HIGHER

# 4.5 Regression and the PMCC: 1-hour Pack

EDEXCEL 1ST0 · Calculator allowed · about 60 minutes

Total Marks

Name: _____ Date: ____ / ____ / _____

**Answer ALL questions. Show all your working.**

**1** The data below could be plotted on a scatter graph. A dealer records the age in years of a car (x) and its value in hundreds of pounds (y).

| x | 1 | 2 | 5 | 6 | 12 | 16 |
|---|---|---|---|---|---|---|
| y | 101 | 98 | 80 | 79 | 55 | 44 |

A line of best fit is drawn so that it passes through the points (2, 98) and (12, 55).

(a) Work out the gradient of the line.

(b) Find the equation of the line in the form $y = mx + c$.

(c) Use your equation to estimate y when $x = 24$.

(d) State whether your estimate in (c) is interpolation or extrapolation, and whether it is likely to be reliable.

**(Total for Question 1 is 7 marks)**


© Revision Library | revisionlibrary.co.uk

**2** For 7 pairs of data (x, y): sum of x = 49, sum of y = 122, sum of xy = 626, sum of $x^2$ = 451, sum of $y^2$ = 2784.

(a) Work out $S_{xy}$ = (sum of xy) - (sum of x)(sum of y)/n and $S_{xx}$ = (sum of $x^2$) - (sum of x)$^2$/n.

(b) Work out the gradient b = $S_{xy}$ / $S_{xx}$ of the regression line y = a + bx, to 3 significant figures.

(c) The mean of x is 7 and the mean of y is 17.43. Find a, then use the line to estimate y when x = 6. Give each answer to 3 significant figures, using unrounded values in your working.

**(Total for Question 2 is 6 marks)**

---

**3** The data below could be plotted on a scatter graph. A taxi firm records distance travelled in kilometres (x) and the fare in pounds (y).

| x | 7 | 8 | 10 | 11 | 13 | 16 |
|---|---|---|---|---|---|---|
| y | 53 | 56 | 65 | 63 | 63 | 74 |

A line of best fit is drawn so that it passes through the points (8, 56) and (13, 63).

(a) Work out the gradient of the line.

(b) Find the equation of the line in the form y = mx + c.

**(Total for Question 3 is 4 marks)**


© Revision Library | revisionlibrary.co.uk

**4** The data below could be plotted on a scatter graph. A teacher records hours spent revising (x) and test mark (y).

| x | 1 | 2 | 6 | 8 | 9 | 13 | 17 |
|---|---|---|---|---|---|---|---|
| y | 51 | 54 | 69 | 73 | 74 | 84 | 102 |

A line of best fit is drawn so that it passes through the points (2, 54) and (13, 84).

(a) Work out the gradient of the line.

(b) Find the equation of the line in the form $y = mx + c$.

**(Total for Question 4 is 4 marks)**

---

**5** For 7 pairs of data (x, y): sum of x = 60, sum of y = 316, sum of xy = 3157, sum of $x^2$ = 620, sum of $y^2$ = 16790.

(a) Work out $S_{xy}$ = (sum of xy) - (sum of x)(sum of y)/n and $S_{xx}$ = (sum of $x^2$) - (sum of x)$^2$/n.

(b) Work out the gradient $b = S_{xy} / S_{xx}$ of the regression line $y = a + bx$, to 3 significant figures.

(c) The mean of x is 8.57 and the mean of y is 45.14. Find a, then use the line to estimate y when x = 7. Give each answer to 3 significant figures, using unrounded values in your working.

**(Total for Question 5 is 6 marks)**


© Revision Library | revisionlibrary.co.uk

**6** The data below could be plotted on a scatter graph. A dealer records the age in years of a car (x) and its value in hundreds of pounds (y).

| x | 2 | 7 | 8 | 14 | 16 | 17 | 20 |
|---|---|---|---|---|---|---|---|
| y | 81 | 72 | 65 | 51 | 44 | 46 | 41 |

A line of best fit is drawn so that it passes through the points (7, 72) and (17, 46).

(a) Work out the gradient of the line.

