More 2 variable statistics IB questions 2014

1. Oral tests are conducted by three examiners A, B and C separately. The results of the examination are classified as Credit, Pass or Fail. A χ^2 test is applied to the data collected in order to test whether or not the examiners differ in their standard of awards.

(a) State the null hypothesis, H_0, for this data.

(b) Write down the number of degrees of freedom.

Of the 135 students who sit the exam, 30 get Credit and 45 are tested by examiner A.

(c) Calculate the expected number of students who get a Credit and are tested by examiner A.

Using a 5% level of significance, the p-value is found to be 0.0327 correct to 3 s.f.

(d) State whether H_0 should be accepted. Justify your answer.

Working:

Answers:

(a) .....

(b) .....

(c) .....

(d) .....

(Total 6 marks)

2. Several candy bars were purchased and the following table shows the weight and the cost of each bar.

	Yummy	Chox	Marz	Twin	Chunx	Lite	BigC	Bite
Weight (g)	60	85	80	65	95	50	100	45
Cost (Euros)	1.10	1.50	1.40	1.20	1.80	1.00	1.70	0.90

(a) Find the correlation coefficient, r, giving your answer correct to 3 decimal places.

(1)

(b) Describe the correlation between the weight of a candy bar and its cost.

(1)

(c) Calculate the equation of the regression line for y on x.

(2)

(d) Use your equation to estimate the cost of a candy bar weighing 109 g (2)

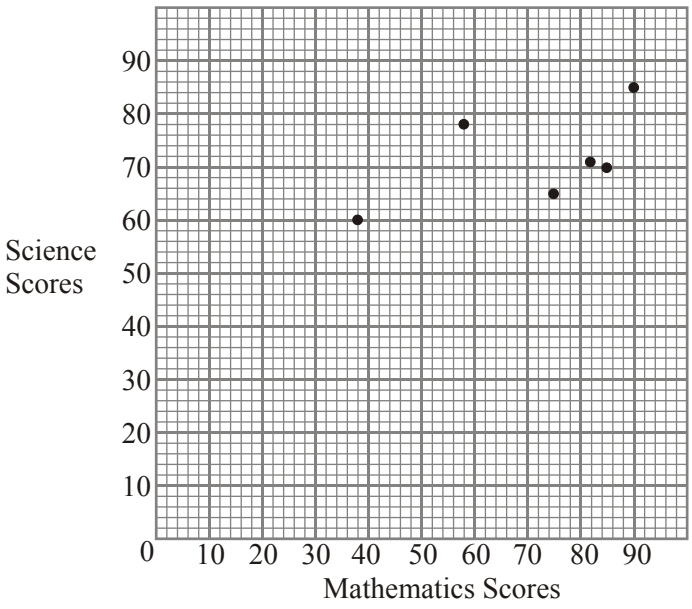
(e) State whether this estimate reliable or not, and give a reason to support your answer

(2)
(Total 8 marks)

3. The following are the results of a survey of the scores of 10 people on both a mathematics (x) and a science (y) aptitude test:

Student	**Mathematics (x)**	**Science (y)**	
1	90	85	
2	38	60	
3	58	78	$\bar{x} = 73$
4	85	70	$\bar{y} = 78$
5	73	65	
6	82	71	
7	56	80	
8	73	90	
9	95	96	
10	80	85	

(a) Copy the graph below on graph paper and fill in the missing points for students 7–10 on the graph.



(4)

(b) Plot the point M ($\bar{x}$, $\bar{y}$) on the graph.

(1)

(c) Find the equation of the regression line of y on x in the form $y = ax + b$.

(2)

(d) Graph this line on the above graph.

(2)

(e) Given that a student receives an 88 on the mathematics test, what would you expect this student's science score to be? Show how you arrived at your result.

(2)

(Total 11 marks)

4. A survey was conducted in a company to determine whether position in upper management was independent of gender. The results of this survey are tabulated below.

	Managers	Junior executives	Senior executives	**Totals**
Male	95	130	75	**300**
Female	65	110	25	**200**
Totals	**160**	**240**	**100**	**500**

The table below shows the expected number of males and females at each level, if they were represented proportionally to the total numbers of males and females employed.

	Managers	Junior executives	Senior executives	**Totals**
Male	*a*	*c*	60	**300**
Female	*b*	*d*	40	**200**
Totals	**160**	**240**	**100**	**500**

(a) (i) Show that the expected number of **Male Managers** (*a*) is 96.

(ii) Hence find the values of *b, c* and *d*.

(5)

(b) (i) Write a suitable null hypothesis for this data.

(ii) Write a suitable alternate hypothesis for this data.

