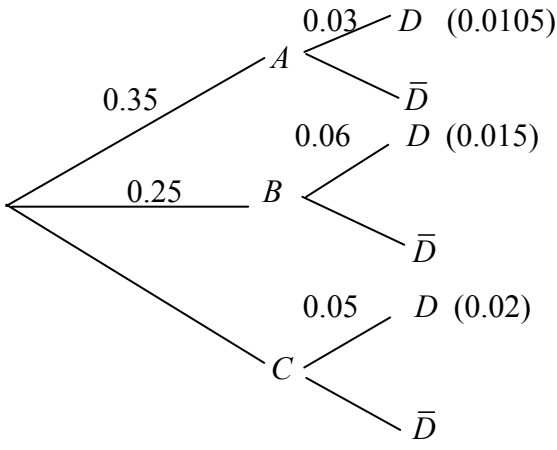


January 2007
6683 Statistics S1
Mark Scheme

Question number	Scheme	Marks
1. (a)	(£) 17	Just 17
(b)	$\sum t = 212$ and $\sum m = 61$ (Accept as totals under each column in qu.)	B1, B1
	$S_{tm} = 2485 - \frac{61 \times 212}{10}, = 1191.8$	awrt 1190 or 119 (3sf)
	$S_{tt} = 983.6$ (awrt 984) and $S_{mm} = 1728.9$ (awrt 1730)	(or 98.4 and 173)
(c)	$r = \frac{1191.8}{\sqrt{983.6 \times 1728.9}}$ $= 0.913922\dots$	M1, A1 M1, A1ft. awrt 0.914
(d)	0.914 (Must be the same as (c) or awrt 0.914)	A1 (3)
	e.g. linear transformation, coding does not affect coefficient (or recalculate)	B1ft. ($ r < 1$) dB1 (2)
(e)	0.914 suggests longer spent shopping the more spent. (Idea more time, more spent)	B1
	0.178 different amounts spent for same time.	B1 (2)
(f)	e.g. might spend short time buying 1 expensive item <u>OR</u> might spend a long time checking for bargains, talking, buying lots of cheap items.	B1g (1)
15 marks		
(b)	M1 for one correct formula seen, f.t. their $\sum t, \sum m$ [Use 1 st A1 for 1 correct, 2 nd A1 for 2 etc]	
(c)	M1 for attempt at correct formula, $\frac{2485}{\sqrt{2101 \times 5478}}$ scores M1A0A0 A1ft f.t. their values for S_{tt} etc from (b) but don't give for $S_{tt} = 5478$ etc (see above) Answer only (awrt 0.914) scores 3/3, 0.913 (i.e. truncation) can score M1A1ft by implication.	
(d)	2 nd B1 dependent on 1 st B1 Accept $\sum m = 261, \sum m^2 = 8541, \sum tm = 6725 \rightarrow 0.914$	
(e)	One mark for a sensible comment relating to each coefficient For 0.178 allow "little or no link between time and amount spent". Must be in context. Just saying 0.914 is strong +ve correlation between amount spent and time shopping and 0.178 is weak correlation ...scores B0B0.	
(f)	B1g for a sensible, practical suggestion showing that other factors might affect the amount spent. E.g. different day (weekend vs weekday) or time of day (time spent queuing if busy)	

