Bivariate measurement data

Scholarship Statistics

Covers AS 91581 (Statistics 3.9)

Chapter 2

Investigating bivariate measurement data

Bivariate data involves paired values for two variables in a data set.

A statistical investigation is carried out (using the PPDAC cycle) in order to explore the relationship (if any) between the two variables. The **purpose** of the investigation should be stated clearly, including the units of the variables, and a description of the population.

The context of the data should be well understood – including the **data source**, the **population** from which the sample was selected, the definition of the **variables** and the nature of the measures. A suitably large sample should be randomly chosen from a well-defined population.

Assumptions (if any) about the data should be considered, e.g. that data collection was carried out under the same conditions, using the same measuring instruments and methods.

Bivariate data are plotted as ordered pairs (points) on a scatter graph.

- The **explanatory** variable (the variable which can be altered or controlled, and which provides information about the other variable) is plotted on the *x*-axis
- The **response** variable (which is affected by changes in the explanatory variable) is plotted on the *y*-axis

Note: In some situations there may be no obvious choice for the explanatory or response variables, so either axis can be chosen for each variable

Plotting points on a graph allows features and patterns in the data set to be seen, such as whether there is a linear relationship between the variables, its strength and direction, or whether there are any groupings or unusual values.

It is important to make a visual assessment of the relationship between the variables prior to carrying out a statistical analysis, such as fitting a linear trend line or evaluating the correlation coefficient.

You may decide that the relationship is better described using a **piecewise model** (with two or more straight-line components), or there may be a **non-linear** (curved) **model** that fits better (you should check that the properties of the curve are appropriate to the context).

If there is an obvious **outlier** in the data set, whose validity cannot be ascertained, then carry out two investigations, one with the outlier included and one without the outlier. You should then compare the two investigations, commenting on the impact that the outlier has had on the trend line.

Least squares regression line and gradient

For bivariate data with a linear relationship, a **least squares regression line** is used to summarise the linear relationship (or linear trend) between the two variables. This is the line with the smallest sum of the squares of the **residuals** (in bivariate data, a residual is the difference between a raw data value and the trend line value).

Two other properties of the least-squares regression line are:

- The sum of the residuals is zero.
- The line always passes through the point $(\bar{x}, \bar{y})$, where $\bar{x}$ is the mean of all *x*-coordinates of points on the scatter graph, and $\bar{y}$ is the mean of all *y*-coordinates of points.

The **equation of the trend line** can be obtained automatically (using software such as iNZight). The **gradient** of the trend line gives information about the average change in the response variable for each unit of increase in the explanatory variable.

For example, suppose a set of bivariate data involves an athlete's height (cm) and weight (kg). If the equation of the trend line is: weight = 1.1171 × height – 126.19, then on average for each extra centimetre of height, the weight increases by 1.1171 kg.

There may be inherent restrictions in the range of values for which the trend line applies (e.g. inappropriate *y*-intercepts, or negative *y*-values for certain *x*-values), e.g. the athletes' trend line in the example above gives a weight of –126.19 kg when an athlete's height is zero!

Making predictions

Predictions of the values of response variables for various values of the explanatory variable can be made by substituting into the equation of the trend line.

The confidence that you have in the reliability of the prediction will depend on the amount of scatter there is about the trend line – for a particular *x*-value there may be a considerable range of *y*-values in the raw data, which increases the uncertainty of the prediction for that *x*-value.

You should also consider the appropriateness of the value of the explanatory variable being used for the prediction – is it **interpolation** (a prediction made within the range of explanatory values of the data set) or **extrapolation** (predictions outside the range of values of the explanatory variable of the data set)? Great care should be taken when extrapolating to consider if it is reasonable to expect that the trend in the data set can be continued outside the data set.

The correlation coefficient

A strong linear relationship is one in which the points representing the bivariate data lie close to a line. Pearson's product-moment **correlation coefficient**, *r*, is a number between –1 and 1 which represents the strength and direction of the linear relationship between two variables:

- *r* has no units.
- If $r = \pm 1$, this indicates a perfect linear relationship (with positive slope for $r = 1$ or with negative slope for $r = -1$).
- The closer *r* is to 1 (or the closer *r* is to –1), the stronger the linear relationship.
- The closer *r* is to zero, the weaker the linear relationship (no linear relationship if $r = 0$).
- A correlation coefficient is unaffected by swapping axes, or changing measurement units.

Correlation coefficients are automatically supplied by spreadsheet and iNZight. Note that correlation coefficients only apply to *linear* relationships.

