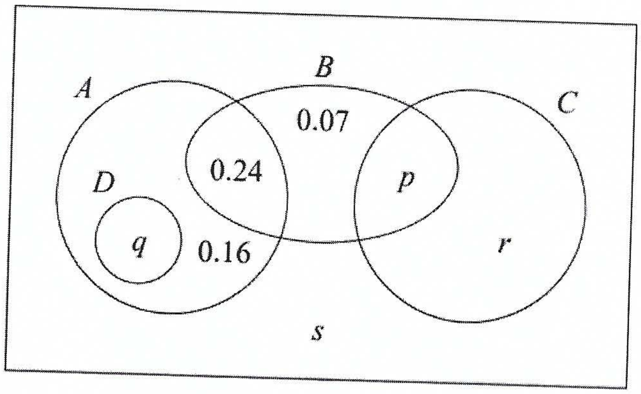
1. The Venn diagram shows the probabilities associated with four events, A, B, C and D



(a) Write down any pair of mutually exclusive events from A, B, C and D

(1)

Given that $P(B) = 0.4$

(b) find the value of p

(1)

Given also that A and B are independent

(c) find the value of q

(2)

Given further that $P(B'|C) = 0.64$

(d) find

(i) the value of r

(ii) the value of s

(4)

Solution 1a

A, C or D, B or D, C

Solution 1b

$P(B) = 0.4 \Rightarrow 0.24 + 0.07 + p = 0.4$

$\Rightarrow p = 0.09$

Solution 1c

If A and B are independent, $P(A \cap B) = P(A)P(B)$

Since $P(A \cap B) = 0.24 \Rightarrow 0.24 = P(A) \times 0.4$

$\Rightarrow P(A) = 0.6$

From the diagram, $0.24 + 0.16 + q = 0.6$

$\Rightarrow q = 0.6 - 0.24 - 0.16$

$\Rightarrow q = 0.2$

Solution 1di

$$\boxed{P(B' \mid C)} = \boxed{0.64}$$

But $\boxed{P(B' \mid C)} = \dfrac{P(B' \cap C)}{P(C)}$

$$= \frac{r}{p+r} = \boxed{\frac{r}{0.09+r}}$$

Hence $\boxed{0.64} = \boxed{\dfrac{r}{0.09+r}}$

$$\Rightarrow 0.64(0.09+r) = r$$

$$\Rightarrow 0.64 \times 0.09 + 0.64r = r$$

$$\Rightarrow 0.36r = 0.0576$$

$$\Rightarrow r = 0.16$$

Solution 1dii

$$S = 1 - (\boxed{q} + 0.16 + 0.24 + 0.07 + \boxed{p} + \boxed{r})$$

$$= 1 - (\boxed{0.2} + 0.16 + 0.24 + 0.07 + \boxed{0.09} + \boxed{0.16})$$

$$= 0.08$$

2. A random sample of 15 days is taken from the large data set for Perth in June and July 1987.

The scatter diagram in Figure 1 displays the values of two of the variables for these 15 days.

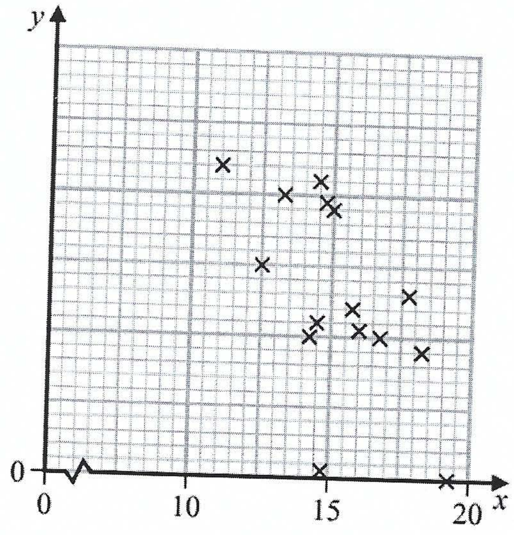


Figure 1

(a) Describe the correlation.

