**Pearson Edexcel Level 3 Certificate in Mathematics in Context**

**Practice questions:**

**Statistics**

**Topic practice questions**

These materials have been gathered together to help provide opportunities for skills practice on some of the maths topics within the content of the Mathematics in Context specification. The materials comprise four sets of questions organised by topic area as follows:

**A Statistics**

B Probability and Venn Diagrams

C Linear Programming

D Sequences

The majority of the questions have been taken from past exam papers in GCE Statistics 1. Some questions in the Series and sequences strand have been written afresh for this purpose to broaden the range of topics covered.

None of the questions is intentionally written in the style of Mathematics in Context exam questions. You and your students may however find them useful for classroom discussion, group work and/or individual practice on some of the maths skills within the specification.

**A. STATISTICS**

**1.** The table shows data on the number of visitors to the UK in a month, *v* (1000s), and the amount of money they spent, *m* (£ millions), for each of 8 months.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Number of visitors  *v* (1000s) | 2450 | 2480 | 2540 | 2420 | 2350 | 2290 | 2400 | 2460 |
| Amount of money spent  *m* (£ millions) | 1370 | 1350 | 1400 | 1330 | 1270 | 1210 | 1330 | 1350 |

You may use

*Svv* = 42587.5 *Svm* = 31512.5 *Smm* = 25187.5 Σ*v =* 19390 Σ*m* = 10610

(*a*) Find the product moment correlation coefficient between *m* and *v*.

**(2)**

(*b*) Give a reason to support fitting a regression model of the form *m* = *a* + *bv* to these data.

**(1)**

(*c*) Find the value of *b* correct to 3 decimal places.

**(2)**

(*d*) Find the equation of the regression line of *m* on *v*.

**(2)**

(*e*) Interpret your value of *b*.

**(2)**

(*f*) Use your answer to part (*d*) to estimate the amount of money spent when the number of visitors to the UK in a month is 2 500 000.

**(2)**

(*g*) Comment on the reliability of your estimate in part (*f*). Give a reason for your answer.

**(2)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**2.** A large company is analysing how much money it spends on paper in its offices every year. The number of employees, *x*, and the amount of money spent on paper, *p*(£ hundreds), in 8 randomly selected offices are given in the table below.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *x* | 8 | 9 | 12 | 14 | 7 | 3 | 16 | 19 |
| *p* (£ hundreds) | 40.5 | 36.1 | 30.4 | 39.4 | 32.6 | 31.1 | 43.4 | 45.7 |

(You may use Σ*x*2 = 1160 Σ*p* = 299.2 Σ*p*2 = 11 422 Σ*xp* = 3449.5)

(*a*) Show that *Spp* = 231.92 and find the value of *Sxx* and the value of *Sxp*.

**(5)**

(*b*) Calculate the product moment correlation coefficient between *x* and *p*.

**(2)**

The equation of the regression line of *p* on *x* is given in the form *p* = *a + bx*.

(*c*) Show that, to 3 significant figures, *b* = 0.824 and find the value of *a*.

**(4)**

(*d*) Estimate the amount of money spent on paper in an office with 10 employees.

**(2)**

(*e*) Explain the effect each additional employee has on the amount of money spent on paper.

**(1)**

Later the company realised it had made a mistake in adding up its costs, *p*. The true costs were actually half of the values recorded. The product moment correlation coefficient and the equation of the linear regression line are recalculated using this information.

(*f*) Write down the new value of

(i) the product moment correlation coefficient,

(ii) the gradient of the regression line.

**(2)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**3.** A meteorologist believes that there is a relationship between the height above sea level, *h* m, and the air temperature, *t* °C. Data is collected at the same time from 9 different places on the same mountain. The data is summarised in the table below.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *h* | 1400 | 1100 | 260 | 840 | 900 | 550 | 1230 | 100 | 770 |
| *t* | 3 | 10 | 20 | 9 | 10 | 13 | 5 | 24 | 16 |

[You may assume that

∑*h* = 7150, ∑*t* =110, ∑*h*2 = 7171500, ∑*t*2 = 1716, ∑*th* = 64 980 and S*tt* = 371.56]

(*a*) Calculate S*th* and S*hh*. Give your answers to 3 significant figures.

**(3)**

(*b*) Calculate the product moment correlation coefficient for this data.

**(2)**

(*c*) State whether or not your value supports the use of a regression equation to predict the air temperature at different heights on this mountain. Give a reason for your answer.

**(1)**

(*d*) Find the equation of the regression line of *t* on *h* giving your answer in the form *t*= *a*+ *bh*.