Note: If a non-linear model is fitted to the data, then EXCEL™ supplies the value of the **coefficient of determination**, R^2, which is the proportion of the variation in the response variable which is explained by the regression model. The closer R^2 is to 1, the better the fit of the model. In linear regressions, the coefficient of determination is equal to the square of the correlation coefficient, i.e. $R^2 = r^2$.

Correlation, causality and lurking variables

If two variables are **correlated**, then changes in the values of one variable are *associated* with changes in the values of the other variable.

If two variables have a **causal** relationship, then changes in the values of one variable *cause* changes in the values of the other variable.

Correlation does not imply causality – there may be another **lurking variable** which is affecting the values of both variables. For example, a larger marketing budget may result in more profit for a company, but simply putting more money into marketing will not *cause* larger profit (if the money is being misused for

example) it is the way the money is being spent – the quality and positioning of the advertising, and the competence and experience of the marketing department – that is causing the improvement in profits.

Comparing reports

There may be related reports available for comparison. You should compare the graphs visually, looking for differences in the amount of scatter, and compare correlation coefficients (both of which affect the reliability of predictions).

- If two reports involve the same variables, you may be able to see if the relationship has weakened, strengthened or stayed the same over time.
- If two different explanatory variables are used for the same response variable, you may be able to determine which explanatory variable has the closer relationship with the response variable.

Further investigations

There may be different relationships between the two variables for different categories within the data set. For example, there may be a stronger/weaker relationship between height and weight for males than for females. Various features of iNZight allow bivariate data for two or more categories within the data set to be highlighted on one graph, or displayed separately, each with its own trend line and correlation coefficient. The strengths of the relationships can then be compared between the categories.

Relevance and usefulness of the investigation

A random sample of sufficient size from a well-defined population will probably allow for generalisations from the relationship observed in the sample to a wider context.

However, if a sample of bivariate data has special characteristics, the findings will probably only apply to a population with those characteristics, and there may be limited opportunities to generalise findings to a wider population – thus limiting the usefulness of the investigation.

Writing a report

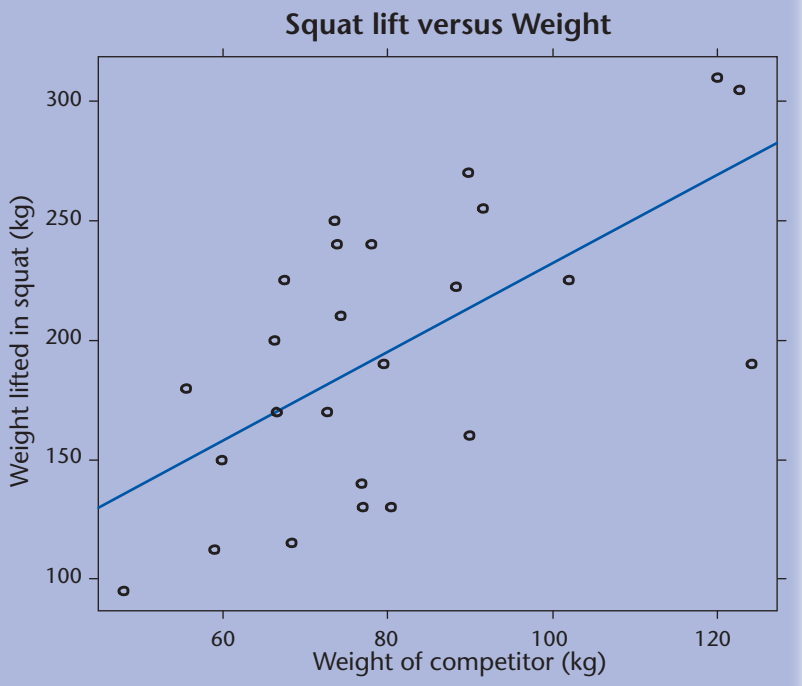
A full report would be based on the Problem-Plan-Data-Analysis-Conclusion (PPDAC) cycle, and may include a discussion of:

- analysis and interpretation of data and graphs
- informed contextual knowledge
- sources and consequences of uncertainty
- applying appropriate models
- stating and evaluating assumptions of models used
- explaining concepts and processes
- offering competing explanations and important follow-up questions
- evaluating claims in statistically based reports.

Example

Power lifting is growing in international popularity. It is a weightlifting sport that involves 3 types of lift: the dead lift, the squat and the bench press. Below is some information from a Power lifting competition. Use the data and the supplied statistical output to answer the questions that follow.