(1)

The variable on the x-axis is Daily Mean Temperature measured in °C.

(b) Using your knowledge of the large data set,

(i) suggest which variable is on the y-axis,

(ii) state the units that are used in the large data set for this variable.

(2)

Stav believes that there is a correlation between Daily Total Sunshine and Daily Maximum Relative Humidity at Heathrow.

He calculates the product moment correlation coefficient between these two variables for a random sample of 30 days and obtains $r = -0.377$

(c) Carry out a suitable test to investigate Stav's belief at a 5% level of significance. State clearly

- your hypotheses
- your critical value

(3)

On a random day at Heathrow the Daily Maximum Relative Humidity was 97%

(d) Comment on the number of hours of sunshine you would expect on that day, giving a reason for your answer.

(1)

Solution 2a Negative

Solution 2bi Rainfall

Solution 2bii mm

Solution 2c

$H_0 : \rho = 0$

$H_1 : \rho \neq 0$

Notice here the question states that Stav believes there is a correlation – not that there is a negative correlation.

Table

Product Moment Correlation

Level	Sample size, n
0.025	30
$\boxed{0.3610}$	

Significance level $= 0.05$

Observed $r = -0.377$

Critical value $= \boxed{-0.3610}$

As we are considering $\rho \neq 0$, we must halve $5\% = 0.05$ to give 0.025 and then read from tables

Since $-0.377 < -0.3610$ (ie Observed < Critical), we can reject H_0 and conclude that there is sufficient evidence at the 5% level to suggest that there is a correlation.

Solution 2d

Since humidity is high and $r < 0$, we would expect a low amount of sunshine

Solution 3d

$$\sigma = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

$$= \sqrt{\frac{17623.25}{27} - 22.5^2}$$

$$= 12.1$$

Solution 3e

Outlier is above (or greater than)

$\text{mean} + 3\sigma = 22.5 + 3 \times 12.1 = 58.8$

~~or~~ $\Rightarrow$ There is only one outlier

~~Outlier is below (or less than)~~

~~$\text{mean} - 3\sigma = 22.5 - 3\times 12.1$~~

Solution 3f

Mean does not change.
Therefore, they must have $a+b = 2\times 22.5 = 45$
Median increases, so both are above the median, so >20. For instance $a=23, b=22$. These values satisfy $a > b > 20$ and $a+b=45$.

Solution 3g

Since a and b are less than 1 standard deviation from the mean, the overall standard deviation will be lower.

4. The discrete random variable D has the following probability distribution

d	10	20	30	40	50
$P(D = d)$	$\frac{k}{10}$	$\frac{k}{20}$	$\frac{k}{30}$	$\frac{k}{40}$	$\frac{k}{50}$

where k is a constant.

(a) Show that the value of k is $\frac{600}{137}$

(2)

The random variables D_1 and D_2 are independent and each have the same distribution as D.

(b) Find $P(D_1 + D_2 = 80)$
Give your answer to 3 significant figures.

(3)

A single observation of D is made.

The value obtained, d, is the common difference of an arithmetic sequence.

The first 4 terms of this arithmetic sequence are the angles, measured in degrees, of quadrilateral Q

(c) Find the exact probability that the smallest angle of Q is more than 50°

(5)

Solution 4a

$$\sum P(D=d) = 1$$

$$\Rightarrow \frac{k}{10} + \frac{k}{20} + \frac{k}{30} + \frac{k}{40} + \frac{k}{50} = 1$$

$$\Rightarrow \frac{60k}{600} + \frac{30k}{600} + \frac{20k}{600} + \frac{15k}{600} + \frac{12k}{600} = 1$$

$$\Rightarrow \frac{137k}{600} = 1 \Rightarrow k = \frac{600}{137}$$

Solution 4b

$P(D_1+D_2=80)$

$= P(D_1=30, D_2=50) + P(D_1=40, D_2=40) + P(D_1=50, D_2=30)$

$= \left(\frac{k}{30}\right)\left(\frac{k}{50}\right) + \left(\frac{k}{40}\right)\left(\frac{k}{40}\right) + \left(\frac{k}{50}\right)\left(\frac{k}{30}\right)$

