Solutions – Data Handling
Probability & Statistics

Simple Probability

1. a) P(7) = $\frac{5}{50} \rightarrow \frac{1}{10}$

   b) P(Blue 7) = $\frac{1}{50}$

2. 12 face cards, P(face card) = $\frac{12}{52} \rightarrow \frac{3}{13}$

3. a) P(green pencil) = $\frac{11}{20}$

   b) P(blue pencil) = $\frac{7}{19}$ (only 19 pencils left)

4. P(green AND red) = $\frac{25}{50} \times \frac{10}{50} \rightarrow \frac{1}{2} \times \frac{1}{5} \rightarrow \frac{1}{10}$

   (for independent events MULTIPLY probabilities).

5. a) P(W or D) = 0.2 + 0.5 = 0.7
   (add probabilities for mutually exclusive events)

   b) P(Lose) = 0.3

6. a) P(miss) = 0.2 (20%)

   b) P(3 hits in a row) = 0.8 x 0.8 x 0.8 = 0.512

   c) P(H,M,M) = 0.8 x 0.2 x 0.2 = 0.032

7. a) P(not defective) = 0.85 (85%)

   b) 5000 x 0.85 = 4250 should not be defective.

8. a) P(M,M,M) = ½ x ½ x ½ = $\frac{1}{8}$

   b) P(3 boys – first IS a boy) = 1 x ½ x ½ = ¼

Probability from relative Frequency

1. a) P(< 3 yrs old) = $\frac{310}{600} \rightarrow \frac{31}{60}$

   b) $4200 \times \frac{10}{600} \rightarrow \frac{42000}{600} \rightarrow \frac{420}{6} \rightarrow 70$

2. a) P(scenery) = $\frac{80}{500} \rightarrow \frac{8}{50} \rightarrow \frac{4}{25}$

   b) P(25 & facilities) = $\frac{23}{500}$

   c) P(not cost) = $\frac{215}{500} \rightarrow \frac{43}{100}$

3. a) P(still water) = $\frac{35}{110} \rightarrow \frac{7}{22}$

   b) P(< 20 & Fizzy) = $\frac{10}{110} \rightarrow \frac{1}{11}$

4. a) P(new car) = $\frac{40}{120} \rightarrow \frac{1}{3}$

   b) P(18-40 & used car) = $\frac{30}{120} \rightarrow \frac{1}{4}$

Statistical Diagrams

1. 25% contain fewer than 50 matches.
   (Lower quartile is 25%)

2. **A = 25 B = 29 C = 43**
   Range of men = 60 – 18 = 42
   Range of ladies = 21
   Low of Ladies = 22, so **C** = 21 + 22 = **43**

   Men's median = 44, Ladies median = 44 – 15 = 29

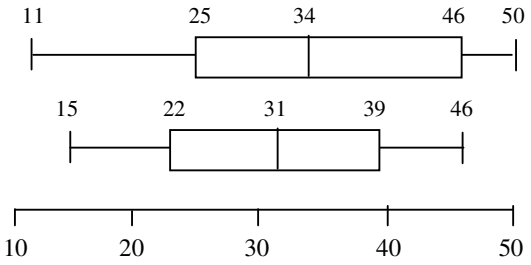
   S.I.R of men = (50 – 34) ÷ 2 = 16 ÷ 2 = 8
   S.I.R Ladies = ¾ of men's so it is 6
   So Ladies IQR = 6 x 2 = 12
   Subtract 12 from UQ to get 37 – 12 = 25.

