Edexcel
advancing learning, changing lives

Mark Scheme (Results)
January 2011

GCE

GCE Statistics S1 (6683) Paper 1

A PEARSON COMPANY

Edexcel is one of the leading examining and awarding bodies in the UK and throughout the world. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers.

Through a network of UK and overseas offices, Edexcel's centres receive the support they need to help them deliver their education and training programmes to learners.

For further information, please call our GCE line on 0844 576 0025, our GCSE team on 0844 576 0027, or visit our website at www.edexcel.com.

If you have any subject specific questions about the content of this Mark Scheme that require the help of a subject specialist, you may find our Ask The Expert email service helpful.

Ask The Expert can be accessed online at the following link:

http://www.edexcel.com/Aboutus/contact-us/

General Instructions for Marking

1. The total number of marks for the paper is 75.

2. The Edexcel Mathematics mark schemes use the following types of marks:

   - **M** marks: method marks are awarded for 'knowing a method and attempting to apply it', unless otherwise indicated.
   - **A** marks: Accuracy marks can only be awarded if the relevant method (M) marks have been earned.
   - **B** marks are unconditional accuracy marks (independent of M marks)
   - Marks should not be subdivided.
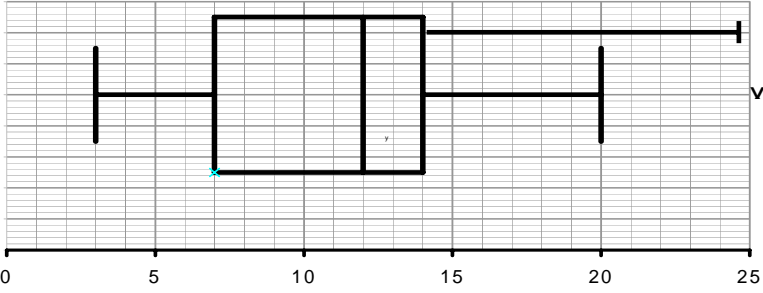
3. Abbreviations

   These are some of the traditional marking abbreviations that will appear in the mark schemes.

   - bod – benefit of doubt
   - ft – follow through
   - the symbol $\sqrt{}$ will be used for correct ft
   - cao – correct answer only
   - cso - correct solution only. There must be no errors in this part of the question to obtain this mark
   - isw – ignore subsequent working
   - awrt – answers which round to
   - SC: special case
   - oe – or equivalent (and appropriate)
   - dep – dependent
   - indep – independent
   - dp decimal places
   - sf significant figures
   - $*$ The answer is printed on the paper
   - $\lfloor$ The second mark is dependent on gaining the first mark

January 2011
Statistics S1 6683
Mark Scheme

Question Number	Scheme	Marks
1. (a)	$S_{ll} = 327754.5 - \dfrac{4027^2}{50} = 3419.92$ $S_{lw} = 29330.5 - \dfrac{357.1 \times 4027}{50} = 569.666$	M1 A1 A1 (3)
(b)	$r = \dfrac{569.666}{\sqrt{3419.92 \times 289.6}} = 0.572$ awrt 0.572 or 0.573	M1 A1 (2)
(c)	As the length of the salmon increases the weight increases	B1ft (1) **[6]**
	Notes	
(a)	M1 for at least one correct expression 1st A1 for S_{ll} = awrt 3420 (Condone S_{xx} =… or even S_{yy} =…) 2nd A1 for S_{lw} = awrt 570 (Condone S_{xy} =…)	

(b)	M1 for attempt at correct formula.
	          Must have their S_{ll}, S_{lw} and given S_{ww} in the correct places
	          If S_{ll}, S_{lw} are correct and an answer of awrt 0.57 is seen then award M1A0
	          M0 for $\dfrac{29330.5}{\sqrt{327754.5 \times 289.6}}$	

(c)	B1ft for a comment mentioning "length" and "weight", not just l and w, and the idea of longer salmon weighing more.
	          e.g. "positive correlation between weight and length" is B0 since the idea of positive correlation is not explained.
	          Allow "larger" instead of "heavier" or "longer"
	          Ignore any spurious values mentioned such as 0.572
	          If their r is negative (but must be $r > -1$) ft an appropriate comment.
	          Condone $r > 1$ if comment is correct.
	          If $|r| < 0.4$ allow a comment of no or little relationship between weight and length but for $0 < r < 0.4$ the printed answer is still acceptable too.
	
