Stats 101/101G/108 Workshop

Observational studies, experiments, polls and surveys [OSE]

2019

by Leila Boyle

Stats 101/101G/108 Workshops

The Statistics Department offers workshops and one-to-one/small group assistance for Stats 101/101G/108 students wanting to improve their statistics skills and understanding of core concepts and topics.

Leila's website for Stats 101/101G/108 workshop hand-outs and information is here: **www.tinyURL.com/stats-10x**

Resources for this workshop, including pdfs of this hand-out and Leila's scanned slides showing her working for each problem are available here: **www.tinyURL.com/stats-OSE**

Leila Boyle
Undergraduate Statistics Assistance, Department of Statistics
Room 303.320 (third floor of the Science Centre, Building 303)
l.boyle@auckland.ac.nz; (09) 923-9045; 021 447-018

Want help with Stats?

Stats 101/101G/108 appointments

Book your preferred time with Leila here: **www.tinyURL.com/appt-stats**, or contact her directly (see above for her contact details).

Stats 101/101G/108 Workshops

Workshops are run in a relaxed environment, and allow plenty of time for questions. In fact, this is encouraged ☺

Please make sure you bring your calculator with you to all of these workshops!

- **Preparation at the beginning of the semester:**

  Multiple identical sessions of a preparation workshop are run at the beginning of the semester to get students off to a good start – come along to whichever one suits your schedule!

  - Basic Maths and Calculator skills for Statistics **www.tinyURL.com/stats-BM**

- **First half of the semester**

  Five theory workshops are held during the first half of the semester:

  - Exploratory Data Analysis **www.tinyURL.com/stats-EDA**
  - Proportions and Proportional Reasoning **www.tinyURL.com/stats-PPR**
  - Observational Studies, Experiments, Polls and Surveys **www.tinyURL.com/stats-OSE**
  - Confidence Intervals: *Means* **www.tinyURL.com/stats-CIM**
  - Confidence Intervals: *Proportions* **www.tinyURL.com/stats-CIP**

- **Second half of the semester**

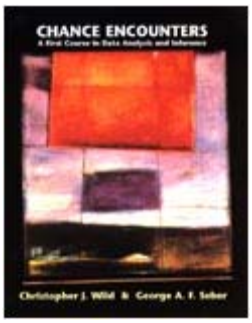
  Five theory workshops and one computing workshop are held during the second half of the semester:

- **Statistics Theory Workshops**
  - Hypothesis Tests: *Proportions* **www.tinyURL.com/stats-HTP**
  - Hypothesis Tests: *Means (part 1)* **www.tinyURL.com/stats-HTM**
  - Hypothesis Tests: *Means (part 2)* **www.tinyURL.com/stats-HTM**
  - Chi-Square Tests **www.tinyURL.com/stats-CST**
  - Regression and Correlation **www.tinyURL.com/stats-RC**
- **Computer Workshop:** Hypothesis Tests in *SPSS* **www.tinyURL.com/stats-HTS**
- **Useful Computer Resource:**

  If you haven't used SPSS before, you may find it useful to work your way through this self-paced workshop: **www.tinyURL.com/stats-IS**

Observational studies, experiments, polls and surveys [OSE]: Notes

To really understand this material, I recommend you read the entirety of *Chapter 1 – What is Statistics?* from



"CHANCE ENCOUNTERS: A First Course in Data Analysis and Inference"

by Chris Wild and George Seber

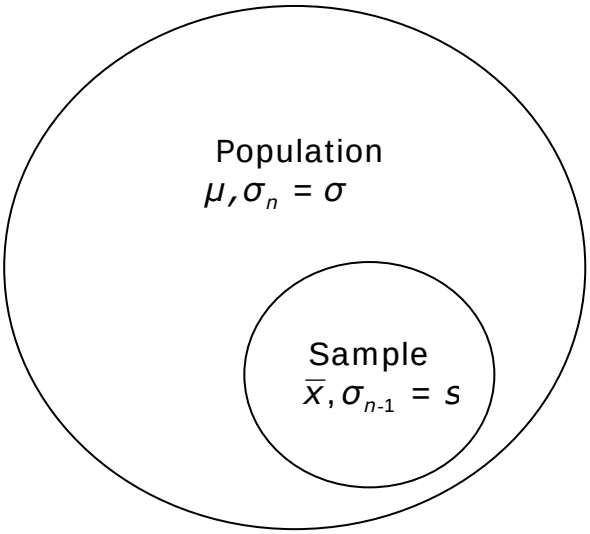
Published by John Wiley & Sons, New York

You can get this on short loan (*2 hours*) or on a 28-day loan from the University of Auckland General Library. Use the call number *519.5 W668c*.

- **Observational studies and experiments**
  - **Unit**
    The individual "thing" data is being recorded about.
    (Alternative names: *entity*, *individual*, *case*)
  - **Variable**
    A characteristic of each unit that we measure or record.
  - **Response variable**
    The outcome variable on which comparisons are made.
- **Observational Studies**
  - A study where the data comes from observing and recording things as they are in the world, or as they unfold over time, without the investigator actively changing anything.
  - A study in which researchers simply compare units that happen to have received each of the levels of the factor of interest. Unit/individual/case "decides" what level they want/get.
  - Should use some form of random sampling → representative samples.
  - CANNOT prove cause and effect – often useful for identifying possible causes of effects, but cannot reliably establish causation.
  - Types of observational studies include:
    - ***Cross-sectional***:
      A study which observes a group of individuals or units at a point in time. It is a descriptive study, providing a "snapshot" at a particular point in time.
    - ***Longitudinal***:
      A study which observes the same group of individuals or units over a long period of time. Comprised of a series of cross-sectional studies.

- **Experiments**
  - A study in which the researcher controls (or manipulates or changes) the conditions experimental units experience, that is the investigator allocates treatment/s to units.
    - **Treatment**
      Something researchers administer to experimental units.
  - Well-designed experiments use randomisation for treatment allocation
  - Can prove cause and effect
  - **Control group**
    May receive no treatment, a placebo or some other existing treatment.
  - **Placebo** – inert (inactive) "dummy" treatment
  - **Placebo effect** – common response where people show signs of "improvement" when they believe they have taken the real treatment.
  - **Blinding**
    A group of people who don't know which treatment the subjects receive.
    - **(Single) Blind** – subjects themselves
    - **Double Blind** – subjects and researchers in contact with subjects
  - **Block** – A group of units which have something in common.
  - **Blocking** – Grouping units into similar groups (blocks) in an experiment.
  - Types of experiments include:
    - **Completely Randomised Design**
      Allocate treatments to units entirely by chance to try to make the treatment groups as similar as possible.
    - **Randomised Block Design**
      Group (block) units by some known factor, then randomly allocate treatments to units within each block to try to balance out the unknown factors.