(2)

(c) (i) Perform a chi-squared test of independence for this data to show the value of χ^2 is 12.8 to 3 significant figures.

(ii) Calculate the number of degrees of freedom,

The critical value of χ^2 at the 5% significance level is 5.991.

(iii) What conclusion can be drawn regarding gender and position in upper management?

(6)
(Total 13 marks)

More 2 variable statistics IB answers

1. (a) H_0 = The standard of award is independent of the examiner (or equivalent) (A1) (C1)

(b) 4 (A1) (C1)

(c) $f_e = \frac{30 \times 45}{135}$ (M1)

$f_e = 10$ (A1) (C2)

(d) No, because the p-value is less than the significance level. (A2)

OR

No, because 0.0327 < 0.05 (A2) (C2)

[6]

2. (a) $r = \frac{S_{xy}}{(S_x S_y)} = \frac{5.81}{(19.2 \times 0.307)}$ (M1)

= 0.986 (A1) 2

Note: *Award (G2) for 0.985 from GDC.*

(b) Strong, positive correlation (A1) 1

(c) $y = 0.182 + 0.0158x$ (G3)

OR

$y - 1.325 = \frac{5.81}{19.2^2}(x - 72.5)$ (M1)(A1)

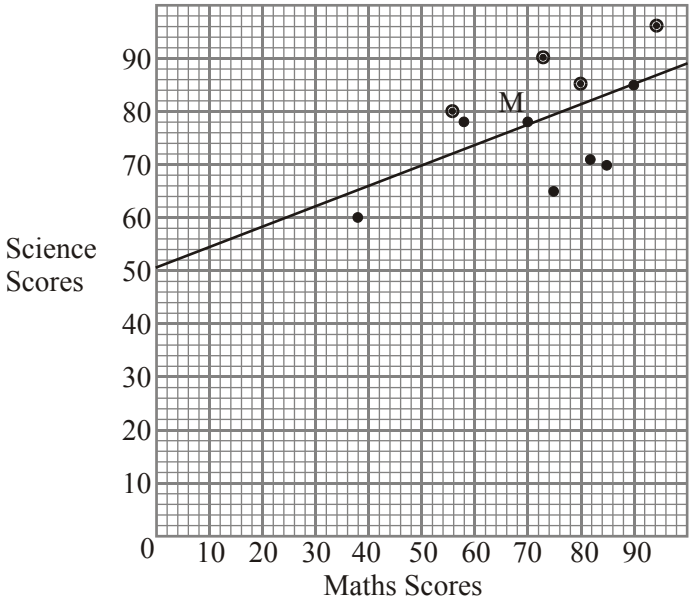
$y = 0.0158x + 0.182$ (A1) 3

(d) $y = 0.0158 \times 109 + 0.182$ (M1)

= 1.90 euros (A1)2

[8]

3. (a)



(A4) 4

Note: *Award (A1) for each circled point plotted correctly.*

(b) Point M plotted correctly. (A1) 1

(c) $y = 0.359x + 51.8$ (A1)(A1) 2

(d) For reasonable line of best fit. (A2) 2

Note: *Award (A1) for going through M, (A1) for y intercept anywhere from 50 to 54 or* ***ft*** *from (c). Extrapolate line to y-axis is necessary.*

(e) $y = 0.359 \times 88 + 51.8$ (M1)
$y = 83$ (A1)

OR

$y = 83$ (±2) if read from the graph and method is shown. (A2) 2

[11]

4. (a) (i) Expected number of male managers

$= \frac{160}{500} \times \frac{300}{500} \times 500 = \frac{160 \times 300}{500}$ (M1)(A1)

$= 96$ (AG)

(ii) $b = 160 - 96 = 64$ (A1)

$c = 300 - 96 - 60 = 144$ (A1)

$d = 240 - 144 = 96$ (A1) 5

(b) (i) H_0: Position is independent of gender (A1)

(ii) H_1: Position is dependent on gender (A1) 2

(c) (i) $\chi^2_{\text{calc}} = \frac{(96-95)^2}{96} + \frac{(64-65)^2}{64} + \frac{(114-130)^2}{144} + \frac{(96-110)^2}{96}$

$+ \frac{(60-75)^2}{60} + \frac{(40-25)^2}{40}$ (M1)(A2)

Note: *Award (M1) for using* $\sum \frac{(f_e - f_0)^2}{f_e}$*, (A2) for all values correct.*

Special case: Award (A1) if only 1 value is incorrect.

$= 12.8$ (AG)

(ii) 2 degrees of freedom (A1)

$\chi^2_{2,0.05} = 5.991$ (A1)

(iii) Any of: (then reject H_0, accept H_1) (R1) 6

Position is dependent on gender.

[13]