Question Number	Scheme	Marks
2. (a)	 <p data-bbox="829 313 1228 627">Correct tree shape A, B and C and 0.35 and 0.25 D (x3) and 0.03, 0.06, 0.05 (May be implied by seeing $P(A \cap D)$ etc at the ends)</p>	<p data-bbox="1244 313 1292 347">M1</p> <p data-bbox="1244 392 1292 425">A1</p> <p data-bbox="1244 481 1292 515">A1 (3)</p>
(b)(i)	$P(A \cap D) = 0.35 \times 0.03, = \underline{\mathbf{0.0105}}$ or $\frac{21}{2000}$	<p data-bbox="1244 795 1340 828">M1, A1</p>
	$P(C) = 0.4 \text{ (anywhere)}$	<p data-bbox="1244 884 1292 918">B1</p>
(ii)	$P(D) = (i) + 0.25 \times 0.06 + (0.4 \times 0.05)$ $= \underline{\mathbf{0.0455}}$ or $\frac{91}{2000}$	<p data-bbox="1244 929 1292 963">M1</p> <p data-bbox="1244 1008 1292 1041">A1 (5)</p>
(c)	$P(C D) = \frac{P(C \cap D)}{P(D)}, = \frac{0.4 \times 0.05}{(ii)}$ $= 0.43956... \text{ or } \frac{40}{91}$	<p data-bbox="1244 1164 1372 1198">M1, A1ft</p> <p data-bbox="1244 1265 1292 1299">A1 (3)</p> <p data-bbox="973 1265 1212 1299">0.44 or awrt 0.440</p>
[Correct answers only score full marks in each part]		11 marks
(a)	<p data-bbox="159 1400 1436 1444">M1 for tree diagram, 3 branches and then two from each. At least one probability attempted.</p> <p data-bbox="159 1456 1037 1500">1st M1 for 0.35×0.03. Allow for equivalent from <u>their</u> tree diagram.</p> <p data-bbox="159 1512 1372 1556">B1 for $P(C) = 0.4$, can be in correct place on tree diagram or implied by 0.4×0.05 in $P(D)$.</p> <p data-bbox="159 1568 1436 1668">2nd M1 for all 3 cases attempted and <u>some</u> correct probabilities seen, including +. Can fit their tree. Condone poor use of notation if correct calculations seen. E.g. $P(C D)$ for $P(C \cap D)$.</p> <p data-bbox="159 1680 1404 1836">(c) M1 for attempting correct ratio of probabilities. There must be an attempt to substitute some values in a correct formula. If no correct formula and ratio not correct ft score M0. Writing $P(D C)$ and attempting to find this is M0. Writing $P(D C)$ but calculating correct ratio – ignore notation and mark ratios.</p> <p data-bbox="159 1904 861 1948">A1ft must have their 0.4×0.05 divided by their (ii).</p> <p data-bbox="159 1960 1436 2004">If ratio is incorrect ft (0/3) unless correct formula seen and part of ratio is correct then M1.</p>	

