Regression - exam questions

Question 1: Jan 2007

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

Stan is a retired academic who supplements his pension by mowing lawns for customers who live nearby.

As part of a review of his charges for this work, he measures the areas, $x\,m^2$, of a random sample of eight of his customers' lawns and notes the times, y minutes, that it takes him to mow these lawns. His results are shown in the table.

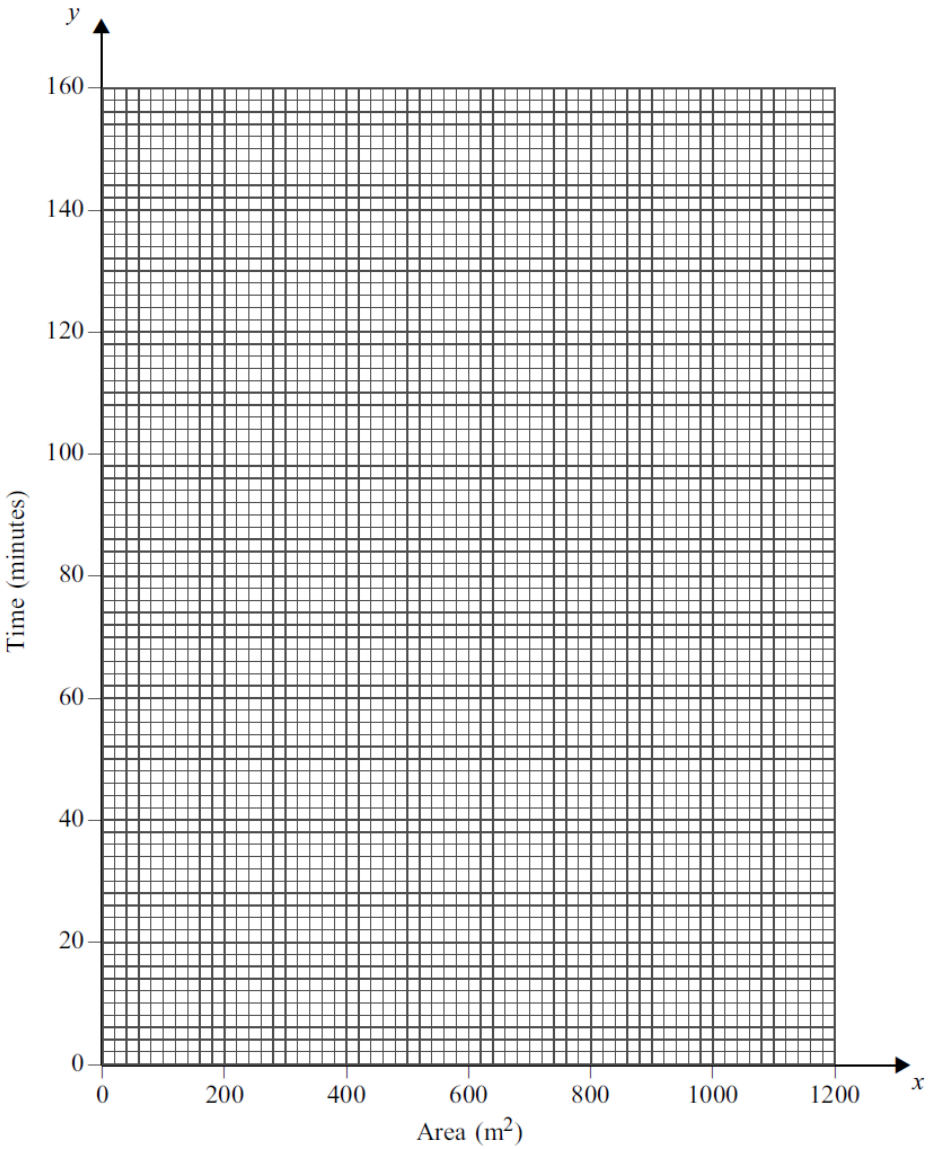
Customer	A	B	C	D	E	F	G	H
x	360	140	860	600	1180	540	260	480
y	50	25	135	70	140	90	55	70

(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. Draw your line on **Figure 1**. *(6 marks)*