3. Draw a boxplot for each one

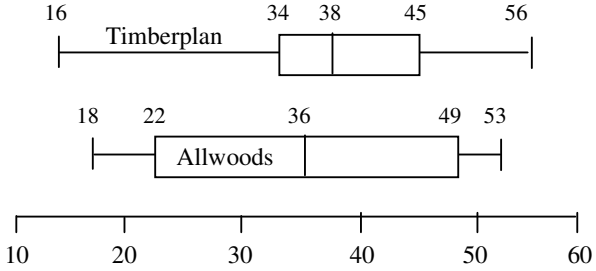
   Use Box plot (or back to back stem & leaf)

   1st Set:
   Lo = 11, Q_1 = 25, Q_2 = 34, Q_3 = 46, Hi = 50

   2nd Set:
   Lo = 15, Q_1 = 22, Q_2 = 31, Q_3 = 39, Hi = 46



4. a) Draw box plots



   b) The semi-interquartile range of timberplan is much lower than that of Allwoods, hence they are more consistent in their deliveries.

5. Draw a **Pie Chart**
   There are 30 pupils.
   Each one can be represented by 12°
   Walk = 156°, Bus = 108°, Car = 72°, Cycle = 24°
   (Check these **total 360°).**
   Then draw a **NEAT** pie chart, and **label** it.

6. a) There are 50 scores, so the median lies between the 25th and 26th scores.
i.e. between 73 and 75. **Median = 74**

b) (UQ – LQ) ÷ 2 LQ is 13th item
UQ is 38th item
So **S.I.R.** = (83 – 69) ÷ 2 = 14 ÷ 2 **= 7**

c) Lo = 63, Hi = 98, Q_1=69, Q_2=74, Q_3=83

7. Put into order
6, 7, 9, 9, 12, 13, **16**, 18, 18, 20, 22, 24, 28
LQ Median UQ

13 items: **Median** is 7th item = **16**
LQ= 9 UQ = 21
Transfer onto sketch.

6 9 16 21 28

Standard Deviation

1. Use formula $s=\sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$

Use 3 columns: x, $(x-\bar{x})$ $(x-\bar{x})^2$

$\sum x = 276$, mean = $276 \div 6 = 46$

$\sum(x-\bar{x})^2 = 84$, SD = $\sqrt{\frac{84}{5}} = 4.098...$

Mean = 46p Standard Deviation = 4.1p

Sugar prices more consistent compared to milk or milk prices more variable than sugar prices.

2. Use formula $s=\sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$

Use 3 columns: x, $(x-\bar{x})$ $(x-\bar{x})^2$

$\sum x = 102$, mean = $102 \div 8 = 12.75$

$\sum(x-\bar{x})^2 = 111.5$, SD = $\sqrt{\frac{111.5}{7}} = 3.991...$

A better formula to use is $s=\sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}}$
to avoid a lot of decimal calculations

$\sum x = 102$, $\sum x^2 = 1412$, $(\sum x)^2 = 10404$
This also gives SD = 3.991..

Mean = 12.75 hrs Standard Deviation = 3.99 hours

Alloa High School were more variable in the hours they spent in study time than Alloa Academy.

3. Use formula $s=\sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$

Use 3 columns: x, $(x-\bar{x})$ $(x-\bar{x})^2$

$\sum x = 750$, mean = $750 \div 5 = £150$

$\sum(x-\bar{x})^2 = 15200$, SD = $\sqrt{\frac{15200}{4}} = £61.64$

Mean = £150 Standard Deviation = £61.64

4. A better formula to use is $s=\sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}}$
to avoid a lot of decimal calculations

Mean = 84.33 pence Standard Deviation = 1.28 pence

The rural garages had a higher average price and the prices were more variable.

5. Use formula $s=\sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$

Use 3 columns: x, $(x-\bar{x})$ $(x-\bar{x})^2$

$\sum x = 36$, mean = $36 \div 6 = 6$

$\sum(x-\bar{x})^2 = 84$, SD = $\sqrt{\frac{80}{5}} = 4$

Mean = 6 Standard Deviation = 4

6. Use formula $s=\sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$

Use 3 columns: x, $(x-\bar{x})$ $(x-\bar{x})^2$

$\sum x = 78$, mean = $78 \div 6 = 13$

$\sum(x-\bar{x})^2 = 76$, SD = $\sqrt{\frac{76}{5}} = 3.898...$

Mean = 13 Standard Deviation = 3.9