- **Polls and Surveys**
  - A poll or survey is the process of collecting data from a **sample** of the population to determine information about the whole **population**.
  - A **population** is the complete set of individuals, objects or units that we want information about. E.g., the population of all adults in Australia, the population of all students studying at the University of Auckland.
  - A **sample** is a portion, or subset, of units in the study population which information is collected on.



- **Parameter**
  Numerical characteristic of the population or distribution.
- **Statistic**
  A number calculated from the data, often used to estimate an unknown parameter.
- Public opinion polls are commonly reported in New Zealand's news media. Surveys provide important information for all kinds of research fields (e.g. marketing, psychology, health, sociology). A poll or survey may focus on preferences (e.g. voting), behaviour (e.g. smoking/drinking habits), or factual information (e.g. income), depending on its purpose.
- Since polls and surveys are always based on a **sample** of the **population**, the success of it depends on how **representative** the **sample** is of the **population of interest**. A **representative sample** is a sample that reflects the characteristics of the population.
- Why sample? It is cheaper, faster and more practical (easier) than taking a **census** (attempt to collect information about the whole population).
- **Random sampling**
  Technique where each unit is selected entirely by chance. The larger the sample, the more effective randomisation is at producing a representative sample.
- **Simple Random Sample (SRS)**
  Sampling <u>without</u> replacement

  Steps to carry out an SRS:

  (a) Work out population size, *N*.

  (b) Number each population member from 1 to *N*.

  (c) Use random numbers to pick the sample, *n.*

  (d) As we are sampling <u>without</u> replacement, discard any repeated observations.
- Bear in mind that polls and surveys, even if they are representative, are a snapshot of the population of interest at the point in time they are taken. Opinions and population characteristics do change over time!
- **Pilot survey**
  Small survey carried out before the main survey. Used to identify any potential problems with the survey design.

- **Errors in Polls and Surveys:**

1. Sampling errors

- Arise from taking a sample rather than a census, unavoidable.
- Also known as chance or random errors.
- Usually bigger in smaller samples than larger ones.
- Size may be estimated by statistical methods.

2. Non-sampling errors

- Errors that occur during the data collection process → try to minimise in design of survey by using a pilot survey.
- Can be much larger than sampling errors – always present
- Can be virtually impossible to correct for after data is collected
- Virtually impossible to determine how badly they can will affect the result

➢ **Selection bias**
Population sampled is not exactly the population of interest.

➢ **Non-response bias**
People who have been specifically chosen to be surveyed do not respond. Non-respondents may behave differently from respondents.

➢ **Self-selection**
People who have not been specifically chosen to be surveyed – anyone may take part. (i.e. people themselves decide whether or not to participate).

➢ **Question effects**
Variations in wording can an effect on responses.

➢ **Interviewer effects**
Different interviewers asking the same question can obtain different results.

➢ **Behavioural considerations**
People tend to answer questions in a way they consider to be socially desirable.

➢ **Transferring findings**
Taking data from one population and transferring the results to another.

➢ **Survey format effects**
Factors such as type of survey (mail/phone/face-to-face interview), question order, layout of written survey, self-administered questionnaire or interviewer, … etc., can affect the results.

Check you understand! Questions

1. A study examined the effects of stress on the subsequent onset of diabetes. People living in stressful environments were selected and followed through time to see if they developed diabetes. All instances of diabetes were recorded as they were diagnosed. This type of study is **best** called:
   (1) a longitudinal study.
   (2) a cross-sectional study.
   (3) a census.
   (4) a completely randomised design study.
   (5) a randomised block design study.

2. Which **one** of the following statements is **false**?
   (1) 'Cross-sectional' and 'longitudinal' studies are observational studies.
   (2) An observational study is often useful for establishing the cause of an effect.
   (3) In an observational study, the researcher compares units that happen to have received each of the treatments.
   (4) An experiment can be conducted only when it is possible (ethically, financially, etc.) for the researcher to allocate treatments to the experimental units.
   (5) A well designed, and well executed experiment can reliably establish causation.

3. We need to determine whether a new treatment (treatment X) is effective in treating a certain disease. Which **one** of the following experiments will give the **clearest** information?
   (1) All the patients are given treatment X.
   (2) Those patients who have had a recent attack of the disease are given treatment X. The others are given a placebo.
   (3) Patients that ask for treatment X are randomly allocated to either treatment X or a placebo. Those that do not ask for treatment X are given a placebo.
   (4) Patients arriving in the morning receive treatment X and those arriving on the afternoon receive a placebo. It is randomly decided whether to give treatment X in the morning or in the afternoon.
   (5) Half the patients randomly selected to receive the treatment. The other half receives a placebo.

Question 4 refers to the following information.

Chi and Snyder (2011) from the University of Sydney investigated the effect of electrical stimulation in problem solving. In part of this study 40 participants were trained to solve problems of a particular kind in a certain way. During training, half of the participants were randomly allocated to receive non-invasive electrical stimulation of the brain and the other half to receive sham stimulation. The sham stimulation mimicked electrical stimulation without prolonged use of electrical current. The participants could not tell, and were not told, which one of the two types of stimulation they received.

They were then asked to solve an unfamiliar problem that required fresh insight. As shown in Figure 2, 12 participants in the electrical stimulation group solved the problem successfully whereas only 4 participants in the sham stimulation group solved the problem successfully. There is a 0.4 difference in success rates between the two groups.

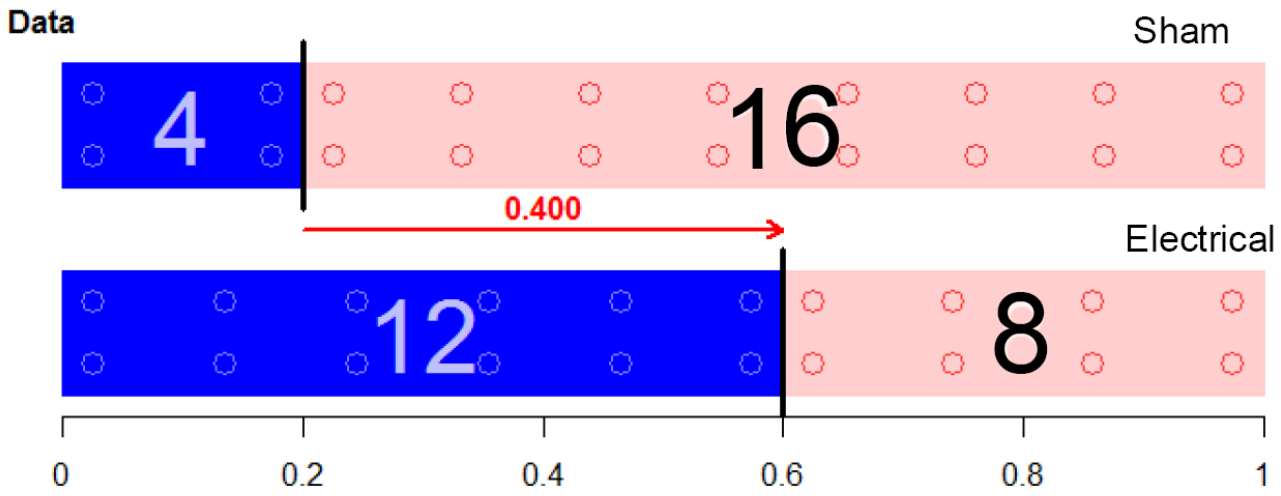


Figure 2: Observed numbers and difference in success rates

4. Which one of the following statements is false?

(1) This study is an example of an experiment.