Competitor weight (kg)	Squat lift (kg)
76.8	140
73.8	240
47.9	95
120	310
90	160
59	112.5
67.5	225
59.8	150
66.2	200
89.8	270
66.5	170
72.7	170
88.3	222.5
74.3	210
68.4	115
122.7	305
79.5	190
78	240
73.6	250
80.4	130
55.5	180
124.2	190
91.5	255
77	130
102	225



Linear Trend

Squat = 1.8564 * Weight + 46.49

Correlation = 0.62766

Sample size: 25

Q. 1. Using the graph and other information, describe the relationship between the weight of a competitor and the weight he can lift in a squat for this group of competitors.

2. Predict what weight a 105 kg person would lift and discuss its precision.

3. Discuss any generalisations that could be made from this data, and their validity.

A. 1. Visually, there appears to be a positive linear relationship between the two variables so that, on average, the heavier a lifter is, the more they can lift. However, the amount of scatter in the data would suggest that the relationship is not strong, so that there is a range of lift weights for a competitor of a particular weight, e.g. a competitor weighing 90 kg could lift between 150 kg and 275 kg.

There are no obvious outliers, although the point (124.2,190), which corresponds to a competitor of weight 124.2 kg lifting a weight of 190 kg, is further than average from the trend line. This would pull the trend line down a little, and may affect how well it fits the data and how reliable a model it is.

The equation of the trend line is Squat = 1.8564 × Weight + 46.49. This means that on average, for each extra kilogram of weight the competitor has, he can lift an extra

1.86 kg. While this may be valid for a certain range of values of the explanatory variable (competitor weight), it is likely that the weight that can be lifted will 'plateau' after a certain competitor weight, so that the relationship is no longer linear, but more realistically may have a non-linear model, such as a logarithmic trend.

The correlation coefficient is 0.62766 which indicates that the linear relationship is moderate. This is confirmed by the large amount of scatter of points about the trend line.

Note: The validity of the point (124.2,190) should be checked so that an explanation can be found for its increased distance from the trend.

An analysis of the data without the point (124.2,190) gives a trend line with equation and trend line as shown below.

Linear Trend

Squat = 2.383 * Weight + 8.84

Correlation = 0.71619

Sample size: 24

As can be seen, the trend line relationship is now stronger (correlation coefficient is now 0.71619) and the equation: Squat = 2.383 × Weight + 8.84 indicates that, for each extra kilogram of weight the competitor has, he can lift an extra 2.38 kg, which is more that half a kilogram more than the increase in lift weight predicted by the previous trend line.

A. 2. Substituting Weight = 105 kg into the equation Squat = 1.8564 × Weight + 46.49 gives a prediction of Squat = 241.4 kg. However, due to the large amount of scatter about the trend line, it is more reliable to give a prediction interval – this would be estimated from the graph to be approximately 241.4 ± 60 kg, i.e. it is very likely that a competitor weighing 105 kg will do a squat lift in the range 180 kg to 300 kg. This interval of possible lift weight values is quite wide, but due to the scatter of points about the trend, this range of values cannot be reduced without making the prediction less reliable. This is realistic, since even though two competitors may weigh the same, factors such as age, skill, experience, practice levels, motivation, health on the day, etc. all contribute to the weight the competitor can lift, adding considerable variability to the weight that each competitor can lift, and thus making a very specific prediction using a point estimate such as 241.4 kg (or even a rounded prediction such as 240 kg) of little use.

A. 3. It has not been made clear if this group of lifters is considered to be a random sample of lifters from a particular population, such as all competitors of a particular age range competing in squat power lifting competitions in a particular year (or years) in New Zealand. (It would be important to define competitor age so that this variable is controlled; and a date for the competition, as techniques are constantly improving in sports, giving improved outcomes.)

This group of results may not be a random sample, or may have particular features which make generalisations to a wider population inappropriate.

Even if this were a random sample, the sample size of 25 is too small for reliable conclusions to be drawn, due to sampling variability. Another group of 25 squat weight power lifters may produce a distribution of points with very different characteristics.

A broad conclusion that may be drawn, is that increased weight in a competitor may be associated with heavier lifts in squat power lifting competitions, but that there is likely to be considerable variability in the amount a competitor of a given weight can lift.

Questions
Bivariate measurement data

Question One

Year 2013
Ans. p. 114

Different breeds of cow are known to have different characteristics in the milk they produce. Jersey and Ayrshire are two common breeds in New Zealand. During the 2012–2013 milking year, approximately 350 000 Jersey cows and 18 000 Ayrshire cows were tested four times.

The age and breed of each cow was recorded and, on the four test days, the amounts of milk produced and the results of a chemical analysis of the milk were recorded. The analysis included the amounts of fat and protein in the milk.