(b) Find the equation of the line in the form $y = mx + c$.

(c) Use your equation to estimate y when $x = 18$.

(d) State whether your estimate in (c) is interpolation or extrapolation, and whether it is likely to be reliable.

**(Total for Question 6 is 7 marks)**

---

**7** The data below could be plotted on a scatter graph. A dealer records the age in years of a car (x) and its value in hundreds of pounds (y).

| x | 3 | 8 | 9 | 13 | 14 | 16 | 19 |
|---|---|---|---|---|---|---|---|
| y | 24 | 47 | 44 | 67 | 64 | 77 | 83 |

A line of best fit is drawn so that it passes through the points (8, 47) and (16, 77).

(a) Work out the gradient of the line.

(b) Find the equation of the line in the form $y = mx + c$.

**(Total for Question 7 is 4 marks)**


© Revision Library | revisionlibrary.co.uk

**8** For 6 pairs of data (x, y): sum of x = 49, sum of y = 289, sum of xy = 2893, sum of $x^2$ = 549, sum of $y^2$ = 16219.

(a) Work out $S_{xy}$ = (sum of xy) - (sum of x)(sum of y)/n and $S_{xx}$ = (sum of $x^2$) - (sum of x)$^2$/n.

(b) Work out the gradient b = $S_{xy}$ / $S_{xx}$ of the regression line y = a + bx, to 3 significant figures.

(c) The mean of x is 8.17 and the mean of y is 48.17. Find a, then use the line to estimate y when x = 6. Give each answer to 3 significant figures, using unrounded values in your working.

**(Total for Question 8 is 6 marks)**

---

**9** The data below could be plotted on a scatter graph. A homeowner records the temperature in degrees Celsius (x) and gas used in units (y).

| x | 4 | 8 | 13 | 15 | 16 | 18 |
|---|---|---|---|---|---|---|
| y | 92 | 75 | 58 | 50 | 41 | 32 |

A line of best fit is drawn so that it passes through the points (8, 75) and (16, 41).

(a) Work out the gradient of the line.

(b) Find the equation of the line in the form y = mx + c.

**(Total for Question 9 is 4 marks)**


© Revision Library | revisionlibrary.co.uk

**10** The data below could be plotted on a scatter graph. A shop records the number of rainy days in a month (x) and umbrellas sold (y).

| x | 1 | 2 | 9 | 12 | 18 | 20 |
|---|---|---|---|---|---|---|
| y | 30 | 32 | 50 | 61 | 78 | 75 |

A line of best fit is drawn so that it passes through the points (2, 32) and (18, 78).

(a) Work out the gradient of the line.

(b) Find the equation of the line in the form $y = mx + c$.

**(Total for Question 10 is 4 marks)**

---

**11** For 5 pairs of data (x, y): sum of x = 41, sum of y = 248, sum of xy = 2335, sum of $x^2$ = 429, sum of $y^2$ = 13470.

(a) Work out $S_{xy}$ = (sum of xy) - (sum of x)(sum of y)/n and $S_{xx}$ = (sum of $x^2$) - (sum of x)$^2$/n.

(b) Work out the gradient $b = S_{xy} / S_{xx}$ of the regression line $y = a + bx$, to 3 significant figures.

(c) The mean of x is 8.2 and the mean of y is 49.6. Find a, then use the line to estimate y when x = 10. Give each answer to 3 significant figures, using unrounded values in your working.

**(Total for Question 11 is 6 marks)**


© Revision Library | revisionlibrary.co.uk

# Mark scheme · 4.5 Regression and the PMCC: 1-hour Pack

## Question 1

| | |
|---|---|
| change in y divided by change in x | **M1** |
| -4.3 cao | **A1** |
| substitutes a point to find c | **M1** |
| y = -4.3x + 106.6 cao | **A1** |
| substitutes x = 24 | **M1** |
| 3.4 cao | **A1** |
| extrapolation and less reliable as outside the data range | **B1** |