(2) The participants were randomly allocated to either an electrical stimulation group or a sham stimulation group.

(3) A randomisation test can be used to see if the difference between the two groups could have occurred just by chance alone.

(4) No placebo was used in this study.

(5) This study's design is an example of a completely randomised design.

Question 5 refers to the following information. A study (Chijiiwa *et al.*, 2015) was conducted to investigate whether or not domestic dogs evaluate humans interacting with one another. Fifty-four domestic dogs and their owners participated in the study. The owners were not told the purpose of the study.

Each dog, along with its owner, was randomly allocated to one of three groups of 18: a control group, a helper group and a nonhelper group. Each dog and its owner participated in four trials under identical conditions.

In each trial the owner and their dog sat between an actor and a neutral person. The owner then tried to take the lid off a container. For those in the helper and nonhelper groups, the owner had been instructed to ask the actor for help to take the lid off. Those in the helper group received help from the actor, while for those in the nonhelper group the actor refused to help. Those in the control group tried to take the lid off but did not ask for help. For all three groups the container with the lid on was then put down and, while the owner watched, both the actor and the neutral person offered the dog a treat at the same time.

The number of times, out of the four trials, that each dog chose to take the treat from the actor (chose the actor) is shown in Figure 2. Also shown is the mean for each group (solid line) and the overall mean (dashed line).

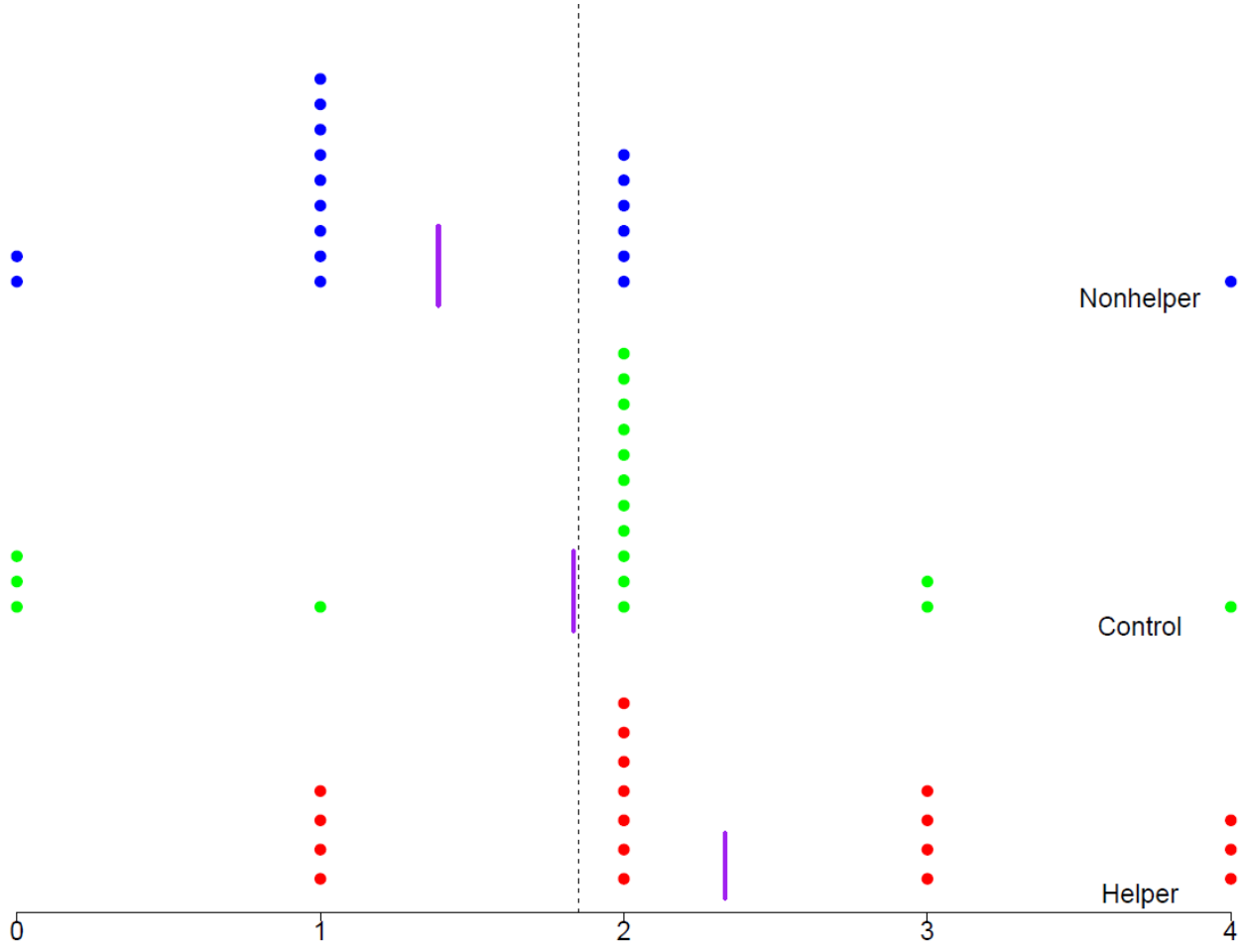


Figure 2: Number of times actor chosen

5. Which **one** of the following statements about this study is **false**?

   (1) The units in this study are the dogs, because data is being recorded about their behaviour.

   (2) This study uses a randomised block design because each dog and owner participated in four identical trials.

   (3) The aim of randomly allocating each dog, along with its owner, was to make comparisons between the groups fair.

   (4) This study is an experiment because each dog, along with its owner, was randomly allocated to a group.

   (5) The response variable is the number of times the dog chose to take the treat from the actor.

6. Which **one** of the following statements about experiments is **false**?

   (1) If blocking is used in the design of an experiment, then nothing is gained from using randomisation in that design.

   (2) Only a well-designed and well-executed experiment can reliably establish causation.

   (3) In experiments on people, we should be aware of the possibility of a placebo effect.

   (4) An observational study can be useful for identifying possible causes of effects.

   (5) Randomisation does not guarantee that the treatment groups are exactly balanced with respect to unknown factors.

7. A researcher wanted to test whether knowledge of letter sounds helps children to learn to read. To do this, a sample of children was divided into two groups. One group consisted of children with no knowledge of letter sounds; they were given training in the sounds of letters. The second group of children had some knowledge of letter sounds; they were not given any special training. The two groups of children were then tested on their ability to learn to read and their scores compared to see if there was a difference. The **main** problem with this design is (select **one** only):

   (1) Insufficient attention is paid to the placebo effect item.