(c) Calculate the value of the residual for Customer H and indicate how your value is confirmed by your scatter diagram. *(3 marks)*

(d) Given that Stan charges £12 per hour, estimate the charge for mowing a customer's lawn that has an area of $560\,m^2$. *(4 marks)*



Question 2: Jun 2006

A new car tyre is fitted to a wheel. The tyre is inflated to its recommended pressure of 265 kPa and the wheel left unused. At 3-month intervals thereafter, the tyre pressure is measured with the following results:

Time after fitting (x months)	0	3	6	9	12	15	18	21	24
Tyre pressure (y kPa)	265	250	240	235	225	215	210	195	180

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

(ii) Interpret in context the value for the gradient of your line. *(2 marks)*

(iii) Comment on the value for the intercept with the y-axis of your line. *(2 marks)*

(b) The tyre manufacturer states that, when one of these new tyres is fitted to the wheel of a car and then inflated to 265 kPa, a suitable regression equation is of the form

$$y = 265 + bx$$

The manufacturer also states that, as the car is used, the tyre pressure will decrease at twice the rate of that found in part (a).

(i) Suggest a suitable value for b. *(2 marks)*

(ii) One of these new tyres is fitted to the wheel of a car and inflated to 265 kPa. The car is then used for 8 months, after which the tyre pressure is checked for the first time.

Show that, accepting the manufacturer's statements, the tyre pressure can be expected to have fallen below its minimum safety value of 220 kPa. *(2 marks)*

Question 3: Jan 11

5 Craig uses his car to travel regularly from his home to the area hospital for treatment. He leaves home at x minutes after 7.30 am and then takes y minutes to arrive at the hospital's reception desk.

His results for 11 mornings are shown in the table.

x	0	5	10	15	20	25	30	35	40	45	50
y	31	42	32	58	47	56	79	68	89	95	85

(a) Explain why the time taken by Craig between leaving home and arriving at the hospital's reception desk is the response variable. *(1 mark)*

(b) Calculate the equation of the least squares regression line of y on x, writing your answer in the form $y = a + bx$. *(5 marks)*

(c) On a particular day, Craig needs to arrive at the hospital's reception desk no later than 9.00 am. He leaves home at 7.45 am.

Estimate the number of minutes **before** 9.00 am that Craig will arrive at the hospital's reception desk. Give your answer to the nearest minute. *(5 marks)*

(d) (i) Use your equation to estimate y when $x = 85$. *(1 mark)*

(ii) Give **one** statistical reason and **one** reason based on the context of this question as to why your estimate in part **(d)(i)** is unlikely to be realistic. *(2 marks)*

Question 4: Jun 2007

Bob, a gardener, measures the time taken, y minutes, for 60 grams of weedkiller pellets to dissolve in 10 litres of water at different set temperatures, x °C. His results are shown in the table.

x	16	20	24	28	32	36	40	44	48	52	56
y	4.7	4.3	3.8	3.5	3.0	2.7	2.4	2.0	1.8	1.6	1.1

(a) State why the explanatory variable is temperature. *(1 mark)*

(b) Calculate the equation of the least squares regression line $y = a + bx$. *(4 marks)*

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

(ii) Explain why no sensible practical interpretation can be given for your value of a. *(2 marks)*

(d) (i) Estimate the time taken to dissolve 60 grams of weedkiller pellets in 10 litres of water at 30 °C. *(2 marks)*

(ii) Show why the equation cannot be used to make a valid estimate of the time taken to dissolve 60 grams of weedkiller pellets in 10 litres of water at 75 °C. *(2 marks)*

Question 5: Jun 2008

The table shows the times taken, y minutes, for a wood glue to dry at different air temperatures, x °C.

x	10	12	15	18	20	22	25	28	30
y	42.9	40.6	38.5	35.4	33.0	30.7	28.0	25.3	22.6

(a) Calculate the equation of the least squares regression line $y = a + bx$. *(4 marks)*

(b) Estimate the time taken for the glue to dry when the air temperature is 21 °C. *(2 marks)*

Question 6: Jan 2008

[**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, $x\,m^2$, 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