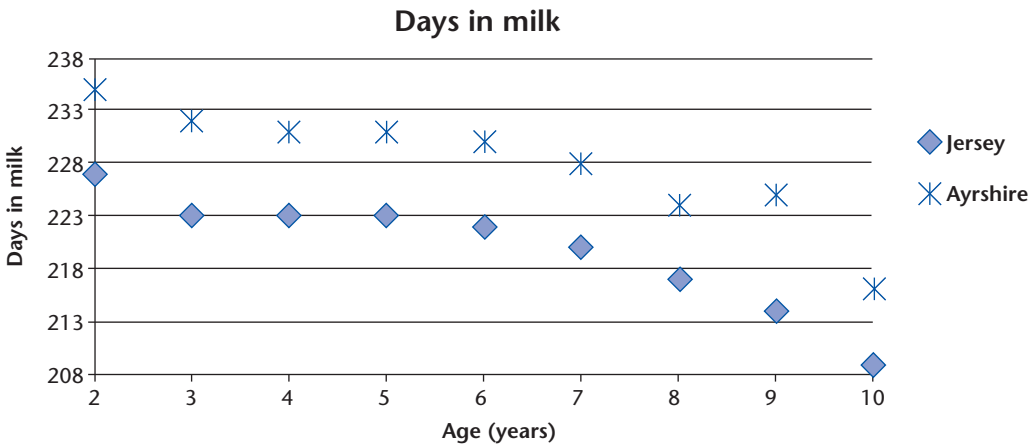
From the data, the annual milk production of each cow was estimated.

a. For each breed and age, the following were calculated:

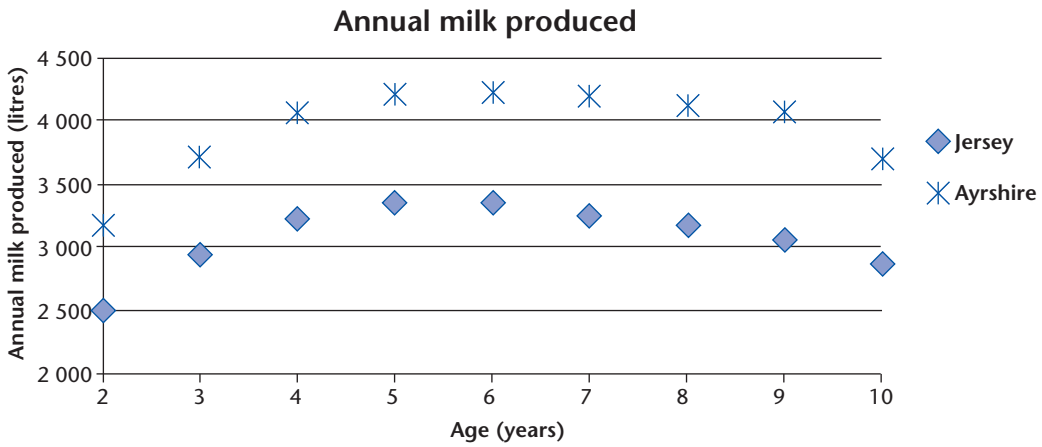
Days in milk	The average number of days in the milking year that a cow produces milk.
Annual milk produced	The average amount of milk produced in the milking year per cow, in litres.
Milk fat percentage	The average percentage of fat in milk.

Graphs 1 to 3 show this information.

Graph 1



Graph 2



Experiments using experimental design principles

Scholarship Statistics

Covers AS 91583 (Statistics 3.11)

Chapter 4

Conducting an experiment to investigate a situation using experimental design principles

A statistical experiment investigates **causality** – the aim is to determine whether an **intervention** or **treatment** (the value of the explanatory variable that is chosen to be given to each individual in a group) *causes* a change in a dependent variable.

The context of the experiment should be well understood and a reasoned prediction considered. For example, a question may be, 'Does the playing of classical music during milking increase milk production per cow for Jersey cows in New Zealand?' An experiment will be carried out to investigate – a sample of New Zealand Jersey cows are the **experimental units**; the playing (or not) of classical music during milking is the treatment.

Research will inform the design of the experiment and your prediction. It is important to identify factors other than the treatment that may have an impact on the response variable and cause extra variability – these factors should be controlled (held constant or removed), so that any differences observed in the response variable are solely attributable to the treatment.

For the example above, the lighting and milking techniques in the cowshed, the health of the cows and their grazing conditions, the timing of milking, etc. should be held constant – the only difference from normal conditions should be the playing (or not) of the music.

Experimental design

Paired comparisons may be used, in which outcomes from one treatment are compared with outcomes from another treatment (or no treatment) for the same group of participants. In this case there will be **bivariate data** (a measurement from each treatment for each participant). Paired dot plots (with related dots joined by arrows) can be used for comparisons, or a box-and-whisker plot of the observed differences.

