edexcel
advancing learning, changing lives

Mark Scheme (Results)
January 2009

GCE

GCE Mathematics (6683/01)

Edexcel Limited. Registered in England and Wales No. 4496750
Registered Office: One90 High Holborn, London WC1V 7BH

A PEARSON COMPANY

January 2009
6683 Statistics S1
Mark Scheme

Question Number	Scheme	Marks
1 (a)	$S_{xx} = 57.22 - \frac{(21.4)^2}{10} = 11.424$	M1 A1
	$S_{xy} = 313.7 - \frac{21.4 \times 96}{10} = 108.26$	A1 (3)
(b)	$b = \frac{S_{xy}}{S_{xx}} = 9.4765...$	M1 A1 M1
	$a = \bar{y} - b\bar{x} = 9.6 - 2.14b = (-10.679...)$	A1 (4)
(c)	$y = -10.7 + 9.48x$ Every (extra) hour spent using the programme produces about 9.5 marks improvement	B1ft (1)
(d)	$y = -10.7 + 9.48 \times 3.3, = 20.6$ awrt 21	M1,A1 (2)
(e)	Model may not be valid since [8h is] outside the range [0.5 - 4].	B1 (1) **[11]**

(a) M1 for a correct expression
1st A1 for AWRT 11.4 for S_{xx}
2nd A1 for AWRT 108 for S_{xy}
Correct answers only: One value correct scores M1 and appropriate A1, both correct M1A1A1

(b)
1st M1 for using their values in correct formula
1st A1 for AWRT 9.5
2nd M1 for correct method for *a* (minus sign required)
2nd A1 for equation with *a* and *b* AWRT 3 sf (e.g. $y = -10.68 + 9.48x$ is fine)
Must have a full equation with *a* and *b* correct to awrt 3 sf

(c)
B1ft for comment conveying the idea of *b* marks per hour. Must mention value of *b* but can ft their value of *b*. No need to mention "extra" but must mention "marks" and "hour(s)" e.g. " …9.5 times per hour …" scores B0

(d)
M1 for sub $x = 3.3$ into their regression equation from the end of part (b)
A1 for awrt 21

(e) B1 for a statement that says or implies that it may not be valid because outside the range. They do not have to mention the values concerned here namely 8 h or 0.5 - 4

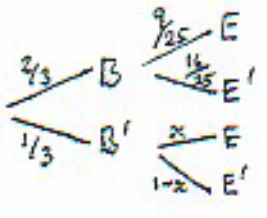
Question Number	Scheme	Marks
2 (a)	E = take regular exercise B = always eat breakfast $\mathrm{P}(E \cap B) = \mathrm{P}(E \mid B) \times \mathrm{P}(B)$ $= \frac{9}{25} \times \frac{2}{3} = 0.24$ or $\frac{6}{25}$ or $\frac{18}{75}$	M1 A1 (2)
(b)	$\mathrm{P}(E \cup B) = \frac{2}{3} + \frac{2}{5} - \frac{6}{25} = \frac{62}{75}$ or $\mathrm{P}(E' \mid B') = \frac{13}{25}$ or $\mathrm{P}(B' \cap E) = \frac{12}{75}$ or $\mathrm{P}(B \cap E') = \frac{32}{75}$ $\mathrm{P}(E' \cap B') = 1 - \mathrm{P}(E \cup B) = \frac{13}{75}$ or $0.17\dot{3}$	M1 A1 M1 A1 (4)
(c)	$\mathrm{P}(E \mid B) = 0.36 \neq 0.40 = \mathrm{P}(E)$ or $\mathrm{P}(E \cap B) = \frac{6}{25} \neq \frac{2}{5} \times \frac{2}{3} = \mathrm{P}(E) \times \mathrm{P}(B)$ So E and B are <u>not</u> statistically independent	M1 A1 (2) **[8]**

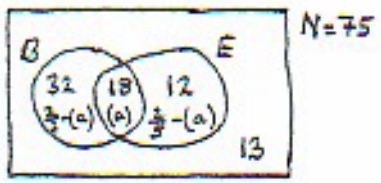
(a) M1 for $\frac{9}{25} \times \frac{2}{3}$ or $\mathrm{P}(E|B) \times \mathrm{P}(B)$ <u>and</u> at least one correct value seen. A1 for 0.24 or exact equiv.
NB $\frac{2}{5} \times \frac{2}{3}$ alone or $\frac{2}{5} \times \frac{9}{25}$ alone scores M0A0. Correct answer scores full marks.