	          Treat mention of "skewness" as ISW if a correct interpretation is given	

Question Number	Scheme	Marks
2. (a)	$2.8 + 5.6 + 2.3 + 9.4 + 0.5 + 1.8 + 84.6 = 107$ mean $= 107 / 28 \ (= 3.821\ldots)$ (**awrt 3.8**)	M1 A1 (2)
(b)	It will have no effect since one is 4.5 under what it should be and the other is 4.5 above what it should be.	B1 dB1 (2) [4]
	Notes	

(a)	M1 for a clear attempt to add the two sums. Accept a full expression or
	       $2.8 + 5.6 + \ldots + 84.6 = x$ where $100 < x < 110$
	       i.e. seeing at least two correct terms of Keith's and the 84.6 with a slip.
	 A1 for awrt 3.8 (Condone 1 dp/2sf here since data is given to 1 dp or 2 sf)
	       Accept $\dfrac{107}{28}$ or $3\dfrac{23}{28}$ or any exact equivalent
	 Correct answer implies M1A1	

(b)	1st B1 for clearly stating that it will have no effect. ("roughly the same" is B0 B0)
	 2nd dB1 for a supporting reason that mentions the fact that the increase and decrease are the
	       same and gives some numerical value(s) to support this.
	 e.g. Sum of Keith's observations is still 22.4 (or mean is still 3.2)
	 <u>or</u> Sum is still 107
	 <u>or</u> $9.4 - 4.9 = 5 - 0.5$ (o.e.)
	       This second B1 is dependent on their saying there is no effect so B0B1 is not possible.	

Question Number	Scheme	Marks
3.		
(a)	Outliers $14 + 1.5 \times (14 - 7) = 24.5$ $7 - 1.5 \times (14 - 7) = -3.5$ Outlier 25 either upper limit acceptable on diagram Sales in £'000	M1 A1 M1 A1ft B1 (5)
(b)	Since $Q_3 - Q_2 < Q_2 - Q_1$. Allow written explanation negatively skew	B1 dB1 (2)
(c)	not true since the lower quartile is 7000 and therefore 75% above 7000 not 10000 or 10 is inside the box or any other sensible comment	B1 dB1 (2) [9]
	Notes	

(a)	**A fully correct box-plot (either version) with no supporting work scores 5/5. Otherwise read on**
	1st M1 for at least one correct calculation seen
	1st A1 for 24.5 and -3.5 (or just negative noted) seen. May be read off the graph.
	          If both values are seen but no calculation is given then M1A1,one value M1A0.
	2nd M1 for a box with an upper and a lower whisker(s) with at least 2 correct values
	          (condone no median marked)
	2nd A1ft for 3,7, 12, 14 and 20 or 24.5 in appropriate places and readable off their scale
	          If both upper whiskers are seen A0
	          Apply ft for their <u>whiskers</u> being compatible with their <u>outlier limits</u>
	          e.g. if their lower limit is + 3.5 then a lower whisker ending at 4 or 3.5 is OK
	B1 for only one outlier appropriately marked at 25
	Apply $\pm$ 0.5 square accuracy for diagram	

(b)	1st B1 for $Q_3 - Q_2 < Q_2 - Q_1$ statement or an equivalent statement in words
	          Use of $Q_3 - Q_2 < Q_2 - Q_1$ does not require differences to be seen.
	2nd dB1 for "negative skew" dependent on suitable reason given above. "correlation" is B0
	          "positive skew" with a supporting argument based on whiskers can score B1B1
	          e.g. "right hand whisker is longer than LH one so positive skew"
	          $Q_3 - Q_2 < Q_2 - Q_1$ followed by "positive skew" is B1B0	

(c)	1st B1 for rejecting the company's claim 2nd dB1 for an appropriate supporting reason. Dependent on rejecting company's claim.	