Question Number	Scheme	Marks														
3. (a)	<p>N.B. Part (a) doesn't have to be in a table, could be a list $P(X = 1) = \dots$ etc</p> <table border="1" data-bbox="236 349 912 510"> <tr> <td>x</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> <td>6</td> </tr> <tr> <td>$P(X = x)$</td> <td>$\frac{1}{36}$</td> <td>$\frac{3}{36}$</td> <td>$\frac{5}{36}$</td> <td>$\frac{7}{36}$</td> <td>$\frac{9}{36}$</td> <td>$\frac{11}{36}$</td> </tr> </table> <p>0.0278, 0.0833, 0.139, 0.194, 0.25, 0.306 (Accept awrt 3 s.f)</p> <p>(b) $P(3) + P(4) + P(5) = \frac{21}{36}$ or $\frac{7}{12}$ or awrt 0.583</p> <p>(c) $E(X) = \frac{1}{36} + 2 \times \frac{3}{36} + \dots = \frac{161}{36}$ or 4.472 or $4\frac{17}{36}$</p> <p>(d) $E(X^2) = \frac{1}{36} + 2^2 \times \frac{3}{36} + \dots = \frac{791}{36}$ or full expression or $21\frac{35}{36}$ or awrt 21.97</p> <p>$\text{Var}(X) = \frac{791}{36} - \left(\frac{161}{36}\right)^2 = \underline{\underline{1.9714\dots}}$ *</p> <p>(e) $\text{Var}(2 - 3X) = 9 \times 1.97$ or $(-3)^2 \times 1.97 = 17.73$ awrt <u>17.7</u> or $\frac{2555}{144}$</p>	x	1	2	3	4	5	6	$P(X = x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$	<p>B1, B1, B1 (3)</p> <p>M1, A1 (2)</p> <p>M1, A1 (2)</p> <p>M1, A1</p> <p>M1, A1c.s.o. (4)</p> <p>M1, A1 (2)</p> <p>13 marks</p>
x	1	2	3	4	5	6										
$P(X = x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$										
	<p>(a) 1st B1 for $x = 1, \dots, 6$ and at least one correct probability N.B. $\frac{3}{36} = \frac{1}{12}$ and $\frac{9}{36} = \frac{1}{4}$ 2nd B1 for at least 3 correct probabilities 3rd B1 for a fully correct probability distribution.</p> <p>(b) M1 for attempt to add the correct three probabilities, fit their probability distribution</p> <p>(c) M1 for a correct attempt at $E(X)$. Minimum is as printed. Exact answer only scores M1A1. [Division by 6 at any point scores M0, no ISW. Non-exact answers with no working score M0.]</p> <p>(d) 1st M1 for a correct attempt at $E(X^2)$. Minimum as printed. $\frac{791}{36}$ or awrt 21.97 scores M1A1. 2nd M1 for their $E(X^2) - (\text{their } E(X))^2$. 2nd A1 cso needs awrt 1.97 <u>and</u> $\frac{791}{36} - \left(\frac{161}{36}\right)^2$ or $\frac{2555}{1296}$ or any fully correct expression seen. Can accept <u>at least 4 sf</u> for both. i.e. 21.97 for $\frac{791}{36}$, 4.472 for $\frac{161}{36}$, 20.00 for $\left(\frac{161}{36}\right)^2$.</p> <p>(e) M1 for correct use of $\text{Var}(aX + b)$ formula or a <u>full</u> method. NB $-3^2 \times 1.97$ followed by awrt 17.7 scores M1A1 <u>BUT</u> $-3^2 \times 1.97$ alone, or followed by -17.7, scores M0A0.</p>															

Question Number	Scheme	Marks
4. (a)	Positive skew (both bits)	B1 (1)
(b)	$19.5 + \frac{(60-29)}{43} \times 10 = 26.7093\dots$ (N.B. Use of 60.5 gives 26.825... so allow awrt 26.8)	awrt 26.7 M1, A1 (2)
(c)	$\mu = \frac{3550}{120} = 29.5833\dots$ or $29\frac{7}{12}$ $\sigma^2 = \frac{138020}{120} - \mu^2$ or $\sigma = \sqrt{\frac{138020}{120} - \mu^2}$ $\sigma = 16.5829\dots$ or ($s = 16.652\dots$)	awrt 29.6 B1 M1 awrt 16.6 (or $s = 16.7$) A1 (3)
(d)	$\frac{3(29.6 - 26.7)}{16.6}$ $= 0.52\dots$ (N.B. 60.5 in (b) ...awrt 0.499[or with s awrt 0.497])	M1A1ft awrt 0.520 (or with s awrt 0.518) A1 (3)
(e)	$0.520 > 0$ So it is consistent with (a)	correct statement about their (d) being >0 or <0 ft their (d) B1ft dB1ft (2)
(f)	Use <u>Median</u> Since the data is skewed <u>or</u> less affected by outliers/extreme values	B1 dB1 (2)
(g)	If the data are <u>symmetrical</u> or <u>skewness is zero</u> or <u>normal/uniform distribution</u> (“mean =median” or “no outliers” or “evenly distributed” all score B0)	B1 (1) 14 marks
(b)	M1 for $(19.5 \text{ or } 20) + \frac{(60-29)}{43} \times 10$ or better. Allow 60.5 giving awrt 26.8 for M1A1 Allow their $0.5n$ [or $0.5(n+1)$] instead of 60 [or 60.5] for M1.	
(c)	M1 for a correct expression for σ, σ^2, s or s^2 . NB $\sigma^2 = 274.99$ and $s^2 = 277.30$ Condone poor notation if answer is awrt 16.6 (or 16.7 for s)	
(d)	M1 for attempt to use this formula using their values to any accuracy. Condone missing 3. 1 st A1ft for using their values to at least 3sf. Must have the 3. 2 nd A1 for using accurate enough values to get awrt 0.520 (or 0.518 if using s) NB Using only 3 sf gives 0.524 and scores M1A1A0	
(e)	1 st B1 for saying or implying correct sign for their (d). B1g and B1ft. Ignore “correlation” if seen. 2 nd B1 for a comment about consistency with their (d) and (a) being positive skew, ft their (d) only This is dependent on 1 st B1: so if (d) >0 , they say yes, if (d) <0 they say no.	
(f)	2 nd B1 is dependent upon choosing median.	