**(4)**

(*e*) Interpret the value of *b*.

**(1)**

(*f*) Estimate the difference in air temperature between a height of 500 m and a height of 1000 m.

**(2)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**4.** A teacher asked a random sample of 10 students to record the number of hours of television, *t*, they watched in the week before their mock exam. She then calculated their grade, *g*, in their mock exam. The results are summarised as follows.

∑ *t* = 258 ∑ *t*2 = 8702 ∑ *g* = 63.6 S*gg* = 7.864 ∑ *gt* = 1550.2

(*a*) Find S*tt* and S*gt* .

**(3)**

(*b*) Calculate, to 3 significant figures, the product moment correlation coefficient between *t*and *g*.

**(2)**

The teacher also recorded the number of hours of revision, *v*, these 10 students completed during the week before their mock exam. The correlation coefficient between *t* and *v* was –0.753.

(*c*) Describe, giving a reason, the nature of the correlation you would expect to find between *v*and *g*.

**(2)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**5.** A biologist is comparing the intervals (*m* seconds) between the mating calls of a certain species of tree frog and the surrounding temperature (*t* °C). The following results were obtained.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *t* °C | 8 | 13 | 14 | 15 | 15 | 20 | 25 | 30 |
| *m* secs | 6.5 | 4.5 | 6 | 5 | 4 | 3 | 2 | 1 |

(You may use ∑ *tm* = 469.5, S*tt* = 354, S*mm* = 25.5)

(*a*) Show that S*tm* = –90.5.

**(4)**

(*b*) Find the equation of the regression line of *m* on *t* giving your answer in the form *m*= *a + bt*.

**(4)**

(*c*) Use your regression line to estimate the time interval between mating calls when the surrounding temperature is 10 °C.

**(1)**

(*d*) Comment on the reliability of this estimate, giving a reason for your answer.

**(1)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**6.** A scientist is researching whether or not birds of prey exposed to pollutants lay eggs with thinner shells. He collects a random sample of egg shells from each of 6 different nests and tests for pollutant level, *p*, and measures the thinning of the shell, *t*. The results are shown in the table below.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *p* | 3 | 8 | 30 | 25 | 15 | 12 |
| *t* | 1 | 3 | 9 | 10 | 5 | 6 |

[You may use ∑*p*2 = 1967 and ∑*pt* = 694]

(*a*) On graph paper, draw a scatter diagram to represent these data.

**(2)**

(*b*) Explain why a linear regression model may be appropriate to describe the relationship between *p* and *t*.

**(1)**

(*c*) Calculate the value of *Spt* and the value of *Spp* .

**(4)**

(*d*) Find the equation of the regression line of *t* on *p*, giving your answer in the form *t = a + bp*.

**(4)**

(*e*) Plot the point () and draw the regression line on your scatter diagram.

**(2)**

The scientist reviews similar studies and finds that pollutant levels above 16 are likely to result in the death of a chick soon after hatching.

(*f* ) Estimate the minimum thinning of the shell that is likely to result in the death of a chick.

**(2)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**7.** The age, *t* years, and weight, *w* grams, of each of 10 coins were recorded. These data are summarised below.

∑ *t*2 = 2688 ∑ *tw* = 1760.62 ∑ *t* = 158 ∑*w* = 111.75 *Sww* = 0.16

(*a*) Find *Stt* and *Stw* for these data.

**(3)**

(*b*) Calculate, to 3 significant figures, the product moment correlation coefficient between *t* and *w*.

**(2)**

(*c*) Find the equation of the regression line of *w* on *t* in the form *w* = *a* + *bt*.

**(4)**

(*d*) State, with a reason, which variable is the explanatory variable.

**(2)**

(*e*) Using this model, estimate

(i) the weight of a coin which is 5 years old,

(ii) the effect of an increase of 4 years in age on the weight of a coin.

**(2)**

It was discovered that a coin in the original sample, which was 5 years old and weighed 20 grams, was a fake.

(*f*) State, without any further calculations, whether the exclusion of this coin would increase or decrease the value of the product moment correlation coefficient. Give a reason for your answer.

**(2)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**8.** A teacher took a random sample of 8 children from a class. For each child the teacher recorded the length of their left foot, *f* cm, and their height, *h* cm. The results are given in the table below.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *f* | 23 | 26 | 23 | 22 | 27 | 24 | 20 | 21 |
| *h* | 135 | 144 | 134 | 136 | 140 | 134 | 130 | 132 |

(You may use ∑ *f* =186 ∑*h* =1085 S*ff* = 39.5 S*hh* =139.875 ∑ *fh* = 25 291)

(*a*) Calculate S*fh*.