Question Number	Scheme	Marks
4. (a)	$b = \dfrac{1.688}{5.753} = 0.293$ $a = 3.22 - 4.42 \times 0.293 = 1.9231...$ $p = 1.92 + 0.293v$	M1A1 M1 A1 (4)
(b)	$v = \dfrac{85-5}{10} = 8$ $p = 1.92 + 0.293 \times 8 = 4.3$ **(awrt 4.3)**	M1 A1 (2) **[6]**

<div align="center"><u>Notes</u></div>

<div align="center">Can ignore (a) and (b) labels here</div>

(a) 1st M1   for a correct expression for b. $\dfrac{1.688}{1.168}$ is M0

     1st A1   for awrt 0.29

     2nd M1   for use of $a = \overline{p} - b\overline{v}$ follow through their value of b(or even just the letter b)

     2nd A1   for a complete equation with a = awrt 1.92 and b = awrt 0.293

          y or $p = 1.92 + 0.293x$ is A0

<div align="center">Correct answer with no working is 4/4</div>

(b)   M1   for an attempt to find the value of v when $x = 85$ (at least 2 correct terms in

     $\pm\dfrac{85-5}{10}$)

          or for an attempt to find an equation for p in terms of x and using $x = 85$

          Attempt at equation of p in x requires $p = 1.92 + 0.293\dfrac{(x-5)}{10}$

     A1   for awrt 4.3 (award when first seen and apply ISW)

         N.B. $p = 1.92 + 0.293 \times 85$ (o.e.) is M0A0

Question Number	Scheme	Marks
5. (a)	Median = 32/2 = 16th term (16.5) $$\frac{x-39.5}{49.5-39.5}=\frac{16-14}{25-14} \text{ or } x=39.5+\left(\frac{2}{11}\times10\right)$$ Median = 41.3 (use of $n+1$ gives 41.8) **(awrt 41.3)**	M1 A1 (2)
(b)	Mean$=\dfrac{1414}{32}=44.1875$ **(awrt 44.2)** Standard deviation $=\sqrt{\dfrac{69378}{32}-\left(\dfrac{1414}{32}\right)^2}$ $= 14.7$ (or $s = 14.9$)	B1 M1 A1 (3)
(c)	mean > median therefore <u>positive skew</u>	B1ft B1ft (2) **[7]**
	Notes	
(a)	M1 for an attempt to use interpolation to find the median. Condone use of 39 or 40 for 39.5 e.g. allow $39+\dfrac{2}{11}\times10$ (o.e.) or $40+\dfrac{2}{11}\times10$ (o.e.) to score M1A0 but must have the 10 A1 for awrt 41.3 (or awrt 41.8 if using ($n+1$))	
(b) Mid-points	B1 for awrt 44.2 M1 for a correct expression including square root. (Allow ft of their mean) A1 for awrt 14.7 (If using s for awrt 14.9) You may see $\sum t=1339 \rightarrow \bar{t}=41.8$ and $\sum t^2 = 62928 \rightarrow \sigma$ 14.7 or $s = 14.9$ this scores B0 for the mean but can score M1 for a correct st.dev expression and A1 for ans. **Correct answer only in (a) and (b) can score full marks but check (n +1) case in (a)**	
(c) Quartiles	1st B1ft for a correct comparison of their mean and their median (may be in a formula) Calculating median – mean as negative is OK for this B1 but must say +ve skew for 2nd B1 Only allow comparison to be ≈ 0 if $|\text{mean} - \text{median}| \leq 0.5$ 2nd B1ft for a correct description of skewness <u>based on their values of mean and median</u>. ft their values for mean and median not their previous calculation/comparison Must be compatible with their previous comparison (if they have one) "Positive skew" with no reason is B0B1 provided you can see their values that imply that. Description should be "positive" or "negative" or "no" skew or "symmetric" "Positive correlation" is B0 1st B1ft if Q_1 = awrt 32 <u>and</u> Q_3 = awrt 49 seen and a correct comparison made. ft Q_2 2nd B1ft if Q_1 = awrt 32 <u>or</u> Q_3 = awrt 49 seen and a correct description based on their quartiles and their comparison is made. (Should get "negative skew")	