Alternatively, participants are randomly separated into two groups, using **randomisation** techniques. One group is given a new treatment while the other group is given a **placebo** (a neutral treatment) or is used as a **control group** (a group receiving no treatment, or receiving an existing or established treatment). In this situation results from the two groups are not linked: comparisons can be made using side-by-side dot plots and box-and-whisker plots.

Group sizes should be as large as is practical to better balance the characteristics of the two groups and so reduce variability in the results.

For some experiments, it may be appropriate to carry out repeated measurements (a process called **replication**). Taking repeated measurements of the response variable for each selected value of the explanatory variable is good experimental practice because it provides insight into the variability of the response variable.

It may also be of interest to determine whether the effects of a treatment diminish or change over time.

Re-randomisation under chance acting alone

Although randomisation is used to create two groups with characteristics which are as similar as possible, the observed difference in the means (or medians) from a single experiment may not necessarily be reliable for drawing conclusions about whether the treatment caused a difference in the response variable.

To investigate further, **re-randomisation under chance acting alone** is carried out (this can be done automatically using **iNZightVIT**), in which participants' results are repeatedly and randomly re-grouped (regardless of the intial groupings) and the re-sample differences in means (or medians) calculated and plotted. A re-sample distribution of these differences between means (or medians) is formed, and the number of re-samples with a difference of size equal to or greater than the observed difference is calculated (a tail proportion is supplied by the software).

- If the proportion in the tail is less than 10% (i.e. fewer than 100 out of 1 000 re-samples had a size difference equal to or larger than the observed difference), then it is unlikely that the observed difference occurred by chance alone. There is evidence that chance is not acting alone, and that the treatment has affected the response variable.
- If the proportion in the tail is larger than 10% then it is possible that the observed difference was due to chance alone. Alternatively, the difference could be due to chance and the treatment working together.

Making a causal inference

When producing a report on an investigation using experimental design principles, your report should be structured according to the PPDAC cycle.

Preferably, the conclusions you reach about your experimental units will be able to be generalised. In order to make a **causal inference**, in which results from an experiment are generalised to make a causal statement concerning a wider population, it is important that the sample on which the experiment was based was a suitably sized random selection from a well-understood population. It may even be possible to extend the results to other populations which share the characteristics of the population from the experiment (this needs to be done with care).

The study would need to have been designed and executed according to correct experimental design principles, with consistent measures taken in controlled conditions for two suitably-sized, balanced groups, so that comparisons are fair and justified.

At all times, factors that may influence results should be identified and their effects should be controlled or minimised.

It should be noted that when dealing with people, various psychological elements can come into play – such as the **placebo effect**, when people improve because they believe they have been treated, when in fact they haven't. Rather than changes resulting from the treatment itself, improved performances can arise simply because an activity is being observed, or through competitiveness when one group is compared with another, etc.

Note: To avoid inadvertently influencing the outcome of an experiment, participants can be kept uninformed of the purpose of the experiment, which is called **single blinding**. If both the people collecting the information and the participants are kept uninformed, then it is called a **double-blind** experiment.

Example

Does a specific exercise programme lower the age at which an infant first walks unassisted?

The experiment was designed to investigate whether giving very young infants specific exercises lowers the age at which the infants start to walk.

12 very young male infants were randomly allocated to either the exercise group or the control group.

- The parents of the six infants allocated to the exercise group were instructed to give their infant a programme of specific exercises for 12 minutes each day.
- The six infants in the control group had no regular exercise programme.

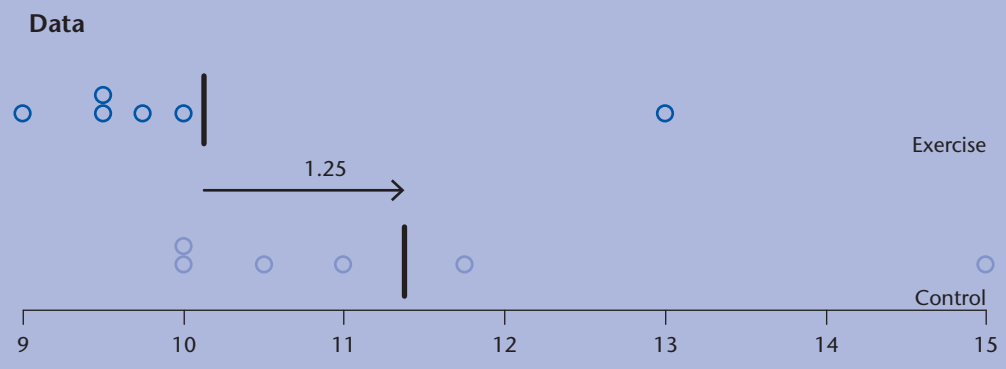
The ages, in months, at which these infants first walked without support was recorded and is shown in the following table.