   (2) The children were not randomly assigned to "treatments".

   (3) The experiment is not an observational study.

   (4) The teachers were not "blind" to the treatments.

   (5) Children were not "blind" to the treatment they received.

Question 8 refers to the following information.

One theory regarding memory is that material is remembered as a function of the extent to which it was processed at the time it was initially presented. Eysenck (Eysenck, 1974) was interested in testing this theory. In part of his study, 40 older participants (between 55 and 65 years old) were randomly assigned to one of four learning groups, each of size n = 10. The four learning processes were:

Counting: Simply count the number of letters in each of the words
(Assumed to be the lowest level of processing)

Rhyming: Think of a word that rhymed with each of the words
(Assumed to be the second lowest level of processing)

Adjective: Think of an adjective that could reasonably be used to modify each of the words
(Assumed to be the second highest level of processing)

Imagery: Form a vivid image of each of the words
(Assumed to be the deepest level of processing)

Each participant was given a list of 27 words. The four groups were not told that they would later be asked to recall the words. After the participants had gone through the list of words three times, they were asked to write down all the words they could remember.

A plot displaying the results is shown in Figure 6.

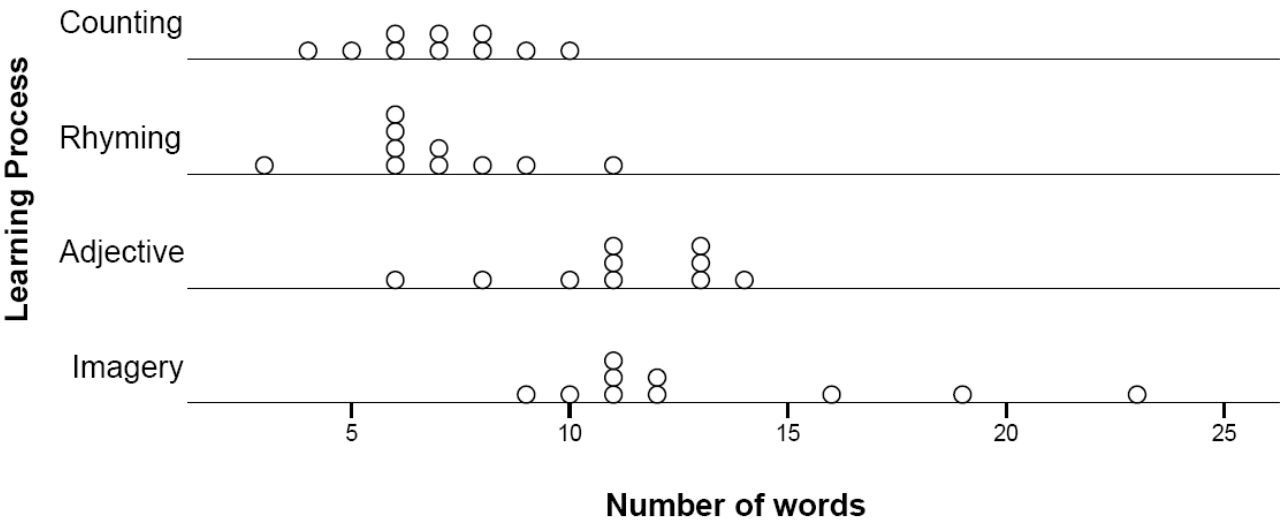


Figure 6: Number of words recalled (out of 27)

8. Which one of the following statements is false?

   (1) Random allocation to groups was needed to ensure that the four groups were all of equal size.

   (2) The response variable was the number of words recalled.

   (3) The treatments were the four types of learning process.

   (4) There was no control group in this experiment.

   (5) This experiment used a completely randomised design.

9. Which one of the following statements about observational studies is false?

   (1) We can only establish causation in an observational study when some form of random sampling has been used.

   (2) In an observational study, the levels of the factors of interest are not randomly assigned to the individuals or units.

   (3) Observational studies are often useful for identifying possible causes of effects.

   (4) A longitudinal study is an observational study that observes the same group of individuals or units over a long period of time.

   (5) In an observational study we compare the responses of the individuals or units that happen to have received each of the levels of the factor of interest.

10. A recent study looked at the type of school a student attends and that student's performance in bursary examinations. It was found that students at single-sex schools perform significantly better in examinations than those students at coeducational schools. The results from this study alone should not be used to argue the case that single-sex schooling, generally, gives rise to better student performance in examinations **mainly** because:

    (1) there may be a difference between male and female performance in examinations.

    (2) there are many more co-educational schools than single-sex schools in New Zealand.

    (3) the designers of this study would not have been able to use any form of blinding.

    (4) the designers of this study have not used a control group.

    (5) the designers of this study did not allocate each student to the school attended.

Question 11 refers to the following information.

It had already been established that increased sexual activity decreases the number of days for which female fruitflies live. Researchers Hanley and Shapiro (1994) designed a study to see if the same were true for male fruitflies. The sexual activity of male fruitflies was manipulated by supplying individual male fruitflies with either one or eight receptive virgin females per day. The lifespan of these males (the number of days they lived for) was compared with the lifespan of males that were supplied daily with one or eight newly inseminated females. Newly inseminated females are not receptive because they will not re-mate for at least two days. There was also a group of males kept with no females.

Thus there were five groups in total and 125 male fruitflies were randomly assigned to one of these five groups. This meant that there were 25 male fruitflies in each group.

The five groups were:

GP1: Male kept alone

GP2: Male supplied daily with 1 newly inseminated female unwilling to mate

GP3: Male supplied daily with 1 receptive virgin female willing to mate

GP4: Male supplied daily with 8 newly inseminated females unwilling to mate

GP5: Male supplied daily with 8 receptive virgin females willing to mate

Dot/box plots of the lifespan for each group are shown in Figure 9:

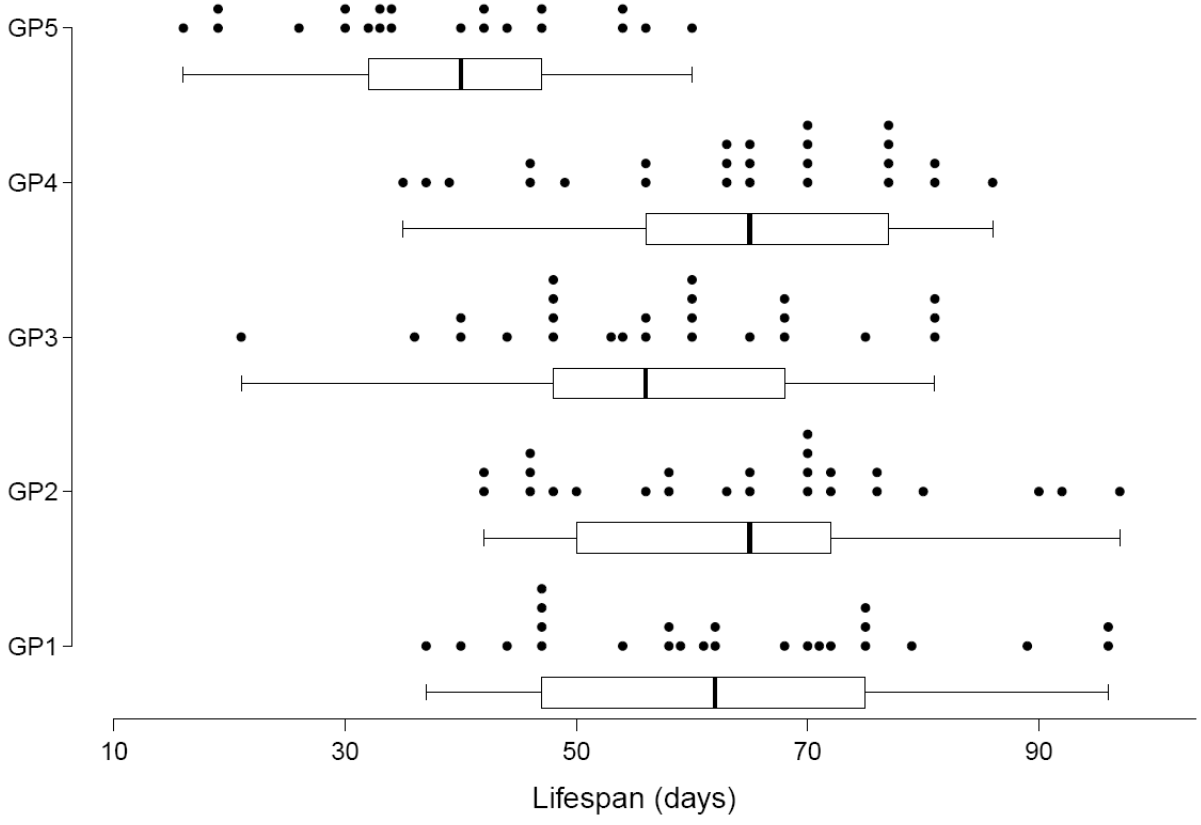


Figure 9: Lifespan of male fruitflies by group

11. Which one of the following statements is false?

    (1) The response variable is lifespan (in days).

    (2) The groups GP1, GP2 and GP4 can be viewed as control groups.

    (3) This study is an experiment with 5 different 'treatment' levels.

    (4) The units in this study are 125 male fruitflies.

    (5) There is blocking in this study design with the female fruitflies blocked as either 'newly inseminated' or 'receptive virgin'.

12. Which **one** of the following statements is **false**?

    (1) Blinding and double blinding are techniques often used by researchers when people are used as experimental units.

    (2) Blocking is used in experiments to ensure fair comparisons with respect to factors the experimenter believes are important.

    (3) In an experiment, the control group always receives no treatment.

    (4) The placebo effect is the response caused in human subjects by the idea that they are being treated.

    (5) Randomisation in experiments allows the calculation of the likely size of sampling errors.

13. The New Zealand Medical Journal has a report on a heart and health investigation that is being carried out by the University of Auckland's Faculty of Medical and Health Sciences. To determine what proportions of the participants were exposed to various risk factors such as smoking, high blood pressure, high cholesterol, etc., each participant was given a medically oriented questionnaire and a physical examination. The aim of the investigation is to follow up what happens to these people in the future. Such an investigation is **best** called a:

    (1) Longitudinal study.

    (2) Cross-sectional study.

    (3) Completely randomised design experiment.

    (4) Double-blind experiment.

    (5) Controlled experiment.

Questions 14 and 15 refer to the following information.

A randomised double-blind study involving 30 elite athletes (male and female) was conducted to assess the effect of taking fish oil supplements on exercise-induced asthma. All of the athletes had experienced symptoms of exercise-induced asthma. The athletes were each randomly assigned to take either a fish oil capsule or a placebo (an olive oil capsule) daily for 3 weeks. For the next 2 weeks the athletes took neither of the capsules. Following this period they then took the other capsule (fish oil or placebo) daily for 3 weeks.

At the end of both 3-week periods each athlete's post-exercise pulmonary (lung) function was assessed by measuring their peak oxygen uptake.

14. Which **one** of the following statements about this study is **false**?

(1) The sample size, $n = 30$, is large enough to allow a valid conclusion to be made about the relationship between taking fish oil supplements and post-exercise pulmonary function.

(2) Randomisation was used to reduce the bias from having all those athletes whose post-exercise pulmonary function naturally improves over time all being given fish oil capsules in the first 3 weeks.

(3) The study is an experiment because of the random assignment of athletes to capsules (fish oil or placebo).

(4) It was not necessary to blind each athlete to the type of capsule received in each 3-week period because they knew they would receive a fish oil capsule during the study.

(5) Results from the study cannot be generalised to apply to all athletes.

15. Which **one** of the following statements about this study is **false**?

(1) Each athlete was part of both the control and treatment groups.

(2) The response is the peak oxygen uptake after the first 3-week period minus the peak oxygen uptake after the second 3-week period.

(3) The groups being compared were based on the capsule being administered (fish oil or placebo).

(4) The athletes selected to participate in the study can be considered to be the study units.

(5) This study did not block on gender.

16. A newspaper reports an Australian scientist's claim that weekdays really are warmer than weekends. Adrian Gordon of the Flinders Institute of Atmospheric and Marine Scientists in Adelaide says, "This is not a greenhouse effect. It is the direct warming effect of fuel burning in factories and vehicles during the week". Gordon's findings follow a study of worldwide satellite data for 5000 days. Sunday, he says, is typically the coldest day, with the temperature rising to a peak on Wednesday and then steadily dropping again.

    This study could **best** be described as:

    (1) an observational study.
    (2) a controlled experiment.
    (3) a double-blind experiment.
    (4) a poll or survey.
    (5) a randomised block design study.

17. Which **one** of the following statements is **false**?

    (1) Blocking is used for experiments to ensure fair comparisons with respect to factors the experimenter knows are important.
    (2) Random sampling errors always have an identifiable cause.
    (3) Experiments on humans should be double blind, if possible.
    (4) The placebo effect is the response in human subjects by the idea that they are being treated.
    (5) Non-sampling errors can be difficult to correct for.

18. It was suggested that the city of Hamilton change its name to Waikato City. TV One ran an opinion poll on its 7pm show. Viewers were asked the question "Should Hamilton retain its name or should it be changed to Waikato City?" and were invited to dial an 0900 number at a cost of $1.49 per minute to record their choice. A total of 4103 votes were received. 77% of these indicated that Hamilton should retain its name and 23% indicated the name should be changed to Waikato City.