(b) 1st M1 for use of the addition rule. Must have 3 terms and some values, can ft their (a)
<u>Or</u> a full method for $\mathrm{P}(E'|B')$ requires 1 - $\mathrm{P}(E|B')$ and equation for $\mathrm{P}(E|B')$: (a) $+\frac{x}{3} = \frac{2}{5}$
<u>Or</u> a full method for $\mathrm{P}(B' \cap E)$ <u>or</u> $\mathrm{P}(B \cap E')$ [or other valid method]
2nd M1 for a method leading to answer e.g. $1 - \mathrm{P}(E \cup B)$
<u>or</u> $\mathrm{P}(B') \times \mathrm{P}(E' \mid B')$ <u>or</u> $\mathrm{P}(B') - \mathrm{P}(B' \cap E)$ <u>or</u> $\mathrm{P}(E') - \mathrm{P}(B \cap E')$
<u>Venn Diagram</u> 1st M1 for diagram with attempt at $\frac{2}{5} - \mathrm{P}(B \cap E)$ or $\frac{2}{3} - \mathrm{P}(B \cap E)$. Can ft their (a)
1st A1 for a correct first probability as listed or 32, 18 and 12 on Venn Diagram
2nd M1 for attempting 75 - their (18 +32 + 12)

(c) M1 for identifying suitable values to test for independence e.g. $\mathrm{P}(E) = 0.40$ and $\mathrm{P}(E|B) = 0.36$
<u>Or</u> $\mathrm{P}(E) \times \mathrm{P}(B) = \ldots$ and $\mathrm{P}(E \cap B) =$ their (a) [but their (a) $\neq \frac{2}{5} \times \frac{2}{3}$]. Values seen somewhere
A1 for correct values and a correct comment

<u>Diagrams</u> You may see these or find these useful for identifying probabilities.





<u>Common Errors</u>
(a) $\frac{9}{25}$ is M0A0
(b) $\mathrm{P}(E \cup B) = \frac{53}{75}$ scores M1A0
$1 - \mathrm{P}(E \cup B) = \frac{22}{75}$ scores M1A0
(b) $\mathrm{P}(B') \times \mathrm{P}(E') = \frac{1}{3} \times \frac{3}{5}$
scores 0/4

Question Number	Scheme	Marks
3 (a)	$E(X) = 0\times0.4+1\times0.3+...+3\times0.1, \quad = \ 1$	M1, A1 (2)
(b)	$F(1.5) = [P(X \le 1.5) =] \ P(X \le 1), \quad = 0.4 + 0.3 = 0.7$	M1, A1 (2)
(c)	$E(X^2) = 0^2\times0.4+1^2\times0.3+...+3^2\times0.1 \quad , = \ 2$ $Var(X) = \ 2-1^2 \ , \ = \ 1 \qquad (*)$	M1, A1 M1, A1cso (4)
(d)	$Var(5-3X) = (-3)^2 \ Var(X), \quad = 9$	M1, A1 (2)
(e)	See table below	

Total	Cases	Probability
4	$(X=3)\cap(X=1)$	$0.1\times0.3=0.03$
	$(X=1)\cap(X=3)$	$0.3\times0.1=0.03$
	$(X=2)\cap(X=2)$	$0.2\times0.2=0.04$
5	$(X=3)\cap(X=2)$	$0.1\times0.2=0.02$
	$(X=2)\cap(X=3)$	$0.2\times0.1=0.02$
6	$(X=3)\cap(X=3)$	$0.1\times0.1=0.01$

B1B1B1

M1

A1

Total probability = 0.03 + 0.03+0.04 +0.02 + 0.02 + 0.01 = 0.15 — A1 (6)

[16]

(a) M1 for at least 3 terms seen. Correct answer only scores M1A1. Dividing by $k(\neq 1)$ is M0.

(b) M1 for $F(1.5) = P(X \le 1)$.[**Beware**: $2\times0.2+3\times0.1=0.7$ but scores M0A0]

(c) 1st M1 for at least 2 non-zero terms seen. $E(X^2) = 2$ alone is M0. Condone calling $E(X^2)$ =Var(X).
1st A1 is for an answer of 2 or a fully correct expression.

ALT
2nd M1 for $-\mu^2$, condone 2 – 1, unless clearly 2 - □. Allow $2-\mu^2$, with □ = 1 even if E(X) ≠ 1
2nd A1 for a fully correct solution with no incorrect working seen, **both** Ms required.