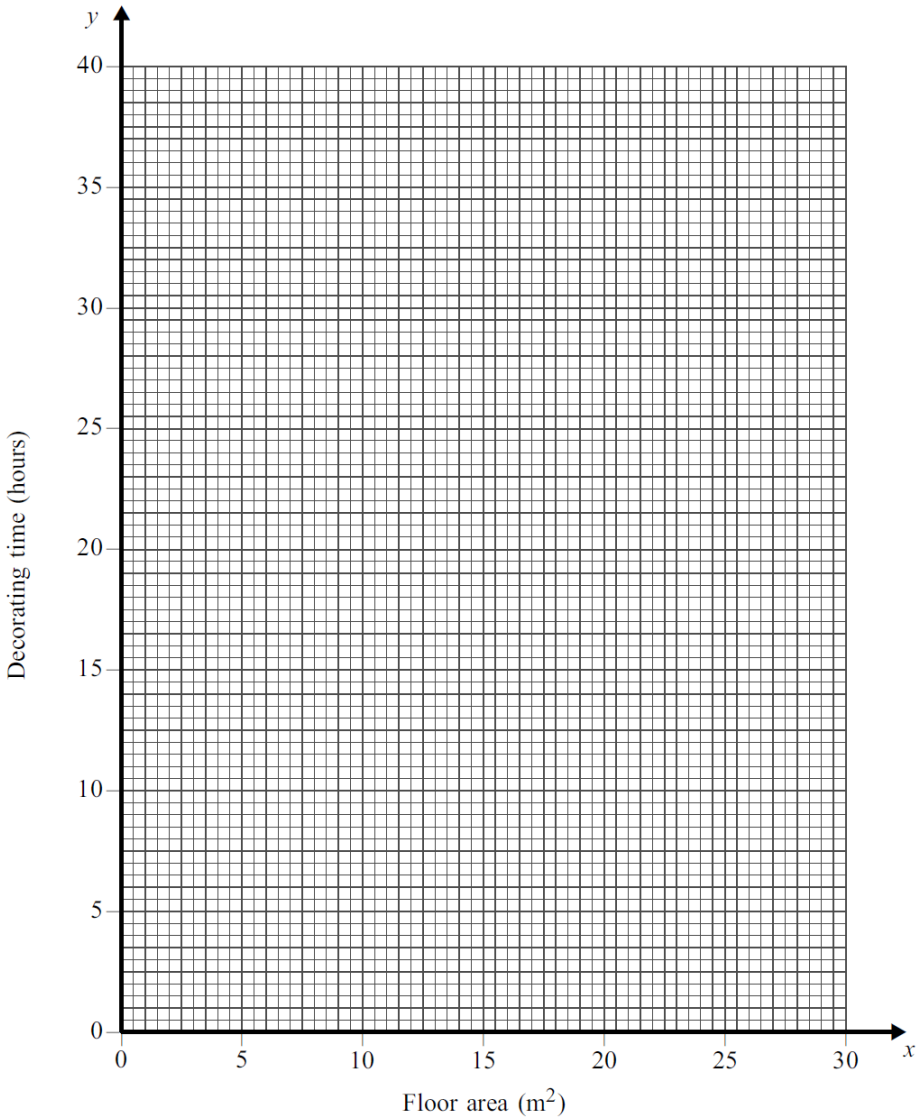
(a) On **Figure 1**, plot a scatter diagram of these data. *(2 marks)*