**(2)**

(*b*) Find the equation of the regression line of *h* on f in the form *h* = *a* + *bf*.

Give the value of *a* and the value of *b* correct to 3 significant figures.

**(5)**

(*c*) Use your equation to estimate the height of a child with a left foot length of 25 cm.

**(2)**

(*d*) Comment on the reliability of your estimate in part (*c*), giving a reason for your answer.

**(2)**

The left foot length of the teacher is 25 cm.

(*e*) Give a reason why the equation in part (*b*) should not be used to estimate the teacher’s height.

**(1)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**9.** A travel agent sells flights to different destinations from *Beerow* airport. The distance *d*, measured in 100 km, of the destination from the airport and the fare £*f* are recorded for a random sample of 6 destinations.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Destination | *A* | *B* | *C* | *D* | *E* | *F* |
| *d* | 2.2 | 4.0 | 6.0 | 2.5 | 8.0 | 5.0 |
| *f* | 18 | 20 | 25 | 23 | 32 | 28 |

[You may use ∑ *d* 2 = 152.09 ∑ *f* 2 = 3686 ∑ *fd* = 723.1]

(*a*) On graph paper, draw a scatter diagram to illustrate this information.

**(2)**

(*b*) Explain why a linear regression model may be appropriate to describe the relationship between *f* and *d*.

**(1)**

(*c*) Calculate *Sdd* and *Sfd*.

**(4)**

(*d*) Calculate the equation of the regression line of *f*on *d* giving your answer in the form *f*= *a*+ *bd*.

**(4)**

(*e*) Give an interpretation of the value of *b*.

**(1)**

Jane is planning her holiday and wishes to fly from *Beerow* airport to a destination *t* km away. A rival travel agent charges 5p per km.

(*f*) Find the range of values of *t* for which the first travel agent is cheaper than the rival.

**(2)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**10.** The volume of a sample of gas is kept constant. The gas is heated and the pressure, *p*, is measured at 10 different temperatures, *t*. The results are summarised below.

Σ *p* = 445 Σ *p*2 = 38 125 Σ *t* = 240 Σ *t* 2 = 27 520 Σ *pt* = 26 830

(*a*) Find S*pp* and S*pt* .

**(3)**

Given that S*tt* = 21 760,

(*b*) calculate the product moment correlation coefficient.

**(2)**

(*c*) Give an interpretation of your answer to part (*b*).

**(1)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**11.** The weight, *w* grams, and the length, *l* mm, of 10 randomly selected newborn turtles are given in the table below.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *l* | 49.0 | 52.0 | 53.0 | 54.5 | 54.1 | 53.4 | 50.0 | 51.6 | 49.5 | 51.2 |
| *w* | 29 | 32 | 34 | 39 | 38 | 35 | 30 | 31 | 29 | 30 |

(You may use S*ll* = 33.381 S*wl* = 59.99 S*ww* = 120.1)

(*a*) Find the equation of the regression line of *w* on *l* in the form *w* = *a* + *bl*.

**(5)**

(*b*) Use your regression line to estimate the weight of a newborn turtle of length 60 mm.

**(2)**

(*c*) Comment on the reliability of your estimate giving a reason for your answer.

**(2)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**12.** A teacher is monitoring the progress of students using a computer based revision course. The improvement in performance, *y* marks, is recorded for each student along with the time, *x*hours, that the student spent using the revision course. The results for a random sample of 10 students are recorded below.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *x* hours | 1.0 | 3.5 | 4.0 | 1.5 | 1.3 | 0.5 | 1.8 | 2.5 | 2.3 | 3.0 |
| *y* marks | 5 | 30 | 27 | 10 | –3 | –5 | 7 | 15 | –10 | 20 |

[You may use ∑ *x* = 21.4, ∑ *y* = 96, ∑ *x*2 = 57.22, ∑ *xy* = 313.7 ]

(*a*) Calculate *Sxx* and *Sxy*.

**(3)**

(*b*) Find the equation of the least squares regression line of *y* on *x* in the form *y = a + bx*.

**(4)**

(*c*) Give an interpretation of the gradient of your regression line.

**(1)**

Rosemary spends 3.3 hours using the revision course.

(*d*) Predict her improvement in marks.

**(2)**

Lee spends 8 hours using the revision course claiming that this should give him an improvement in performance of over 60 marks.

(*e*) Comment on Lee’s claim.

**(1)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**