(Source: Zelazo, P. R., Zelazo, N. A., and Kolb, S. (1972). 'Walking' in the Newborn. *Science*, Vol. 176, pp. 314–315.)

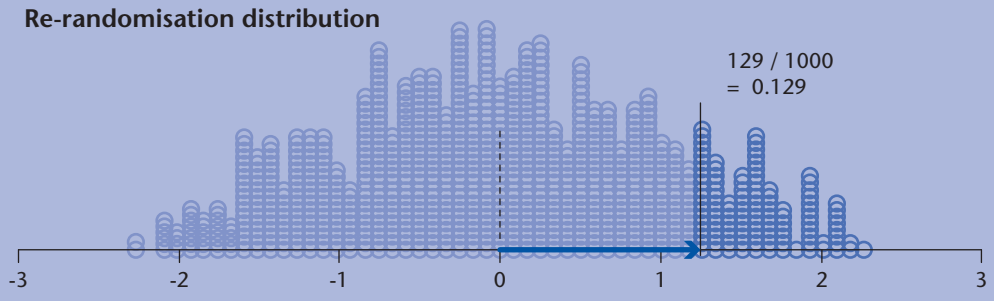
Age at which infant walked unassisted							
Treatment	Age (months)	Age (months)	Age (months)	Age (months)	Age (months)	Age (months)	Mean age (months)
Exercise	9	9.5	9.75	10	13	9.5	10.125
Control	11	10	10	11.75	10.5	15	11.375

The difference in mean age of walking is 11.375 – 10.125 = 1.25 months

Side-by-side dot plots are drawn for the two groups, and vertical lines are used to mark the positions of the group means. The difference between the group means is displayed and its direction marked with an arrow on the plot.



In order to investigate the effects of chance on the experimental results, the following re-randomisation graph was produced using iNZightVIT software.



The tail proportion of 129/1000 means that 129 out of 1 000 re-randomisations, when chance was acting alone, had a difference in mean walking age of 1.25 months or greater.

Therefore the observed difference in mean walking age of 1.25 months is not unusual when chance is acting alone (i.e. when the treatment is not a factor). Therefore chance could be acting alone, or the treatment (doing specific exercises) as well as chance could be acting together – there is not enough information to make a call as to which alternative is true.

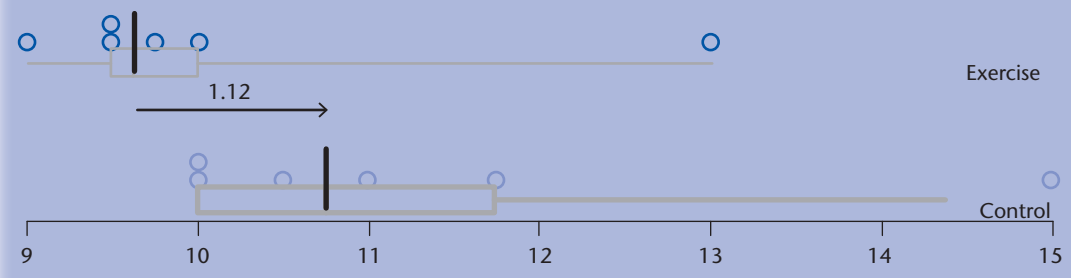
It is interesting to compare these results with those for the differences in median age of walking.

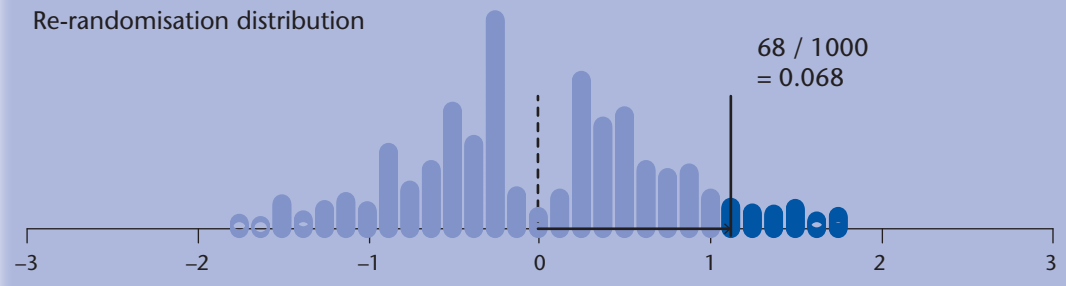
The difference in median age of walking for these samples is 10.75 – 9.625 = 1.125 months

Using iNZightVIT software, the following graphs are produced. Vertical lines mark the positions of the group medians, and the difference (and direction) in group medians is displayed using an arrow. Note that the difference in medians is recorded on the graph to two decimal places as 1.12.