(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 $15\,m^2$. *(2 marks)*

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



Regression - exam questions - MS

Question 1: Jan 2007

7(a)	8 or 7 points plotted accurately (6 or 5 points plotted accurately)	B2 (B1)	2	
(b)	Gradient, $b = 0.114$ to 0.115 ($b = 0.11$ to 0.12)	B2 (B1)		AWFW (0.11469)
	Intercept, $a = 15.9$ to 16.1 ($a = 13$ to 19)	B2 (B1)		AWFW (16.00824)
	Attempt at $\sum x$, $\sum x^2$, $\sum y$ and $\sum xy$ or Attempt at S_{xx} and S_{xy}	(M1)		4420, 3230800, 635 and 441300 788750 and 90462.5
	Attempt at correct formula for b $b = 0.114$ to 0.115 $a = 15.9$ to 16.1	(m1) (A1) (A1)		 AWFW AWFW
	Accept a and b interchanged only if then identified correctly later in question			
	Line plotted accurately (Evidence of correct method for ≥ 2 points)	B2 (M1)	6	At least from $x = 200$ to 1000
(c)	$\text{Res}_H = y_H - Y_H = 70 - (a + b \times 480)$	M1		Used; or implied by **correct** answer; allow for $Y_H - y_H$ **shown**
	$= -1.5$ to -0.5	A1		AWFW (-1.06)
	Point H is (almost) **on** / **just below** the line	B1	3	Accept near / close / just above or equivalent
(d)	$Y = a + b \times 560$ or reading from scatter diagram	M1		Used
	$= 79$ to 81	A1		AWFW (80.2)
	Cost $= Y \times \frac{12}{60}$ or $\frac{Y}{5}$	M1		Used
	$=$ £15.8 to £16.2	A1	4	AWFW; ignore units (£16.05)
	Total		**15**	

Question 2: Jun 2006

3(a)(i)	Gradient, $b = -3.24$ to -3.26	B2		AWFW (–3.25)
	$b = -3.2$ to -3.3	B1		AWFW
	Intercept, $a = 262$ to 264	B2		AWFW $(262.8\dot{8})$
	$a = 260$ to 270	B1		AWFW
	Attempt at Σx Σx^2 Σy Σxy or Attempt at S_{xx} S_{xy}	M1		108, 1836, 2015, 22425 or 540, –1755
	Attempt at a correct formula for b	m1		
	$b = -3.24$ to -3.26	A1		AWRT
	$a = 262$ to 264	A1	4	AWFW
	Accept a & b interchanged only if identified correctly in (b) and (c)			
(ii)	Gradient, b:			
	Decrease in pressure per month	B2		or equivalent
	Change in pressure	B1	2	or better
(iii)	Intercept, a:			
	Initial pressure or pressure at $x = 0$	B1		or equivalent; not y-intercept
	Reference to 265, actual or expected value	B1	2	
(b)(i)	Value for $b = 2 \times$ [gradient or b from (a)(i)]	M1		accept $2b$; ignore sign
	$= -6.4$ to -6.6	A1√	2	AWFW (–6.5) √ from (a)(i) but must be < 0
(ii)	$P_8 = 265 - 6.5 \times 8$	M1		must use 265 and $x = 8$ and $2 \times [b\ (< 0)$ from (a)(i)] or [from (b)(i) (< 0)]
	$= 212$ to 214 (< 220)	A1	2	AWFW AG
	Total		**12**	

Question 3: Jan 11

5(a)	Time taken is dependent upon leaving time	B1	1	Or equivalent
(b)	b (gradient) = 1.28 (or 141/110) b (gradient) = 1.25 to 1.35	B2 (B1)		AWRT; (CAO or equivalent) (1.28182) AWFW Treat rounding of correct answers as ISW
	a (intercept) = 29.95 to 30 (or 659/22) a (intercept) = 29 to 31 Thus $y = 30 + 1.28x$	B2 (B1) B1F	5	AWFW; (CAO or equivalent) (29.95455) AWFW F on a and b
	or Attempt at Σx Σx^2 Σy and Σxy $\left(\Sigma y^2\right)$			275 9625 682 and 20575 (47494) (All four attempted)
	or Attempt at S_{xx} and S_{xy} $\left(S_{yy}\right)$	(M1)		2750 and 3525 (5210) (Both attempted)
	Attempt at correct formula for b gradient b (gradient) = 1.28 (or 141/110) a (intercept) = 29.95 to 30 (or 659/22)	(m1) (A1) (A1)		 AWRT; (CAO or equivalent) AWFW; (CAO or equivalent)
	Thus $y = 30 + 1.28x$	(B1F)		F on a and b
	Accept a and b interchanged only if identified correctly by a clearly shown equation			If a and b are not identified anywhere in the question, then: 1.25 to 1.35 $\Rightarrow$ B1 29 to 30 $\Rightarrow$ B1
(c)	7.45 am $\Rightarrow x = 15$ $\Rightarrow y_{15} = 30 + 1.28 \times 15$ = 47 to 52	B1 M1 A1		CAO; stated, used or implied Use of $10 < x < 20$ AWFW (49.2)
	Time before 9.00 am = 9.00 – (7.45 + c's y_{15}) = 23 to 28	M1 A1	 5	May be implied AWFW (25.8)
	SC Answer of 17 CAO (use of c's y_{15} = 58) gains 2 marks			**NB:** An answer of 8.32 to 8.37 gains B1 M1 A1 M0 A0
d)(i)	$y_{85} = 30 + 1.28 \times 85 = 135$ to 146	B1	1	AWFW (138.9)
(ii)	Extrapolation/ outside/ above range of x-values	B1		Or equivalent
	Implies leaves home at 8.55 so different traffic conditions	B1	2	Or equivalent; 8.55 may be implied by 5 minutes
	Total		**14**	

