**1.** George throws a ball at a target 15 times.
Each time George throws the ball, the probability of the ball hitting the target is 0.48

The random variable $X$ represents the number of times George hits the target in 15 throws.

(a) Find

(i) $P(X = 3)$

(ii) $P(X \geqslant 5)$

**(3)**

George now throws the ball at the target 250 times.

(b) Use a normal approximation to calculate the probability that he will hit the target more than 110 times.

**(3)**

Solution 1ai

$X \sim B(15, 0.48)$

$P(X=3) = 0.019668 = 0.0197$ (3 sf) ← calculator

Solution 1aii

$X \sim B(15, 0.48)$

$P(X \geqslant 5) = 1 - P(X \leqslant 4) = 0.920$ (3 sf) ← calculator

Solution 1b

$\boxed{n = 250}$, $\boxed{p = 0.48}$

$\text{Mean} = \boxed{np} = \boxed{250} \times \boxed{0.48} = \boxed{120}$

$\text{Variance} = \boxed{np(1-p)} = \boxed{250} \times \boxed{0.48} \times 0.52 = \boxed{62.4}$

$0.52$ ← $1 - 0.48$

Let Y represent no. of hits. Then

$Y \sim N(\boxed{120}, \boxed{62.4})$

$P(Y > 110) \approx P(Y > \boxed{110.5}) = P\left(Z > \frac{\boxed{110.5} - \boxed{120}}{\boxed{\sqrt{62.4}}}\right) = 0.885$

$= P(Z > -1.20263)$

$= 0.885$

**2.** A manufacturer uses a machine to make metal rods.

The length of a metal rod, $L$ cm, is normally distributed with

- a mean of 8 cm
- a standard deviation of $x$ cm

Given that the proportion of metal rods less than 7.902 cm in length is 2.5%

(a) show that $x = 0.05$ to 2 decimal places.

**(2)**

(b) Calculate the proportion of metal rods that are between 7.94 cm and 8.09 cm in length.

**(1)**

The **cost** of producing a single metal rod is 20p

A metal rod

- where $L < 7.94$ is **sold** for scrap for 5p
- where $7.94 \leqslant L \leqslant 8.09$ is **sold** for 50p
- where $L > 8.09$ is shortened for an extra **cost** of 10p and then **sold** for 50p

(c) Calculate the expected profit per 500 of the metal rods.
Give your answer to the nearest pound.

**(5)**

The same manufacturer makes metal hinges in large batches.

The hinges each have a probability of 0.015 of having a fault.

A random sample of 200 hinges is taken from each batch and the batch is accepted if fewer than 6 hinges are faulty.

The manufacturer's aim is for 95% of batches to be accepted.

(d) Explain whether the manufacturer is likely to achieve its aim.

**(4)**

Solution 2a

$L \sim N(\boxed{8}, \boxed{x}^2)$

$P(L < \boxed{7.902}) = 0.025 \Rightarrow \dfrac{\boxed{7.902} - \boxed{8}}{\boxed{x}} = -1.96$

from tables

| P | Z |
|---|---|
| 0.0250 | 1.96 |

(for $P(Z>z)=0.025$) so change sign for $P(Z<z)$

$\Rightarrow x = \dfrac{7.902 - 8}{-1.96} = 0.05$

Solution 2b

$L \sim N(8, 0.05^2) \Rightarrow P(7.94 \leq L \leq 8.09) = 0.849$

## Solution 2c

| p | $P(L<7.94)$ | $P(7.94 \le L \le 8.09)$ | $P(L>8.09)$ |
|---|---|---|---|
| | 0.115069 | 0.849 | 0.03593 |
| Profit (£s) | 0.05 | 0.5 | 0.4 |

Expected Income per 500 rods

$= \sum 500 \times \text{Probability} \times \text{Income}$

$= (500 \times \boxed{0.115069 \times 0.05}) + (500 \times \boxed{0.849 \times 0.5}) + (500 \times \boxed{0.03593 \times 0.4})$

$= £222.311$

Expected Profit = Expected Income − Cost

$= £222.311 - 500 \times 0.2$

$= £122.31$

cost of production of a metal rod is 20p

## Solution 2d

$X \sim B(200, 0.015)$

$P(X \le 5) = 0.9176$

Manufacturer is unlikely to achieve their aim since

$0.9176 < 0.95$

**3.** Dian uses the large data set to investigate the Daily Total Rainfall, $r$ mm, for Camborne.

(a) Write down how a value of $0 < r \leqslant 0.05$ is recorded in the large data set.

**(1)**

Dian uses the data for the 31 days of August 2015 for Camborne and calculates the following statistics

$$n = 31 \qquad \sum r = 174.9 \qquad \sum r^2 = 3523.283$$

(b) Use these statistics to calculate

(i) the mean of the Daily Total Rainfall in Camborne for August 2015,

(ii) the standard deviation of the Daily Total Rainfall in Camborne for August 2015.