    The types of non-sampling error likely to cause the **most** bias in the results of this poll are:

    (1) Behavioural effects and non-response.
    (2) Transferring findings and behavioural effects.
    (3) Selection bias and transferring findings.
    (4) Non-response bias and question effects.
    (5) Self-selection bias and selection bias.

Question 19 refers to the following information.

A study (The Star Project) in the United Kingdom (UK) investigated bullying amongst students at school. As part of the study, a random sample of 192 pupils from one school completed a questionnaire.

One of the questions asked how often they were bullied at school: Often, Rarely, or Never.

19. Which **one** of the following errors would **not** be present if the results of this part of the survey were used to determine information about school pupils in the UK?

    (1) Transfer findings
    (2) Sampling error
    (3) Selection bias
    (4) Behavioural considerations
    (5) Question effects

20. A Statistics Department recently carried out a survey to see if the heating was set at a comfortable level. Of the 60 staff, 6 were surveyed. To take a random sample, the 60 staff names were numbered from 1 to 60. The following numbers were extracted from a table of random digits:

    46400 89413 74507 49374 85137 99211 49557 98263

    By starting at the beginning of the line of random digits and taking consecutive pairs of digits, the first six staff selected would be numbered:

    (1) 46 40 41 50 49 37
    (2) 46 40 08 41 37 45
    (3) 46 13 07 49 17 57
    (4) 46 40 08 13 50 11
    (5) 46 13 07 49 17 13

21. The following is a quotation from a New Zealand paper: “New Zealanders have given the thumbs down to Prime Minister’s call for a republic”. A nationwide telephone opinion poll of 1014 people showed only 28% in favour with 46% against and 26% who don’t care or don’t know.

    Which **one** of the following statements about telephone polls is **false**?

    (1) The true error in the poll is larger than the sampling error.
    (2) Retired people may be over represented.
    (3) Busy people may be less likely to respond.
    (4) The errors quoted for polls such as this in the media include non-response errors.
    (5) There may be different responses from different geographical areas.

Question 22 refers to the following information.

Tilburt et al. (2008) conducted a survey in the United States to investigate attitudes and behaviours regarding the use of placebo treatments by physicians who commonly treat patients with chronic conditions; namely, those specialising in internal medicine (internists) and rheumatology (rheumatologists).

A survey research company was used to randomly select 1200 currently practising internists or rheumatologists from the American Medical Association (AMA) database.

Each selected physician was posted a questionnaire, a $20 incentive, and a letter in which they were informed of the voluntary nature of participation and assured that their identities would not be disclosed to investigators.

Those who did not respond within six weeks were sent a second survey. Overall, 679 responses were received.

22. Which **one** of the following statements is **false**?

(1) The offer of a $20 incentive and the sending of a second survey to those who did not respond within six weeks was an attempt to increase the response rate.

(2) The assurance that respondents' identities would not be disclosed to investigators was an attempt to reduce behavioural considerations and to increase the response rate.

(3) The sample of 1200 is a relatively small proportion of practising internists and rheumatologists in the United States and this creates a potential for selection bias.

(4) There is no potential for self-selection bias because of the random selection of 1200 practising internists and rheumatologists from the AMA database.

(5) There is potential for non-response bias because the response rate was approximately 57%.

23. Which **one** of the following statements is **false**?

(1) Nonresponse can cause bias in surveys because non-respondents often tend to behave differently from people who do respond.

(2) Nonsampling errors are often bigger than the random sampling errors in surveys.

(3) Well-designed experiments use randomisation to avoid subjective and other biases.

(4) People will sometimes answer a question differently for different interviewers.

(5) A well-planned observational study is a reliable method for establishing causation.

Questions 24 and **25** refer to the following information.

The Graduate Longitudinal Study NZ (see https://www.glsnz.org.nz) began in 2011. All New Zealand universities were asked to identify students who were intending to complete a bachelors degree or above in 2011. A randomly selected sample of these students were invited to take part in an online survey about their University experiences and future plans. 65.2% (8719) of those who were asked to participate completed the survey. Some other students started the survey but did not complete it. The first follow-up survey was conducted in 2014.

24. Which one of the following statements is false?

(1) There is potential for question effects in this survey.

(2) There is potential for survey-format effects to have been the reason some people began but did not complete the survey.

(3) There is potential for non-response bias because only 65.2% of those who were selected completed the survey.

(4) There will be no sampling error because the participants were randomly selected.

(5) There will be no interviewer effects because the survey was completed online.

25. Suppose that each of the original participants was assigned a unique ID number from 1 to 8719. Suppose that as part of the 2014 follow-up survey it was decided to randomly select 100 of the original participants to take part in a follow-up in-depth interview. Use the following list of random numbers to select the first five participants. Start at the beginning of the row and use consecutive groups of digits.

14168 86300 55727 25351 00206 07672 84073 01281

The identification numbers of the five chosen teachers are:

(1) 1416 55 7272 5351 20

(2) 1416 8630 5572 2535 20

(3) 1416 5572 7253 5100 2060

(4) 1416 572 7253 5102 607

(5) 1416 8863 55 7272 5351

26. A magazine reports on a long-term study of child development conducted by researchers at the University of Otago. In the study, every child born at Queen Mary Hospital between 1 April 1972 and 31 March 1973 has been monitored at two year intervals until the present day. This study could **best** be described as:

    (1) a sample survey.

    (2) a randomised block design study.

    (3) an observational study.

    (4) a double-blind experiment.

    (5) a cross-sectional study.

Question 27 refers to the following information.

Students enrolled in stage one statistics at the University of Auckland were surveyed regarding their access to, and experience with, computers. The survey was included as a question in an assignment, and students were given marks for completing it (irrespective of the answers they gave). Staff administering the courses wished to use the results of this survey to draw conclusions about future stage one statistics students.

One question asked: 'At the start of the course, how would you describe your Excel experience?'. A total of 918 students answered this question. Each of the 918 answers were classified according to the response given by the student, and the stream the student attended. The results are given in Table 4 below, where 101G, 101 and 108 refer to the various streams.

	Stream			
Response	101G	101	108	**Total**
None	15	36	102	153
Very Little	44	89	119	252
Some	74	150	200	424
Lots	9	29	51	89
Total	142	304	472	918

Table: Responses to question regarding Excel experience.

27. Which **one** of the following is **not** a potential source of error in this analysis?

    (1) Sampling error.

    (2) Question effects.

    (3) Nonresponse bias.

    (4) Interviewer effects.

    (5) Behavioural considerations.