Question 4: Jun 2007

5(a)	Time taken depends upon temperature	B1	1	OE; **not** x set values
(b)	b (gradient) = –0.0873 to –0.087	B2		AWFW $\left(-0.087\dot{2}\dot{7}\right)$
	b (gradient) = –0.09 to –0.08	(B1)		AWFW; $-8.73^{-02} \Rightarrow$ B0
	a (intercept) = 5.94 to 5.96	B2		AWFW $\left(5.950\dot{9}\right)$
	a (intercept) = 5.6 to 6.1	(B1)		AWFW
	Attempt at $\sum x$, $\sum x^2$, $\sum y$ and $\sum xy$			396, 16016, 30.9 and 958.8
	or	(M1)		
	Attempt at S_{xx} and S_{xy}			1760 and –153.6
	Attempt at correct formula for b	(m1)		
	b = –0.0873 to –0.087	(A1)		AWFW
	a = 5.94 to 5.96	(A1)	4	AWFW
	Accept a and b interchanged only if then identified correctly later in question			
c)(i)	Each 1 °C rise in temperature results in an	B1		Quantified rise in x (results in)
	(average) decrease of 0.087 m (5 s)	B1	2	Decrease in y
	in time taken for pellets to dissolve			OE
(ii)	a is y–value at $x = 0$ at which water is	B1		Indication that it is y at $x = 0$
	solid/ice/frozen so pellets cannot dissolve	B1	2	Mention of solid or ice or frozen
d)(i)	When x = 30			
	y = 3.3 to 3.4	B2		AWFW $\left(3.33\dot{2}\dot{7}\right)$
	y = 2.9 to 3.7	(B1)		AWFW
	If B0, use of their equation with x = 30	(M1)	2	
(ii)	When x = 75			
	$y < 0$ or negative	B1		OE
	which	↑Dep↑		
	is impossible	B1	2	OE; **not** extrapolation
	Total		**13**	

Question 5: Jun 2008

1(a)	b (gradient) = –1.01 to –1(.00) (b (gradient) = –1.05 to –0.95)	B2 (B1)		AWFW	(–1.00337)
	a (intercept) = 53(.0) to 53.2 (a (intercept) = 52(.0) to 54(.0))	B2 (B1)	4	AWFW	(53.06736)
	OR				
	Attempt at $\sum x$, $\sum x^2$, $\sum y$ and $\sum xy$ **or** Attempt at S_{xx} and S_{xy}	(M1)		180, 3986, 297 and 5552.7 386 and –387.3	
	Attempt at **correct** formula for b (gradient)	(m1)			
	b (gradient) = –1.01 to –1(.00) a (intercept) = 53(.0) to 53.2	(A1) (A1)		AWFW AWFW	
	Accept a and b interchanged only if then identified correctly in part (b), but B2 in (b) does **not** necessarily imply 4 marks in (a)				
(b)	When x = 21,				
	y = 31.7 to 32.2 (y = 29.9 to 34.1)	B2 (B1)	2	AWFW AWFW	(32.0)
	Evidence of use of 21 in c's equation	(M1)			
	Special Cases (if seen):				
	$y = \frac{33.0+30.7}{2}$ = 31.8 to 31.9	(B1)		AWFW; or equivalent	
	y = 31.85 without working	(B1)			
	Total		**6**		

Question 6: Jan 2008

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**	