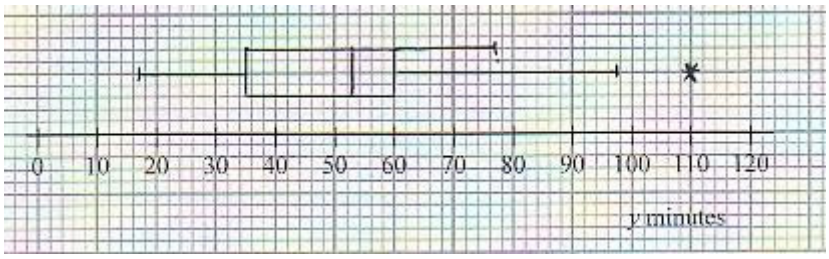
$\sum(x-\mu)^2 \times P(X=x)$

1st M1 for an attempt at a full list of $(x-\mu)^2$ values and probabilities. 1st A1 if all correct
2nd M1 for at least 2 non-zero terms of $(x-\mu)^2 \times P(X=x)$ seen. 2nd A1 for 0.4 + 0.2 + 0.4 = 1

(d) M1 for use of the correct formula. -3^2 Var(X) is M0 unless the final answer is >0.

(e) Can follow through their Var(X) for M1

1st B1 for all cases listed for a total of 4 or 5 or 6 . e.g. (2,2) counted twice for a total of 4 is B0
ALT
2nd B1 for all cases listed for 2 totals }
3rd B1 for a complete list of all 6 cases }These may be highlighted in a table
Using Cumulative probabilities
1st B1 for one or more cumulative probabilities used e.g.2 then 2 or more or 3 then 1 or more
2nd B1 for both cumulative probabilities used. 3rd B1 for a complete list 1, 3; 2, ≥2; 3, ≥1
M1 for one correct pair of correct probabilities multiplied
1st A1 for all 6 correct probabilities listed (0.03, 0.03, 0.04, 0.02, 0.02, 0.01) needn't be added.
2nd A1 for 0.15 or exact equivalent only as the final answer.

Question Number		Scheme	Marks
4	(a)	$Q_2 = 53, \quad Q_1 = 35, \quad Q_3 = 60$	B1, B1,B1 (3)
	(b)	$Q_3 - Q_1 = 25 \Rightarrow Q_1 - 1.5\times 25 = -2.5$ (no outlier) $Q_3 + 1.5\times 25 = 97.5$ (so 110 is an outlier)	M1 A1 (2)
	(c)		M1 A1ft A1ft (3)
	(d)	$\sum y = 461, \sum y^2 = 24\ 219 \therefore\ S_{yy} = 24219 - \frac{461^2}{10}$, $= 2966.9$ (*)	B1, B1, B1cso (3)
	(e)	$r = \frac{-18.3}{\sqrt{3463.6\times 2966.9}}$ or $\frac{-18.3}{3205.64...} = -0.0057$ AWRT - 0.006 or -6×10^{-3}	M1 A1 (2)
	(f)	*r* suggests correlation is close to zero so parent's claim is not justified	B1 (1) **[14]**

(a) 1st B1 for median
2nd B1 for lower quartile
3rd B1 for upper quartile

(b) M1 for attempt to find one limit
A1 for both limits found and correct. No explicit comment about outliers needed.

(c) M1 for a box and two whiskers
1st A1ft for correct position of box, median and quartiles. Follow through their values.
2nd A1ft for 17 and 77 or "their" 97.5 and *. If 110 is not an outlier then score A0 here.
Penalise no gap between end of whisker and outlier. Must label outlier, needn't be with *.
Accuracy should be within the correct square so 97 or 98 will do for 97.5

(d) 1st B1 for $\sum y$ N.B. $\left(\sum y\right)^2 = 212521$ and can imply this mark
2nd B1 for $\sum y^2$ or at least three correct terms of $\sum\left(y-\bar{y}\right)^2$ seen.
3rd B1 for complete correct expression seen leading to 2966.9. So all 10 terms of $\sum\left(y-\bar{y}\right)^2$

(e) M1 for attempt at correct expression for *r*. Can ft their S_{yy} for M1.

(f) B1 for comment rejecting parent's claim on basis of weak or zero correlation
Typical error is "negative correlation so comment is true" which scores B0
Weak negative or weak positive correlation is OK as the basis for their rejection.