**(3)**

Dian believes that the mean Daily Total Rainfall in August is less in the South of the UK than in the North of the UK.
The mean Daily Total Rainfall in Leuchars for August 2015 is 1.72 mm to 2 decimal places.

(c) State, giving a reason, whether this provides evidence to support Dian's belief.

**(2)**

Dian uses the large data set to estimate the proportion of days with no rain in Camborne for 1987 to be 0.27 to 2 decimal places.

(d) Explain why the distribution B(14, 0.27) might **not** be a reasonable model for the number of days without rain for a 14-day summer event.

**(1)**

Solution 3a

It is recorded as 'tr'

Solution 3bi

$$\bar{r} = \frac{\sum r}{n} = \frac{174.9}{31} = 5.6419 \approx 5.64$$

Solution 3bii

$$\text{Standard Deviation} = \sqrt{\frac{\sum r^2}{n} - \bar{r}^2}$$

$$= \sqrt{\frac{3523.283}{31} - 5.64...^2}$$

$$= 9.04559...$$

## Solution 3c

Leuchars is in the north and Camborne is in the South.
The mean is greater for Camborne than Leuchars. Therefore, there is no evidence that Dion's belief is true.

## Solution 3d

The distribution $B(14, 0.27)$ might not be a reasonable model for the number of days without rain for a 14 day summer event because $p = 0.27$ is unlikely to be constant.

**4.** A dentist knows from past records that 10% of customers arrive late for their appointment.

A new manager believes that there has been a change in the proportion of customers who arrive late for their appointment.

A random sample of 50 of the dentist's customers is taken.

(a) Write down

- a null hypothesis corresponding to no change in the proportion of customers who arrive late
- an alternative hypothesis corresponding to the manager's belief

**(1)**

(b) Using a 5% level of significance, find the critical region for a two-tailed test of the null hypothesis in (a)
You should state the probability of rejection in each tail, which should be less than 0.025

**(3)**

(c) Find the actual level of significance of the test based on your critical region from part (b)

**(1)**

The manager observes that 15 of the 50 customers arrived late for their appointment.

(d) With reference to part (b), comment on the manager's belief.

**(1)**

Solution 4a

$H_0 : p = 0.1$

$H_1 : p \neq 0.1$

Solution 4b $X \sim B(50, 0.1)$

| | p = 0.1 |
|---|---|
| n = 50, x = 0 | 0.0052 |
| 1 | 0.0338 |
| 2 | 0.1117 |
| ⋮ | |
| 9 | 0.9755 |
| 10 | 0.9906 |

$\boxed{P(X=0) = 0.0052}$

$P(X \leqslant 9) = 0.9755 \Rightarrow P(X \geqslant 10) = 1 - P(X \leqslant 9)$

$= 1 - 0.9755$

$\Rightarrow \boxed{P(X \geqslant 10) = 0.0245}$

Therefore critical regions are:

$X = 0$ and $X \geqslant 10$

Solution 4c

Actual significance level $= \boxed{0.0052 + 0.0245} = 0.0297$

Solution 4d

Since 15 is in the critical region, there is evidence to support the manager's belief.

**5.** A company has 1825 employees.
The employees are classified as professional, skilled or elementary.

The following table shows

- the number of employees in each classification
- the two areas, $A$ or $B$, where the employees live

| | $A$ | $B$ |
|---|---|---|
| **Professional** | 740 | 380 |
| **Skilled** | 275 | 90 |
| **Elementary** | 260 | 80 |

An employee is chosen at random.

Find the probability that this employee

(a) is skilled,

**(1)**

(b) lives in area $B$ and is not a professional.

**(1)**

Some classifications of employees are more likely to work from home. $H_A^P = 0.65 \times 740 = 481$

- 65% of professional employees in both area $A$ and area $B$ work from home $H_B^P = 0.65 \times 380 = 247$
- 40% of skilled employees in both area $A$ and area $B$ work from home $H_A^S = 0.4 \times 275 = 110$ $H_B^S = 0.4 \times 90 = 36$
- 5% of elementary employees in both area $A$ and area $B$ work from home $H_A^E = 0.05 \times 260 = 13$ $H_B^E = 0.05 \times 80 = 4$
- Event $F$ is that the employee is a professional
- Event $H$ is that the employee works from home
- Event $R$ is that the employee is from area $A$

(c) Using this information, complete the Venn diagram on the opposite page.

