Varianc$e = np(1-p) = 112.8 \times (0.06) = 6.77$ ✓</p> <p>iv) The means both have the same value of 112.8, however the standard deviations differ extensively. The probability of 0.06 is therefore incorrect. The assumption that each letter is sorted independently cannot be commented on by only looking at the mean and variance. We would need more information.</p>	Leave blank B1 M2 3 0 2 B1 B0 B0 9
---	---

Commentary

Whilst most candidates scored quite well on this question, few gained full or nearly full marks. Marks were often lost in part (a) for quoting 0.8801 as the answer to (i) and/or, as here, using one incorrect value in (ii). Whilst it was very rare indeed to see an incorrect answer to part (b)(i), it was equally rare to see a correct answer to part (b)(ii). The above illustrates a typical error of calculating $1 - P(1)$ rather than $1 - P(0)$. In common with this candidate, almost all candidates were able to score full marks in part (b)(iii), though a minority ignored the word '**correctly**'. In order to score any marks in part (b)(iv), a **comparison** of means and/or variances was required. Whilst this candidate has so done for the first mark, subsequent statements are incorrect conclusions.

Mark Scheme

Q	Solution	Marks	Total	Comments
7(a)	$R \sim B(50, 0.15)$			
(i)	$P(R < 10) = 0.791$	B1		AWRT (0.7911)
(ii)	$P(5 \leq R \leq 10) = 0.8801$ or 0.7911 (p_1) minus 0.1121 or 0.2194 (p_2) $= 0.768$	M1 M1 A1		Accept 3 dp accuracy $(1 - p_2) - p_1 \Rightarrow M0 M0 A0$ $p_1 - (1 - p_2) \Rightarrow M1 M0 A0$ only providing result > 0 Accept 3 dp accuracy
	or $B(50, 0.15)$ expressions stated for at least 3 terms within $4 \leq R \leq 10$ gives probability $= 0.768$	(M1) (A2)	4	Can be implied by correct answer AWRT
(b)	Confusion of 22, 35, 120 and/or 0.15, 0.06			Do not treat as misreads
(i)	$S \sim B(22, 0.06)$	M1		Used in (b)(i) as evidenced by any correct binomial term for $S > 0$
	$P(S = 2) = \binom{22}{2} (0.06)^2 (0.94)^{20}$ $= 0.24$ to 0.242	A1 A1	3	Can be implied by correct answer Ignore any additional terms AWFW (0.24125)
(ii)	$P(S \geq 1) = 1 - q^{35}$ where $0.84 \leq q \leq 0.96$ $= 0.885$ to 0.89	M1 (B1)	2	Can be implied by correct answer Award for $(0.94)^{35}$ seen in an expression but not if accompanied by a multiplier $\neq 1$ AWFW (0.88532)
(iii)	Mean = $np = 120 \times 0.94 = 112.8$ or 113 If not identified, assume order is μ then σ^2 Variance = $np(1-p)$ $= 120 \times 0.94 \times 0.06 = 6.76$ to 6.78	B1 B1	2	Either Must clearly state variance value AWFW (6.768)
(iv)	Means are (approximately) the same stated or Variances are (very) different stated Agree with $P(\text{sorts letter incorrectly}) = 0.06$ Disagree with independent from letter to letter	B1 B1 dep B1 dep	3	Must have scored 1 st B1 in (iii) Must have scored 2 nd B1 in (iii) Dependent on 'means same' stated Dependent on 'variances different' stated
		Total	14	