Question Number	Scheme	Marks
6. (a)	$k + 2k + 3k + 4k = 1$ or $10k = 1$ $k = 0.1$ (*) [allow verification with a comment e.g. "so $k = 0.1$"]	B1cso (1)
(b)	$\mathrm{E}(X) = 1 \times 0.1 + 2 \times 0.2 + 3 \times 0.3 + 4 \times 0.4 = 3$	M1 A1 (2)
(c)	$\mathrm{E}(X^2) = 1 \times 0.1 + 4 \times 0.2 + 9 \times 0.3 + 16 \times 0.4 = 10$	M1 A1 (2)
(d)	$\mathrm{Var}(X) = 10 - 9 (= 1)$ $\mathrm{Var}(2 - 5X) = 5^2 \mathrm{Var}(X) = 25$	M1 M1 A1 (3)
(e)	$\mathrm{P}(1,3) + \mathrm{P}(2,2) = 2 \times 0.1 \times 0.3 + 0.2 \times 0.2 = 0.1$ (*)	M1 A1cso (2)
(f)	<table><tr><td>$X_1 + X_2$</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td></tr><tr><td>p</td><td>0.01</td><td>0.04</td><td>0.1</td><td>**0.2**</td><td>0.25</td><td>0.24</td><td>**0.16**</td></tr></table>	B1 B1 (2)
(g)	$\mathrm{P}(2) + \mathrm{P}(3) = 0.05$	M1A1 (2) [14]

Question Number	Scheme	Marks
	Notes	
(a)	B1 for a clear attempt to use sum of probabilities = 1. Must see previous line as well as $k = 0.1$ **A correct expression for** $E(X)$ **or** $E(X^2)$ **that is later divided by 4 scores M0**	
(b)	M1 for a completely correct expression. May be implied by correct answer of 3 or 30k A1 for 3 only.	
(c)	M1 for a completely correct expression. May be implied by correct answer of 10 or 100k A1 for 10 only. [For $E(X^2) = 0.1 + 0.8 + 2.7 + 6.4 - 9 = 1$ scores M0A0 but accept this as Var(X) in (d)]	

(d)	1st M1 for using $Var(X) = E(X^2) - E(X)^2$, f.t their values from (b) and (c)
	       Allow this mark for $Var(X) = 10 - 9$ or better. May be implied if this is seen in (c).
	 2nd M1 for $5^2 Var(X)$ or $25Var(X)$ can f.t. their Var(X). Allow -5^2 if it later becomes +25
	 A1 for 25 only. Dependent upon both Ms
	       Forming distribution for $Y = 2-5X$ gets M1 for E(Y^2)=194 then M1A1 for 194-169=25	

(e)	M1 for correctly identifying (1, 3) <u>or</u> (3, 1) <u>and</u> (2, 2) as required cases ($3k^2 + 4k^2$ or better) A1 cso for 0.1 only but must see evidence for M1	
(f)	1st B1 for 0.2 correctly assigned. May be in table. 2nd B1 for 0.16 correctly assigned. May be in table	
(g)	M1 for P(2) + P(3). May be implied by correct answer of 0.05 A1 for 0.05 only. **Correct answer only can score full marks in parts (b), (c), (f) and (g)**	

Question Number	Scheme	Marks
7. (a)	both $\frac{2}{3},\frac{1}{3}$	B1
	$\frac{4}{9}$	B1
	both $\frac{3}{5},\frac{2}{5}$	B1
	all three of $\frac{4}{9},\frac{4}{9},\frac{5}{9}$	B1 (4)
(b)	$P(A)=P(RR)+P(YY)=\frac{1}{2}\times\frac{2}{5}+\frac{1}{2}\times"\frac{2}{5}"=\frac{2}{5}$ B1 for $\frac{1}{2}\times\frac{2}{5}$ (oe) seen at least once	B1 M1 A1 (3)
(c)	$P(B)=P(RRR)+P(RYR)+P(YRR)+P(YYR)$ M1 for at least 1 case of 3 balls identified. (Implied by 2nd M1)	M1
	$\left(\frac{1}{2}\times\frac{2}{5}\times"\frac{2}{3}"\right)+\left(\frac{1}{2}\times\frac{3}{5}\times\frac{5}{9}\right)+\left(\frac{1}{2}\times"\frac{3}{5}"\times\frac{5}{9}\right)+\left(\frac{1}{2}\times"\frac{2}{5}"\times"\frac{4}{9}"\right)=\frac{5}{9}$ (*)	M1,A1cso (3)
(d)	$P(A\cap B)=P(RRR)+P(YYR)$ M1 for identifying both cases and + probs. may be implied by correct expressions	M1
	$=\left(\frac{1}{2}\times\frac{2}{5}\times\frac{2}{3}\right)+\left(\frac{1}{2}\times\frac{2}{5}\times\frac{4}{9}\right)=\frac{2}{9}$ (*)	A1cso (2)
(e)	$P(A\cup B)=P(A)+P(B)-P(A\cap B)$ Must have some attempt to use	M1
	$="\frac{2}{5}"+\frac{5}{9}-\frac{2}{9}=\frac{11}{15}$	A1cao (2)