$= k^2\left(\frac{1}{1500} + \frac{1}{1600} + \frac{1}{1500}\right)$

$= k^2\left(\frac{16}{24000} + \frac{15}{24000} + \frac{16}{24000}\right)$

$= k^2\left(\frac{47}{24000}\right) = \left(\frac{600}{137}\right)\left(\frac{47}{24000}\right) = 0.0376$

Solution 4c

$S_4 = a + (a+d) + (a+2d) + (a+3d) = 360°$

$\Rightarrow 2a + 3d = 180°$

If $a > 50$ then $2a > 100$

$\Rightarrow 3d < 80$

So $d=10$ or $d=20$ only (since $3d = 90 > 80$ for $d=30$)

Therefore $P(d=10 \text{ or } d=20) = \frac{k}{10} + \frac{k}{20}$

$= \frac{3k}{20} = \frac{3}{20} \times \frac{600}{137}$

$= \frac{90}{137}$

5. A health centre claims that the time a doctor spends with a patient can be modelled by a normal distribution with a mean of 10 minutes and a standard deviation of 4 minutes.

(a) Using this model, find the probability that the time spent with a randomly selected patient is more than 15 minutes.

(1)

Some patients complain that the mean time the doctor spends with a patient is more than 10 minutes.

The receptionist takes a random sample of 20 patients and finds that the mean time the doctor spends with a patient is 11.5 minutes.

(b) Stating your hypotheses clearly and using a 5% significance level, test whether or not there is evidence to support the patients' complaint.

(4)

The health centre also claims that the time a dentist spends with a patient during a routine appointment, T minutes, can be modelled by the normal distribution where $T \sim N(5, 3.5^2)$

(c) Using this model,

(i) find the probability that a routine appointment with the dentist takes less than 2 minutes

(1)

(ii) find $P(T < 2 \mid T > 0)$

(3)

(iii) hence explain why this normal distribution may not be a good model for T.

(1)

The dentist believes that she cannot complete a routine appointment in less than 2 minutes.

She suggests that the health centre should use a refined model only including values of $T > 2$

(d) Find the median time for a routine appointment using this new model, giving your answer correct to one decimal place.

(5)

Solution 5a

$X \sim N(10, 16)$

$P(X > 15) = 0.106$

Solution 5b

$H_0: \mu = 10$

$H_1: \mu > 10$

$\bar{X} \sim N(10, \frac{16}{20}) \Rightarrow \bar{X} \sim N(10, \frac{4}{5})$

Solution 5b

$H_0: \mu = 10$

$H_1: \mu > 10$

$\bar{X} \sim N(10, \frac{4}{5})$

Central Limit Theorem

$X \sim N(10, 16) \xrightarrow{\text{Central Limit Theorem}} \bar{X} \sim N(10, \frac{16}{20})$ (Sample size)

Significance Level = $\boxed{0.05}$

Observed Data = $\boxed{11.5}$

calculator

$P(X \geq \boxed{11.5}) = 0.0468 < \boxed{0.05}$

So we can reject H_0 and conclude there is sufficient evidence to suggest the doctor is spending longer with patients.

Solution 5ci

$T \sim N(5, 3.5^2)$

calculator

$P(T < 2) = 0.1956$

Solution 5cii

$$P(T < 2 \mid T > 0) = \frac{P(0 < T < 2)}{P(T > 0)}$$

$$= \frac{0.119119}{0.923476} = 0.129$$

Solution 5ciii

This model gives a substantial probability of $T > 0$ which does not make sense.

Solution 5d

Median m is such that $P(T > m \mid T > 2) = 0.5$

$$\frac{P(T > m)}{P(T > 2)} = 0.5 \Rightarrow P(T > m) = 0.5P(T > 2)$$

$$= 0.5(1 - 0.1956)$$

$$= 0.5 \times 0.8043 = 0.40$$

$$\Rightarrow t = 5.867 \approx 5.9$$