28. A statistician recently surveyed people in a small town on the mathematics they used in the workplace. Because of financial restraints he could only survey 200 people. To take a random sample he used a list of 661 workplaces in the town (numbered 1…661). The following numbers were extracted from the table of random digits:

    03521 10933 34889 74209 31391 62728 99671 77720 13726

    If he started sampling at the beginning of the line of random digits and took consecutive digits, the first five workplaces selected would be numbered:

    (1) 35 211 93 334 139
    (2) 35 93 139 162 177
    (3) 35 109 348 313 137
    (4) 35 93 93 139 162
    (5) 35 211 93 334 93

29. On the anniversary of the (alleged) death of Elvis Presley, a Dallas record company sponsored a national telephone call-in survey. Listeners of over 1000 American radio stations were asked to phone-in (at a charge of $2.50 per call) to voice an opinion on whether or not Elvis was really dead. It turned out that 56% of the callers felt that Elvis was still alive.

    This is an unreliable estimate of the percentage of all Americans who believed that Elvis was still alive at that time **mainly** because:

    (1) of biases from question effects and survey format.
    (2) of bias from transferring findings.
    (3) a pilot study should have been conducted to find out which radio stations should have been included.
    (4) of sampling errors.
    (5) of selection and self-selection biases.

30. Which **one** of the following statements is **false**?

    (1) Slight changes in the wording of questions can make a measurable difference to the results of a survey.
    (2) There are always statistical procedures available to correct results (at the completion of a survey) when the population from which a sample is taken is different from the population of interest.
    (3) Bias can occur when too many respondents in a survey give an answer which does not reflect their actual behaviour.
    (4) The outcome of a survey which uses personal interviews may be different from the outcome of the same survey if telephone interviews had been used.
    (5) The outcome of a survey may be affected by the race and/or gender of the interviewer.

Question 31 is about the following information.

Data on 56 hospital births from a single week at the Wellington hospital were collected. Researchers wanted to examine the relationship between various characteristics of the mother and the Apgar Score for the baby. This is a score given to the baby in the first minute after birth and measures the overall physical appearance of the baby.

The characteristics that were measured include:

Age	Age of mother in years.
Mass	Mass of baby in grams.
Gravida	Number of pregnancies including this.
Para	Number of births including this.
Term	Time to delivery in weeks.
Apgar	Apgar score out of 10.

Age	Mass	Gravida	Para	Term	Sex	Apgar
18	3850	1	1	40	M	9
22	2590	1	1	34	M	7
23	3500	1	1	41	M	8
29	2850	2	2	36	F	9
26	3480	3	3	41	M	9
28	3210	2	1	40	F	6
30	3310	8	4	39	F	9
30	4220	3	3	42	M	9
27	4400	2	2	41	M	9
27	2900	1	1	40	F	9

Table: Ten observations from the Apgar Study

31. The **main** problem with using this data to draw conclusions about Apgar scores in New Zealand is that:
    (1) not all pregnancies resulted in births
    (2) there was an interviewer effect
    (3) there was no control group
    (4) there was selection bias
    (5) the data contain sensitive personal information

32. An issue of *Consumer* contains figures on the reliability of cars driven in New Zealand. The figures were obtained from a sample of 15,372 *Consumer* readers who responded to a mailed questionnaire.

    The **main** problem with using the results to draw conclusions about the reliability of all cars driven in New Zealand is:
    (1) there was no control group
    (2) there was insufficient attention to the placebo effect
    (3) selection bias
    (4) question effects
    (5) interviewer bias

Questions 33 and **34** are about the following information.

Information on 5387 school children in Caithness, Scotland, was collected by sending out questionnaires to randomly selected schools. In each school, every child was asked to take the questionnaire home and fill it out. Information gathered on the children included their eye and hair colour, and their favourite pastimes.

33. This study is **best** described as:

    (1) a longitudinal study.

    (2) a cross-sectional study.

    (3) a census.

    (4) a completely randomised design study.

    (5) a randomised block design study.

34. Which **one** of the following is **not** a potential source of error in this study?

    (1) Interviewer effects.

    (2) Behavioural considerations.

    (3) Nonresponse bias.

    (4) Question effects.

    (5) Sampling error.

35. All 10 to 14-year-olds at a Hamilton Intermediate School took part in a study of adolescent attitudes towards the television programme *Shortland Street*. Suppose that the number of students involved was 200. Four of the students interviewed did not watch the programme. The study found that 69% of the children felt that the programme dealt with 'real-life' issues, and 23% reported instances where they had learned or been guided in their own lives by stories from the *Street*.

    These results should not be taken as reflecting the opinions of the general population of New Zealand school children **mainly** because (select **one** only):

    (1) of the "placebo" effect.

    (2) of interviewer effects.

    (3) of non-response bias.

    (4) four of the children surveyed did not watch the programme.

    (5) of selection bias.

Question 36 refers to the following information.

Sports Foundation grants for sports which won the right to represent New Zealand at the Sydney Olympics are shown in the table below.

No.	Sport	1998–1999	1999–2000	2000–2001
01	Archery	$16,500	$15,000	$25,000
02	Athletics	$515,340	$485,400	$299,000
03	Basketball	$133,100	$90,000	$40,000
04	Boxing	$166,950	$55,000	$44,350
05	Cycling	$678,500	$747,182	$688,140
06	Equestrian	$691,000	$717,000	$558,620
07	Gymnastics	$94,500	$34,500	$22,400
08	Hockey	$498,500	$478,460	$554,000
09	Judo	$153,650	$93,179	$124,500
10	Rowing	$533,100	$466,700	$707,265
11	Shooting	$327,000	$106,000	$405,616
12	Softball	$251,913	$425,259	$254,542
13	Swimming	$431,470	$205,000	$280,594
14	Table Tennis	$26,250	$3,000	$29,000
15	Triathlon	$343,110	$548,255	$86,300
16	Weightlifting	$98,900	$48,125	$79,500
17	Wrestling	$13,520	$8,000	$15,000
18	Yachting	$947,000	$1,131,000	$622,356

Table: Sports Foundation Grants

36. Suppose we wish to randomly select five of the sports listed in the table on above. The method for randomly selecting the five sports uses the number (in the **No.** column of the table above) associated with each sport and random number digits. Use the row of random digits below to select a simple random sample of five sports. You must start at the beginning of the row and use consecutive pairs of digits.

09874 11018 39090 54804 17130

The five sports selected are:

(1) Judo, Rowing, Yachting, Judo, Cycling
(2) Judo, Shooting, Archery, Wrestling, Swimming
(3) Judo, Shooting, Archery, Cycling, Wrestling
(4) Judo, Rowing, Yachting, Cycling, Boxing
(5) Judo, Hockey, Gymnastics, Boxing, Archery

Questions 37 to **39** refer to the following information.

Last year an Auckland Statistics student was required to survey Statistics students on transport as part of a Diploma project. One aim of the survey was to test whether changing the wording of a question produced a different level of response. One such question was posed in the following two versions:

Q1 (1) *Taking into account the problems and cost associated with parking, do you think that the cost of catching a bus to university is too high? (Answer: Yes/No/Don't know.)*

Q1 (2) *Do you think that the cost of catching a bus to university is too high? (Answer: Yes/No/Don't know.)*

Two questionnaires were prepared, but each questionnaire only included one of the above versions. The two questionnaires were then distributed to two groups of students - one questionnaire to each group.