**Answer: (a) -4.3; (b) y = -4.3x + 106.6; (c) 3.4; (d) extrapolation**

## Question 2

| | |
|---|---|
| $S_{xy}$ or $S_{xx}$ formula used | **M1** |
| $S_{xy}$ = -228 and $S_{xx}$ = 108 | **A1** |
| b = -2.11 | **A1** |
| a = mean of y minus b x mean of x | **M1** |
| a = 32.2 | **A1** |
| 19.5 cao | **A1** |

**Answer: (a) -228, 108; (b) -2.11; (c) a = 32.2, y = 19.5**

## Question 3

| | |
|---|---|
| change in y divided by change in x | **M1** |
| gradient 1.4 cao | **A1** |
| substitutes a point into y = mx + c | **M1** |
| y = 1.4x + 44.8 cao | **A1** |

**Answer: (a) 1.4; (b) y = 1.4x + 44.8**

## Question 4

| | |
|---|---|
| change in y divided by change in x | **M1** |
| gradient 2.73 cao | **A1** |
| substitutes a point into y = mx + c | **M1** |
| y = 2.73x + 48.55 cao | **A1** |

**Answer: (a) 2.73; (b) y = 2.73x + 48.55**

## Question 5

| | |
|---|---|
| $S_{xy}$ or $S_{xx}$ formula used | **M1** |
| $S_{xy}$ = 448.43 and $S_{xx}$ = 105.71 | **A1** |
| b = 4.24 | **A1** |
| a = mean of y minus b x mean of x | **M1** |
| a = 8.78 | **A1** |
| 38.5 cao | **A1** |

**Answer: (a) 448.43, 105.71; (b) 4.24; (c) a = 8.78, y = 38.5**

## Question 6

| | |
|---|---|
| change in y divided by change in x | **M1** |
| -2.6 cao | **A1** |
| substitutes a point to find c | **M1** |
| y = -2.6x + 90.2 cao | **A1** |
| substitutes x = 18 | **M1** |
| 43.4 cao | **A1** |
| interpolation and more reliable as within the data range | **B1** |

**Answer: (a) -2.6; (b) y = -2.6x + 90.2; (c) 43.4; (d) interpolation**

## Question 7

| | |
|---|---|
| change in y divided by change in x | **M1** |
| gradient 3.75 cao | **A1** |
| substitutes a point into y = mx + c | **M1** |
| y = 3.75x + 17 cao | **A1** |

**Answer: (a) 3.75; (b) y = 3.75x + 17**

## Question 8

| | |
|---|---|
| $S_{xy}$ or $S_{xx}$ formula used | **M1** |
| $S_{xy}$ = 532.83 and $S_{xx}$ = 148.83 | **A1** |
| b = 3.58 | **A1** |
| a = mean of y minus b x mean of x | **M1** |
| a = 18.9 | **A1** |
| 40.4 cao | **A1** |

**Answer: (a) 532.83, 148.83; (b) 3.58; (c) a = 18.9, y = 40.4**

## Question 9

| | |
|---|---|
| change in y divided by change in x | **M1** |
| gradient -4.25 cao | **A1** |
| substitutes a point into y = mx + c | **M1** |
| y = -4.25x + 109 cao | **A1** |

**Answer: (a) -4.25; (b) y = -4.25x + 109**

## Question 10

| | |
|---|---|
| change in y divided by change in x | **M1** |
| gradient 2.875 oe (awrt 2.88) | **A1** |
| substitutes a point into y = mx + c | **M1** |
| y = 2.875x + 26.25 oe (awrt y = 2.88x + 26.25) | **A1** |

**Answer: (a) 2.88; (b) y = 2.88x + 26.25**

## Question 11

| | |
|---|---|
| $S_{xy}$ or $S_{xx}$ formula used | **M1** |
| $S_{xy}$ = 301.4 and $S_{xx}$ = 92.8 | **A1** |
| b = 3.25 | **A1** |
| a = mean of y minus b x mean of x | **M1** |
| a = 23.0 | **A1** |
| 55.4 cao | **A1** |

**Answer: (a) 301.4, 92.8; (b) 3.25; (c) a = 23.0, y = 55.4**


© Revision Library | revisionlibrary.co.uk