**(4)**

(d) Find $P(R' \cap F)$

**(1)**

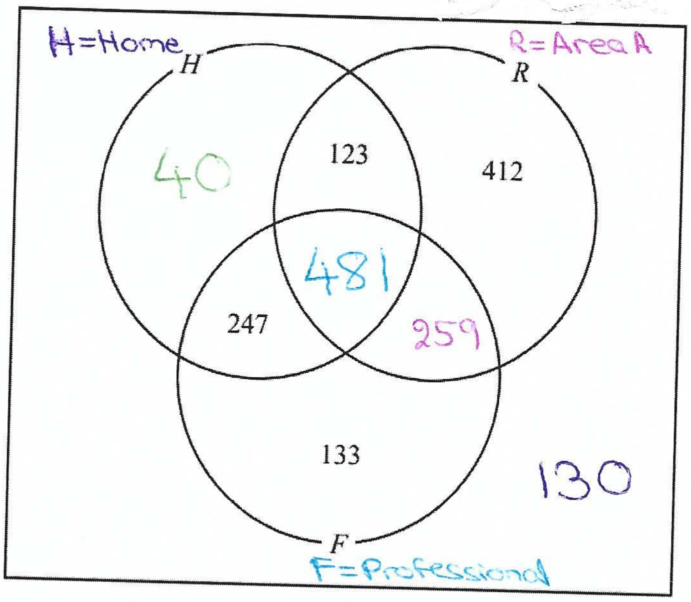
(e) Find $P([H \cup R]')$

**(1)**

(f) Find $P(F \mid H)$

**(2)**

**Question 5 continued** Solution 5c




Solution 5d

$P(R' \cap F) = \frac{247+133}{1825}$

$= \frac{76}{365}$

Solution 5e

$P([H \cup R]') = \frac{133+130}{1825}$

$= \frac{263}{1825}$

Solution 5f

$P(F|H) = \frac{P(F \cap H)}{P(H)}$

$= \frac{247+481}{40+123+481+247} = \frac{728}{891}$

**Turn over for a spare diagram if you need to redraw your Venn diagram.**

Solution 5a

Total no. of employees $= 740+380+275+90+260+80$

$= 1825$

So, P(Employee is skilled) $= \frac{275+90}{1825} = \frac{365}{1825} = \frac{1}{5}$

Solution 5b

P( lives in area B and is not professional)

$= \frac{90+80}{1825} = \frac{170}{1825} = \frac{34}{365}$

Solution 5c

See calculations on question $\Rightarrow (H \cap R \cap F) = \boxed{481}$

Total in Area A $= 740+275+260 = \boxed{1275}$ ← from table

So to find missing value in circle R:

$= \boxed{1275} - (481+123+412) = 259$

From calculations on question,

Total working from home $= 481+247+110+36+13+4$

$= \boxed{891}$

So to find missing value in circle H:

$= \boxed{891} - (481+123+247) = 40$

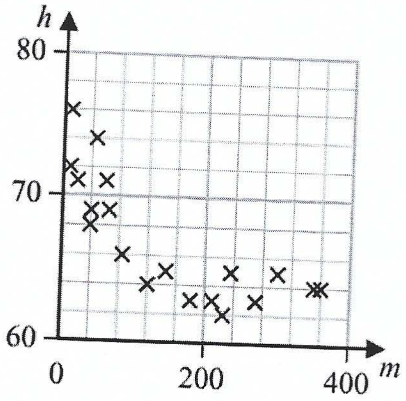
To find missing values outside circles:

$= 1825 - (40+123+481+247+412+259+133) = 130$

6. Anna is investigating the relationship between exercise and resting heart rate. She takes a random sample of 19 people in her year at school and records for each person

- their resting heart rate, $h$ beats per minute
- the number of minutes, $m$, spent exercising each week

Her results are shown on the scatter diagram.




(a) Interpret the nature of the relationship between $h$ and $m$

**(1)**

Anna codes the data using the formulae

$$x = \log_{10} m$$
$$y = \log_{10} h$$

The product moment correlation coefficient between $x$ and $y$ is $-0.897$

(b) Test whether or not there is significant evidence of a negative correlation between $x$ and $y$
You should

- state your hypotheses clearly
- use a 5% level of significance
- state the critical value used

**(3)**

The equation of the line of best fit of $y$ on $x$ is

$$y = -0.05x + 1.92$$

(c) Use the equation of the line of best fit of $y$ on $x$ to find a model for $h$ on $m$ in the form

$$h = am^k$$

where $a$ and $k$ are constants to be found.

**(5)**

## Solution 6a

As the number of minutes spent exercising increases, the resting heart rate decreases.
Negative correlation.

## Solution 6b

$H_0: \rho = 0$

$H_1: \rho < 0$

| Product Moment Co-efficient Level | Sample size, n |
|---|---|
| 0.05 | |
| ⋮ | ⋮ |
| 0.3887 | 19 |

Critical Value = $-0.3887$

Product Moment Corr Co-eff $= -0.897 < -0.3887 = \text{crit}$

There is evidence that the product moment correlation is less than 0. So there is a negative correlation.

## Solution 6c

$$\log_{10} h = -0.05 \log_{10} m + 1.92$$

$$= \log_{10} m^{-0.05} + 1.92$$

$$\Rightarrow \quad h = 10^{\log_{10} m^{-0.05} + 1.92}$$

$$= 10^{1.92} \times 10^{\log_{10} m^{-0.05}}$$

$$\Rightarrow \quad h = 10^{1.92} m^{-0.05}$$