Module: Randomisation test Variable: Age Quantity: median Statistic: difference File: Walking age by treatment.xlsx

Data





The tail proportion of 68/1 000 means that 68 out of 1 000 re-randomisations, when chance was acting alone, had differences in medians of 1.125 months or greater.

This shows that a difference in group medians of 1.125 months or more is unlikely to have occurred through chance acting alone. Therefore there is evidence that chance is not acting alone, and that the specific exercises have affected the age of walking for these infants.

Note: Unlike means, medians are unaffected by outliers (such as a walking age of 15 months for a member of the control group of infants), so there was more variability in the differences in re-sample mean walking ages than in the differences of re-sample median walking ages.

Conclusion

Overall there is evidence, based on the analysis of the difference in median walking times, to conclude that a specific exercise programme may reduce the age at which an infant first walks for this group of young male infants. However, taking the analysis of the difference in mean walking times into account, the evidence is not strong, as it is possible that chance alone caused the observed differences.

In order to generalise these results, a considerably larger sample size should have been used so that there is less likely to be imbalances within the groups (leading to extra variability in outcomes).

Close supervision would also need to have been carried out to ensure that the exercises were administered as described, and that no other significant variables were affecting outcomes (such as illness, or members of the non-exercise group using equipment or carrying out activities that also encouraged their infants to walk earlier).

In order to generalise to a wider community of infants, such as all New Zealand male infants, it would also need to have been established how 'typical' a cross-section of infants was used in the experiment, and at what precise age the exercising began and for how long a period it was carried out. It would be of interest to repeat the experiment with other groups, such as female infants, or infants from other countries.

Questions
Experiments using experimental design principles

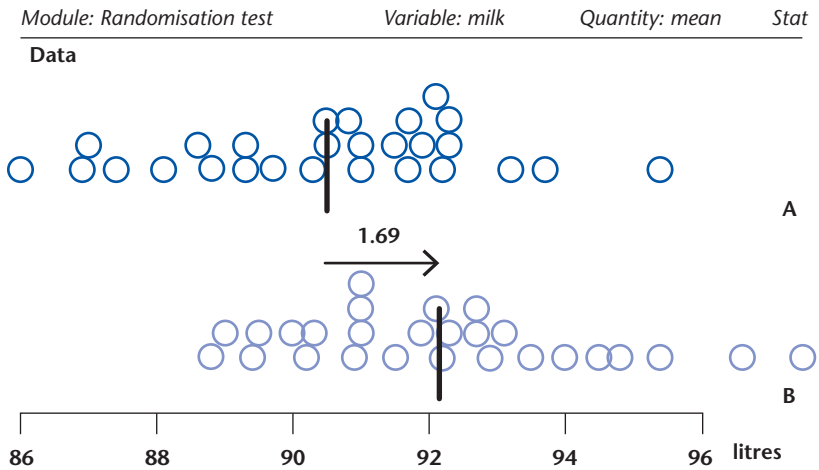
Question One

Year 2013
Ans. p. 118

It is known that the presence of internal parasites adversely affects the health of a cow and the amount of milk it produces. An experiment to test the effectiveness of a new anti-parasite product compared with a currently used product was carried out on a small herd of 54 cows. The cows were assigned randomly to two equally sized groups; group A and group B. Cows in group A were treated with the currently used product, and cows in group B were treated with the new product. The cows were kept as one herd and had the same grazing conditions. Six weeks after receiving the treatment the total amount of milk each cow produced over a period of one week was recorded (in litres).

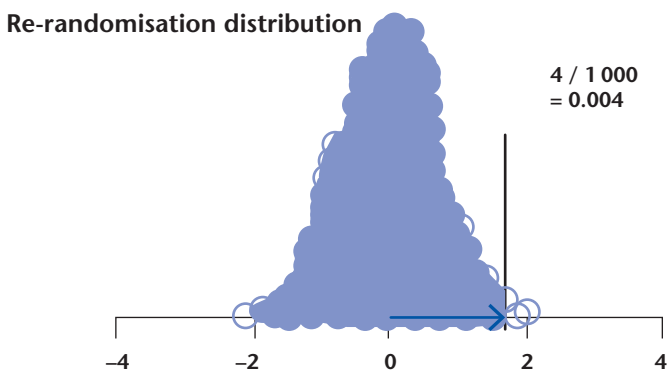
Figure 1 shows a dot plot of the data. The vertical lines show the group means of 90.49 litres for group A and 92.18 litres for group B. The difference in the group means of 1.69 litres is also displayed.