Question Number		Scheme	Marks
5	(a)	8-10 hours: width = 10.5 - 7.5 = 3 represented by 1.5cm 16-25 hours: width = 25.5 - 15.5 = 10 so represented by 5 cm 8- 10 hours: height = fd = 18/3 = 6 represented by 3 cm 16-25 hours: height = fd = 15/10 = 1.5 represented by 0.75 cm	B1 M1 A1 (3)
	(b)	$Q_2 = 7.5 + \frac{(52-36)}{18} \times 3 = 10.2$ $Q_1 = 5.5 + \frac{(26-20)}{16} \times 2 [= 6.25 \text{ or } 6.3]$ or $5.5 + \frac{(26.25-20)}{16} \times 2$ [=6.3] $Q_3 = 10.5 + \frac{(78-54)}{25} \times 5 [= 15.3]$ or $10.5 + \frac{(78.75-54)}{25} \times 5$ [=15.45 \15.5] IQR = (15.3 - 6.3) = 9	M1 A1 A1 A1 A1ft (5)
	(c)	$\sum fx = 1333.5 \Rightarrow \bar{x} = \frac{1333.5}{104} =$ AWRT 12.8	M1 A1
	(d)	$\sum fx^2 = 27254 \Rightarrow \sigma_x = \sqrt{\frac{27254}{104} - \bar{x}^2} = \sqrt{262.05 - \bar{x}^2}$ AWRT 9.88	M1 A1 (4)
	(e)	$Q_3 - Q_2 [= 5.1] > Q_2 - Q_1 [= 3.9]$ or $Q_2 < \bar{x}$ So data is positively skew	B1ft dB1 (2)
		Use median and IQR, since data is skewed or not affected by extreme values or outliers	B1 B1 (2) **[16]**

(a) M1 For attempting both frequency densities $\frac{18}{3}(=6)$ and $\frac{15}{10}$, and $\frac{15}{10} \times \text{SF}$, where $\text{SF} \neq 1$

(b) NB Wrong class widths(2 and 9) gives $\frac{h}{1.66...} = \frac{3}{9} \rightarrow h = \frac{5}{9}$ or 0.55… and scores M1A0

(c) M1 for identifying correct interval and a correct fraction e.g. $\frac{\frac{1}{2}(104)-36}{18}$. Condone 52.5 or 53
1st A1 for 10.2 for median. Using (n + 1) allow awrt 10.3
2nd A1 for a correct expression for either Q_1 or Q_3 (allow 26.25 and 78.75)
3rd A1 for correct expressions for both Q_1 and Q_3
4th A1ft for IQR, ft their quartiles. Using (n + 1) gives 6.28 and 15.45
NB: Must see some method

(d) 1st M1 for attempting $\sum fx$ and $\bar{x}$
2nd M1 for attempting $\sum fx^2$ and $\sigma_x, \sqrt{}$ is needed for M1. Allow s = awrt 9.93

(e) 1st B1ft for suitable test, values need not be seen but statement must be compatible with values used. Follow through their values
2nd dB1 Dependent upon their test showing positive and for stating positive skew
If their test shows negative skew they can score 1st B1 but lose the second

1st B1 for choosing median and IQR. Must mention both. } Award independently
2nd B1 for suitable reason }
e.g. "use median because data is skewed" scores B0B1 since IQR is not mentioned

Question Number	Scheme	Marks
6 (a)	$P(X < 39) = P\left(Z < \frac{39-30}{5}\right)$	M1
	$= P(Z < 1.8) = \underline{0.9641}$ (allow awrt 0.964)	A1 (2)
(b)	$P(X < d) = P\left(Z < \frac{d-30}{5}\right) = 0.1151$	
	$1 - 0.1151 = 0.8849$	M1
	$\Rightarrow \quad z = -1.2$ (allow $\pm$ 1.2)	B1
	$\therefore \frac{d-30}{5} = -1.2 \qquad \underline{d = 24}$	M1A1 (4)
(c)	$P(X > e) = 0.1151$ so $e = \mu + (\mu$ - their $d)$ or $\frac{e-30}{5} = 1.2$ or – their z	M1
	$\underline{e = 36}$	A1 (2)
(d)	$P(d < X < e) = 1 - 2 \times 0.1151$	M1
	$= 0.7698$ AWRT $\underline{0.770}$	A1 (2)
		[10]

Answer only scores all marks in each section BUT check (b) and (c) are in correct order

(a)

M1 for standardising with σ, $z = \pm\frac{39-30}{5}$ is OK

A1 for 0.9641 or awrt 0.964 but if they go on to calculate 1 – 0.9641 they get M1A0

(b)

1st M1 for attempting 1- 0.1151. Must be seen in (b) in connection with finding d

B1 for $z = \pm$ 1.2. They must state $z = \pm$ 1.2 or imply it is a z value by its use. This mark is only available in part (b).

2nd M1 for $\left(\frac{d-30}{5}\right)$ = their negative z value (or equivalent)

(c) M1 for a full method to find e. If they used $z = 1.2$ in (b) they can get M1 for $z = \pm$ 1.2 here
If they use symmetry about the mean $\mu + (\mu$ - their $d)$ then ft their d for M1
Must explicitly see the method used unless the answer is correct.

(d) M1 for a complete method or use of a correct expression e.g. "their 0.8849" - 0.1151
or **If their *d* < their *e*** using their values with $P(X < e) - P(X < d)$
If their $d \geq$ their e then they can only score from an argument like 1 – 2x0.1151
A negative probability or probability > 1 for part (d) scores M0A0