**1.** (*a*)State one disadvantage of using quota sampling compared with simple random sampling.

**(1)**

In a university 8% of students are members of the university dance club.

A random sample of 36 students is taken from the university.

The random variable *X* represents the number of these students who are members of the dance club.

(*b*)Using a suitable model for *X*, find

(i) P(*X* = 4)

(ii) P(*X* ≥ 7)

**(3)**

Only 40% of the university dance club members can dance the tango.

(*c*)Find the probability that a student is a member of the university dance club and can

dance the tango.

**(1)**

A random sample of 50 students is taken from the university.

(*d*)Find the probability that fewer than 3 of these students are members of the

university dance club and can dance the tango.

**(2)**

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

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

**2.** Marc took a random sample of 16 students from a school and for each student recorded

* the number of letters, *x*, in their last name
* the number of letters, *y*, in their first name

His results are shown in the scatter diagram on the next page.

(*a*)Describe the correlation between *x* and *y*.

**(1)**

Marc suggests that parents with long last names tend to give their children shorter first names.

(*b*)Using the scatter diagram comment on Marc’s suggestion, giving a reason for your

answer.

**(1)**

The results from Marc’s random sample of 16 observations are given in the table below.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *x* | 3 | 6 | 8 | 7 | 5 | 3 | 11 | 3 | 4 | 5 | 4 | 9 | 7 | 10 | 6 | 6 |
| *y* | 7 | 7 | 4 | 4 | 6 | 8 | 5 | 5 | 8 | 4 | 7 | 4 | 5 | 5 | 6 | 3 |

(*c*)Use your calculator to find the product moment correlation coefficient between

*x* and *y* for these data.

**(1)**

(*d*)Test whether or not there is evidence of a negative correlation between the number of letters in the last name and the number of letters in the first name.

You should

* state your hypotheses clearly
* use a 5% level of significance

**(3)**

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

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

****

**3.** Stav is studying the large data set for September 2015

He codes the variable Daily Mean Pressure, *x*, using the formula *y* = *x* − 1010

The data for all 30 days from Hurn are summarised by



(*a*)State the units of the variable *x*

**(1)**

(*b*)Find the mean Daily Mean Pressure for these 30 days.

**(2)**

(*c*)Find the standard deviation of Daily Mean Pressure for these 30 days.

**(3)**

Stav knows that, in the UK, winds circulate

* in a **clockwise** direction around a region of **high** pressure
* in an **anticlockwise** direction around a region of **low** pressure

The table gives the Daily Mean Pressure for 3 locations from the large data set on 26/09/2015

|  |  |  |  |
| --- | --- | --- | --- |
| **Location** | Heathrow | Hurn | Leuchars |
| **Daily Mean Pressure** | 1029 | 1028 | 1028 |
| **Cardinal Wind Direction** |  |  |  |

The Cardinal Wind Directions for these 3 locations on 26/09/2015 were, in random order,

W NE E

You may assume that these 3 locations were under a single region of pressure.

(*d*)Using your knowledge of the large data set, place each of these Cardinal Wind

Directions in the correct location in the table.

Give a reason for your answer.

**(2)**

**(Total for Question 3 is 8 marks)**

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

**4.** A large college produces three magazines.

One magazine is about green issues, one is about equality and one is about sports.

A student at the college is selected at random and the events *G*, *E* and *S* are defined as

follows

*G* is the event that the student reads the magazine about green issues

*E* is the event that the student reads the magazine about equality

*S* is the event that the student reads the magazine about sports

The Venn diagram, where *p*, *q*, *r* and *t* are probabilities, gives the probability for each

subset.



(*a*)Find the proportion of students in the college who read exactly one of these

magazines.

**(1)**

No students read all three magazines and P(*G*) = 0.25

(*b*)Find

(i) the value of *p*

(ii) the value of *q*

**(3)**

Given that P(*S* | *E*) = 

(*c*)find

(i) the value of *r*

(ii) the value of *t*

**(4)**

(*d*)Determine whether or not the events (*S* ∩ *E*ʹ) and *G* are independent.

Show your working clearly.

**(3)**

**(Total for Question 4 is 11 marks)**

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

**5.** The heights of females from a country are normally distributed with

* a mean of 166.5 cm
* a standard deviation of 6.1 cm

Given that 1% of females from this country are shorter than *k* cm,

(*a*)find the value of *k*

**(2)**

(*b*)Find the proportion of females from this country with heights between 150 cm

and 175 cm

**(1)**

A female, from this country, is chosen at random from those with heights between

150 cm and 175 cm

(*c*)Find the probability that her height is more than 160 cm

**(4)**

The heights of females from a different country are normally distributed with a standard

deviation of 7.4 cm

Mia believes that the mean height of females from this country is less than 166.5 cm

Mia takes a random sample of 50 females from this country and finds the mean of her

sample is 164.6 cm

(*d*)Carry out a suitable test to assess Mia’s belief.

You should

* state your hypotheses clearly
* use a 5% level of significance

**(4)**

**(Total for Question 5 is 11 marks)**

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

**6.** The discrete random variable *X* has the following probability distribution

|  |  |  |  |
| --- | --- | --- | --- |
| *x* | *a* | *b* | *c* |
| P(*X* = *x*) | log36 *a* | log36 *b* | log36 *c* |

where

* *a*, *b* and *c* are distinct integers (*a* < *b* < *c*)
* all the probabilities are greater than zero

(*a*)Find

(i) the value of *a*

(ii) the value of *b*

(iii) the value of *c*

Show your working clearly.

**(5)**

The independent random variables *X*1 and *X*2 each have the same distribution as *X*

(*b*)Find P(*X*1 = *X*2)

**(2)**

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

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

**TOTAL FOR STATISTICS IS 50 MARKS**