Figure 1



A randomisation test was carried out on the data and the resulting test output is shown in Figure 2. The tail proportion produced by the test is 0.004.

Figure 2



Answers and explanations

Time series data

Question One:

p. 10

a. The graph of milk production per milking year shows steady production levels of between 14 000 and 15 000 million litres per year in milking years 1–5 (Winter 2003 to Autumn 2008). Then in the 6th milking year (Winter 2008 to Autumn 2009) annual production rose sharply to a peak of 18 000 million litres per year, dropping back to just under 16 000 million litres per year in the 7th milking year (Winter 2009 to Autumn 2010). In the final year recorded (8th milking year from Winter 2010 to Autumn 2011) annual production rose again to around 17 000 million litres per year.

The graph of milk production per season shows a strong seasonal pattern. Highest production is usually in Spring with production around 5 000–6 000 million litres (although this has shown a reduction to around 4 500 million litres in milking years 7 and 8). Lowest production is generally in autumn, of around 2 000 million litres between milking years 1 and 5, thereafter around 3 000 million litres in milking years 6 and 7. There was an unexpectedly high Autumn production level in milking year 8 (Autumn 2010) – with Autumn the highest producting season for that year.

Peak production was 6 500 million litres in Spring 2004, and lowest production was in Autumn 2005 at around 1 500 million litres.

The seasonal fluctuations in milk production are high (3 500 million to 5 000 million litres) in the first three milking years (from Winter 2003 to Autumn 2006); then from milking years 4 to 6, seasonal fluctuations reduce to around 3 000 million litres; in the final two milking years seasonal fluctuations are below 2 000 million litres.

Despite reducing Spring peaks, total annual milk production has increased over the 32 seasons because the autumn troughs of production are much higher.

b. Visually extrapolating the line showing milk production per milking year (Graph 1) would produce an estimated annual milk production for 2013 of around 18 000 million litres.

Using Table 1, for milking year 8, the proportion of the total annual production of milk that was produced in Autumn was $\frac{4661}{17\,139}$. Using this same proportion for 2013, a prediction of milk production for Autumn 2013 would be

$\frac{4661}{17\,139} \times 18\,000 = 4\,900$ million litres (2 s.f.)

(Alternatively, using a graphics calculator to fit a trend line to the points in Table 3, the equation of the trend line is: Annual milk production = 480.5 × Milking year + 13 276.5

Substituting Milking year = 10, gives a prediction for annual milk production in 2013 of 18 000 million litres (2 s.f.).

Alternatively, values in Table 2 are increasing on average by $\frac{17\,139 - 14\,746}{3} = 800$ million litres per year. So the prediction for 2013 (milking year 10) would be

17 139 + 2 × 800 = 19 000 million litres (2 s.f.))

An alternative approach would be to use an additive model and note that the average milk production per season for year 8 is $\frac{17\,139}{4} = 4\,285$, so the individual seasonal effect for Autumn production is (4 661 – 4 285) = +376 million litres.

So the prediction for Autumn 2013 would be

$\frac{17\,139}{4} + 376 = 4\,900$ million litres (2 s.f.)

The validity of the prediction is affected by extrapolating two years beyond the last data values – there are many external influences (e.g. weather, disease, politics, export quotas) that could affect milk production after this length of time, making forecasts invalid.

Also there has been quite a bit of variability in the recent milking years on which this estimate was based; in particular, milk production in Autumn 2010 (milking year 8) was unexpectedly high. This situation may be unlikely to be repeated, so that the estimate for Autumn 2013 is too high.

c. All prices are deflated relative to the base year (divide by the CPI then multiply by 1 000).

Year	Milk fat price ($ per kg)	Dairy land sale value ($ per hectare)
1998	38.32	13 265
2002	47.60	15 301
2006	60.17	26 937
2010	55.25	25 634

In real terms (using deflated prices):

Milk fat price has increased by 44% between 1998 and 2010 but the increase has not been steady – there was a 24% increase between 1998 and 2002, a 26% increase between 2002 and 2006, and a 8% drop between 2006 and 2010.

Dairy land sale value has increased by 93% between 1998 and 2010, but the increase has not been uniform – there was a 15% increase between 1998 and 2002, a 76% increase between 2002 and 2006, and a 5% drop between 2006 and 2010.

Dairy land prices increased over the 12 years by more than double the percentage increase in milk fat prices.

Over this period, the CPI increased by 43% (increasing by 15% between 1998 and 2002, by 13% between 2002 and 2006, and by 10% between 2006 and 2010). So over the 12 years, milk fat prices have increased in line with CPI increases, whereas dairy land prices have outstripped CPI increases by more than a factor of two.