Question Number	Scheme	Marks
5. (a)	Time is a <u>continuous</u> variable <u>or</u> data is in a <u>grouped</u> frequency table	B1 (1)
(b)	Area is proportional to frequency <u>or</u> $A \propto f$ or $A = kf$	B1 (1)
(c)	$3.6 \times 2 = 0.8 \times 9$ <p>1 child represented by 0.8</p>	M1 dM1 A1 cso (3)
(d)	$(\text{Total}) = \frac{24}{0.8} = \underline{\underline{30}}$	M1, A1 (2)
7 marks		
(b)	<p>1st B1 for one of these correct statements. “Area proportional to frequency density” or “Area = frequency” is B0</p>	
(c)	<p>1st M1 for a correct combination of any 2 of the 4 numbers: 3.6, 2, 0.8 and 9 e.g. 3.6×2 or $\frac{3.6}{0.8}$ or $\frac{0.8}{2}$ etc BUT e.g. $\frac{3.6}{2}$ is M0</p> <p>2nd M1 dependent on 1st M1 and for a correct combination of 3 numbers leading to 4th. May be in separate stages but must see all 4 numbers</p> <p>A1cso for fully correct solution. Both Ms scored, no false working seen and <u>comment required</u>.</p>	
(d)	M1 for $\frac{24}{0.8}$ seen or implied.	

Question Number	Scheme	Marks
6. (a)	<p>Used to simplify <u>or</u> represent a real world problem Cheaper <u>or</u> quicker <u>or</u> easier (than the real situation) <u>or</u> more easily modified To improve understanding of the real world problem Used to predict outcomes from a real world problem (idea of predictions)</p>	<p>(any two lines) B1 B1 (2)</p>
(b)	(3 or 4) Model used to make predictions. (Idea of predicted values based on the model)	B1
(b)	(4 or 3) (Experimental) data collected	B1
(b)	(7) Model is refined.	B1 (3)
5 marks		
(a)	<p>1st B1 For one line 2nd B1 For a second line Be generous for 1st B1 but stricter for B1B1</p>	
(b)	<p>1st & 2nd B1 These two points can be interchanged. Idea of values from (experimental) data and predicted values based on the model. 1st B1 for predicted values from model e.g. “model used to gain suitable data” 2nd B1 for data collected. Idea of experimental data but “experiment” needn’t be explicitly seen 3rd B1 This should be stage 7. Idea of refinement or revision or adjustment</p>	

Question Number	Scheme	Marks
7. (a)	$P(X < 91) = P\left(Z < \frac{91 - 100}{15}\right)$ $= P(Z < -0.6)$ $= 1 - 0.7257$ $= 0.2743$	Attempt standardisation M1 A1 M1 awrt 0.274 A1 (4)
(b)	$1 - 0.2090 = 0.7910$ $P(X > 100+k) = 0.2090 \quad \text{or} \quad P(X < 100+k) = 0.7910 \quad (\text{May be implied})$ $\frac{100+k-100}{15} = 0.81 \quad (\text{ft their } z = 0.81, \text{ but must be } z \text{ not prob.})$ $\underline{k = 12}$	0.791 B1 M1 B1 M1, A1ft A1 cao (6) 10 marks
(a)	$\pm \frac{(91 - \mu)}{\sigma \text{ or } \sigma^2}$	
(b)	$\frac{X-100}{15}$	
NB	Using 0.7910 instead of 0.81 gives 11.865 which might be rounded to 12. This should score no more than B1M1B0M1A0A0.	