37. A pilot survey should have been used:

(1) Before the full survey to try to determine the precision and bias of the sample estimates.

(2) Before the full survey to try to get good estimates of the proportions and thereby avoiding the cost of having to do a full survey.

(3) After the full survey to make sure that it is an experiment rather than an observational study.

(4) After the full survey to try to determine the precision and bias of the sample estimates.

(5) Before the full survey to try to identify any problems with the questionnaires.

38. Suppose we decide to conduct a pilot survey involving 10 students; 5 students receiving version (i) and 5 receiving version (ii). The first 5 students are to be selected by taking a random sample without replacement from a class of 75 students. To do this, the names of the 75 students are numbered from 1 to 75. The following numbers from our table of random digits are used to select the students:

38905 93159 25252 29004 54972 73607 49557 98263

By starting sampling at the beginning of the above line of random digits and taking consecutive pairs of digits, the first 5 students selected would be numbered:

(1) 38 59 31 59 25

(2) 38 59 31 25 22

(3) 38 15 25 29 00

(4) 38 05 31 59 25

(5) 38 15 25 29 54

39. The two questionnaires are to be distributed to all first and second year Statistics classes. Which one of the following designs will result in the **fairest** allocation of the two versions?

    (1) Take copies of each version to every class and distribute version (i) to the front half of the lecture room and version (ii) to the back half of the lecture room.

    (2) Make sure that version (i) is distributed to all the first year classes and that version (ii) is distributed to all second year classes.

    (3) Take copies of each version to every class and distribute version (i) to all females in the class and version (ii) to all the males in the class.

    (4) Take copies of each version to every class and distribute version (i) to those students who travel to university by bus and version (ii) to the rest of the students.

    (5) Have the pile of questionnaires alternating (i), (ii), (i), (ii) etc., and distribute the pile one row at a time in each class so that about every second student in each class receives version (ii).

Question 40 refers to the following information.

The table below shows the total amount of debt owed by each of a university department's fourth year students. The amount of debt owed by each of the 20 students is given to the nearest hundred dollars. The students are ordered by the size of the debt owed. Each student has been assigned an identification number between 1 and 20.

Student	1	2	3	4	5	6	7	8	9	10
Debts ($00's)	0	0	23	37	49	114	124	125	125	127

Student	11	12	13	14	15	16	17	18	19	20
Debts ($00's)	131	151	193	234	260	267	273	281	342	491

Table: Student debt in hundreds of dollars

40. Forty random digits are given below. Select a simple random sample of five (5) students from the population of 20 students by using these random digits. You must start at the beginning of the line of random digits and take consecutive pairs of digits.

    37182 91306 13472 81283 04972 59607 16557 98144

    The students in the sample are:

    (1) 18 29 13 6 13

    (2) 18 13 6 13 12

    (3) 18 2 9 13 6

    (4) 18 13 6 12 4

    (5) 18 29 13 6 4

41. At around midnight on the 5th March 1997, the Deputy Prime Minister of New Zealand, the Right Honourable Mr. Winston Peters, allegedly assaulted the MP for Whangarei, the Right Honourable Mr. John Banks. The following evening TV3 ran an opinion poll on the question "Is Mr. Peters fit to be Deputy Prime Minister of New Zealand?". Viewers were invited to dial an 0900 number (at a cost of 99c/min) and record either "yes" or "no". 10% of calls said Mr. Peters was fit to be Deputy Prime Minister while 90% said Mr. Peters was not fit to be Deputy Prime Minister.

    The type of non-sampling errors likely to bias the results the **most** are:

    (1) Behavioural effects and nonresponse bias.

    (2) Self-selection bias and selection bias.

    (3) Nonresponse bias and question effects.

    (4) Transferring findings and behavioural effects.

    (5) Selection bias and transferring findings.

42. Which **one** of the following statements about experiments and observational studies is **false**?

    (1) A group of experimental units given no active treatment is called a control group.

    (2) If we know that the gender of a person will make a difference to the outcome of an experiment, we should block by gender and give the females one treatment and the males the other treatment.

    (3) A study in which we simply compare units that happen to have received each level of the factor of interest is called an observational study.

    (4) Although an observational study cannot reliably establish causation, it can be useful for identifying possible causes of effects.

    (5) A study which observes the same group of individuals or units over a long period of time is called a longitudinal study.

43. Which **one** of the following statements is **false**?

    (1) A well designed, and well executed experiment can reliably establish causation.

    (2) Results obtained from subjects treated with a placebo won't always be equivalent to results obtained from comparable subjects receiving no treatment.

(3) Non-response in surveys can cause bias because non-respondents often tend to behave differently from people who do respond.

(4) Random sampling errors always have an identifiable cause.

(5) Observational studies are not reliable for proving causation.

ANSWERS

1.	**(1)**	2.	**(2)**	3.	**(5)**	4.	**(4)**	5.	**(2)**	6.	**(1)**
7.	**(2)**	8.	**(1)**	9.	**(1)**	10.	**(5)**	11.	**(5)**	12.	**(3)**
13.	**(1)**	14.	**(4)**	15.	**(2)**	16.	**(1)**	17.	**(2)**	18.	**(5)**
19.	**(1)**	20.	**(4)**	21.	**(4)**	22.	**(3)**	23.	**(5)**	24.	**(4)**
25.	**(1)**	26.	**(3)**	27.	**(4)**	28.	**(1)**	29.	**(5)**	30.	**(2)**
31.	**(4)**	32.	**(3)**	33.	**(2)**	34.	**(1)**	35.	**(5)**	36.	**(4)**
37.	**(5)**	38.	**(2)**	39.	**(5)**	40.	**(4)**	41.	**(2)**	42.	**(2)**
43.	**(4)**										

WHAT SHOULD I DO NEXT?

Once you've had a go at all of the problems in the handout (check out Leila's scanned slides at www.tinyURL.com/stats-OSE for her additional handwritten notes and workings for every problem), you could:

- Go through the Chapter 2 and 4 blue pages. In each chapter, this includes:
  - **additional notes**, all of which are examinable!
  - a **glossary**
  - **a challenge!** (true/false statements)
  - **Sample Exam / Test** (multi-choice) **Questions**
  - **tutorial** material
- Try Chapter 2 & 4 questions from three of the past five tests – use the *Test questions index* document to identify them! (You can find this useful resource along with the past five tests on Canvas via ***Modules → Past Tests and Exams (with answers)***)
- Make sure you pick up your Assignment 1 answers from the Science Student Resource Centre and check them against the model answers on Canvas. (The Chapter 2 material covered in the first half of this workshop was assessed in Question 5 of Assignment 1.)
- Do Assignment 2 (The Chapter 4 material covered in the second half of this workshop will be assessed in Question 2 of Assignment 2.)