Question Number	Scheme		Marks
(f)	$\dfrac{\text{P}(RRR)}{\text{P}(RRR)+\text{P}(YYY)} = \dfrac{\frac{1}{2}\times\frac{2}{5}\times\frac{2}{3}}{\left(\frac{1}{2}\times\frac{2}{5}\times\frac{2}{3}\right)+\left(\frac{1}{2}\times\frac{2}{5}\times\frac{5}{9}\right)} = \dfrac{6}{11}$	Probabilities must come from the product of 3 probs. from their tree diagram.	M1 A1ft A1 cao (3) [17]

Notes

(b)	M1 for both cases, and +, attempted, ft their values from tree diagram. May be 4 cases of 3 balls.
(c)	2nd M1 for all 4 correct expressions, ft their values from tree diagram. A1 is cso
(e)	M1 for clear attempt to <u>use</u> the correct formula, must have some correct substitution. ft their (b)
(f)	M1 for identifying the correct probabilities and forming appropriate fraction of probs. 1st A1ft for a correct expression using probabilities from their tree **Accept exact decimal equivalents. Correct answer only is full marks except in (c) and (d)**

Question Number	Scheme	Marks
8.		
(a)	$P(X > 168) = P\left(Z > \dfrac{168-160}{5}\right)$	M1
	$= P(Z > 1.6)$	A1
	$= 0.0548$ **awrt 0.0548**	A1 (3)
(b)	$P(X < w) = P\left(Z < \dfrac{w-160}{5}\right)$	
	$\dfrac{w-160}{5} = -2.3263$	M1 B1
	$w = 148.37$ **awrt 148**	A1 (3)
(c)	$\dfrac{160-\mu}{\sigma} = 2.3263$	M1 B1
	$\dfrac{152-\mu}{\sigma} = -1.2816$	B1
	$160 - \mu = 2.3263\sigma$	
	$152 - \mu = -1.2816\sigma$	
	$8 = 3.6079\sigma$	M1
	$\sigma = 2.21....$ **awrt 2.22**	A1
	$\mu = 154.84...$ **awrt 155**	A1 (6)
		[12]

Notes

(a)

M1   for an attempt to standardize 168 with 160 and 5 i.e. $\pm\left(\dfrac{168-160}{5}\right)$ or implied by 1.6

1st A1   for $P(Z > 1.6)$ or $P(Z < -1.6)$ ie $z = 1.6$ and a correct inequality or 1.6 on a shaded diagram

Correct answer to (a) implies all 3 marks

(b)

M1   for attempting $\pm\left(\dfrac{w-160}{5}\right) = $ recognizable z value $(|z| > 1)$

B1   for $z = \pm\, 2.3263$ or better. Should be $z = ...$ or implied so: $1 - 2.3263 = \dfrac{w-160}{5}$ is M0B0

A1   for awrt 148. This may be scored for other z values so M1B0A1 is possible

For awrt 148 only with no working seen award M1B0A1

(c)

M1   for attempting to standardize 160 or 152 with μ and σ (allow $\pm$) and equate to z value ($|z| > 1$)

1st B1   for awrt $\pm\, 2.33$ or $\pm\, 2.32$ seen

2nd B1   for awrt $\pm\, 1.28$ seen

2nd M1   for attempt to solve their two linear equations in μ and σ leading to equation in just one variable

1st A1   for $\sigma = $ awrt 2.22 . Award when 1st seen

2nd A1   for $\mu = $ awrt 155. Correct answer only for part (c) can score all 6 marks.

NB $\sigma = 2.21$ commonly comes from $z = 2.34$ and usually scores M1B0B1M1A0A1

The A marks in (c) require both M marks to have been earned