# Statistics for

## Scholarship and Excellence

### Nicola Ward Petty (Dr Nic)

# Dr Nic

Dr Nicola Ward Petty BSc(Hons), DipTchg, PhD

Lecturer in Operations Research at University of Canterbury for 20 years.

Creator of Statistics Learning Centre resources and videos.

Loves to teach.

# Main Ideas

- Skills needed and attitudes

- Developing fluency

- Answering questions

# Scholarship skills, statistical insight

**What skills are needed to get a scholarship?**

- Talk to the person next to you and come up with three skills.

**What does a teacher need to do in order to help a student develop these skills?**

- Discuss

# Skills

| | | |
|---|---|---|
| Use statistical language correctly | Relate to the PPDAC cycle | Evaluate assumptions of a model |
| Understand the question | Interpret graphs | Pose new questions |
| Basic concepts, such as trend, seasonality, spread, shape, expected value | Identify erroneous thinking | Explain the statistical result, relating it to the context |
| Commit to an answer | Apply an appropriate model | Calculate numeric answers |
| Understand the context | Justify choices | Decide between alternatives |

# Basic Skills

| | | |
|---|---|---|
| Use statistical language correctly | Relate to the PPDAC cycle | Evaluate assumptions of a model |
| Understand the question | Interpret graphs | Pose new questions |
| Basic concepts, such as trend, seasonality, spread, shape, expected value | Identify erroneous thinking | Explain the statistical result, relating it to the context |
| Commit to an answer | Apply an appropriate model | Calculate numeric answers |
| Understand the context | Justify choices | Decide between alternatives |

# Contextual

| | | |
|---|---|---|
| Use statistical language correctly | Relate to the PPDAC cycle | Evaluate assumptions of a model |
| Understand the question | Interpret graphs | Pose new questions |
| Basic concepts, such as trend, seasonality, spread, shape, expected value | Identify erroneous thinking | Explain the statistical result, relating it to the context |
| Commit to an answer | Apply an appropriate model | Calculate numeric answers |
| Understand the context | Justify choices | Decide between alternatives |

# Process

| | | |
|---|---|---|
| Use statistical language correctly | Relate to the PPDAC cycle | Evaluate assumptions of a model |
| Understand the question | Interpret graphs | **Pose new questions** |
| Basic concepts, such as trend, seasonality, spread, shape, expected value | **Identify erroneous thinking** | Explain the statistical result, relating it to the context |
| **Commit to an answer** | Apply an appropriate model | Calculate numeric answers |
| Understand the context | **Justify choices** | **Decide between alternatives** |

# Attitudes

**What attitudes will help scholarship candidates (and everyone really!)**

- Approach new situations with courage

- Try different approaches

- Growth mindset not fixed mindset - Read Dweck, Boaler

# Discussion

How important are the elements in the basic group?

What are the implications for teaching to Excellence and Scholarship?

# Dr Nic says:

**Excellence is not "going one step further".**

**Statistical insight can only be displayed if statistical principles are understood deeply across multiple contexts.**

# Implications for teaching

**Students need to be fluent in language**

- Must know the basics, similar to the way basic number facts fluency helps in algebra.

- Can't be worried about vocabulary

  - T/F questions, discussion, writing, reading

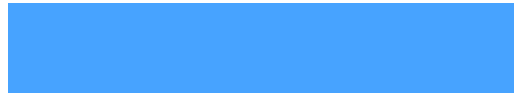**Multiple <span style="color:red">different</span> contexts are essential**

- Must be able to tell the general from the specific

- Practice linking results and context

  - Not just pattern repetition

# Example of multiple contexts

**This is an example of a plog.**

**Is this a plog?**

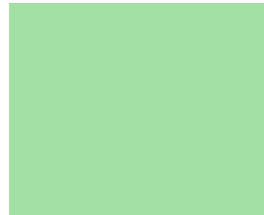# Example of multiple contexts (2)

**These are examples of plogs:**

**Is this a plog?**

# Example of multiple contexts (3)

**These are examples of plogs:**

**Is this a plog?**

# Plan for success

Scholarship exam is like sport or musical or any other performance.

**Three phases**

- Underlying fitness (long term)

  - Basic material – know it automatically

    **1. Know stuff**

- Skill development (medium term)

  - Practice answering questions, and thinking like a statistician

    **2. Think like a statistician**

- Strategy (short term)

  **3. Show to the examiner that you know stuff and think like a statistician.**

  - Memorise checklists, be fit, make sensible choices during the exam.
    Understand "the rules of the game".

# Time

**Time engaged in a subject is one of the best predictors of success.**

**Students need something useful (and fun) to do with their time.**

**Not just "read notes"**

# Quick Quiz Questions(1)
# Answer true or false

1. It is possible to avoid sampling error, by using a simple random sample.

2. You cannot get seasonality in annual data.

3. The strength of a bivariate relationship is shown by how close the points are to the fitted line.

4. Understanding the context is an important part of a statistical study.

5. Sampling error can cause bias in a result.

# **Quick Quiz Answers (1)**

1. It is possible to avoid sampling error, by using a simple random sample.     **False**

2. You cannot get seasonality in annual data.     **True**

3. The strength of a bivariate relationship is shown by how close the points are to the fitted line.     **True**

4. Understanding the context is an important part of a statistical study.     **True**

5. Sampling error can cause bias in a result.

   **False**

# Quick Quiz Questions(2)

6. Inference is when we draw conclusions about a sample from information in the population.

7. Often the triangular distribution can be used instead of the normal distribution to approximate a real-life distribution.

8. You cannot draw conclusions of causation from a randomised experiment.

9. The probability for something not to happen is always 1 minus the probability for it to happen.

10. When you use dice and coins, you can calculate the true probability of an outcome.

# Quick Quiz Answers (2)

6. Inference is when we draw conclusions about a sample from information in the population.

   **False**

7. Often the triangular distribution can be used instead of the normal distribution to approximate a real-life distribution.

   **True**

8. You cannot draw conclusions of causation from a randomised experiment.

   **False**

9. The probability for something not to happen is always 1 minus the probability for it to happen.

   **True**

10. When you use dice and coins, you can calculate the true probability of an outcome.

    **False**

# Ideas

**How could you use this in the classroom?**

**Make their own? Develop a "box" of T/F questions, including definitions.**

**Statistics Learning Centre**

# Writing to a higher standard

**What identifies an answer as Excellent or Outstanding?**

# Scenario taken from the Stuff website on 18 August 2014.

*Nationally, the latest poll put National at 55.1 per cent of the party vote, virtually unchanged from last month. Labour is on 22.5 per cent and the Greens on 11.3 per cent.*

The methodology used by Ipsos is given as: *The poll was conducted by telephone between August 9 and August 13 2014 with a total sample of 1007 people. It has a maximum margin of error of plus or minus 3.1 per cent. Party Vote percentages are only based on eligible and decided voters. Respondents who are undecided, do not intend to vote or are ineligible to do so are excluded from the Party Vote.*

Ipsos states: *Interviewers ask to speak to the person in the household who is aged 18 years or over and is the next to have their birthday. If the person is not immediately available we call back repeatedly giving them every chance of participating.*

## Questions

1. What does it mean to say that the maximum margin of error is plus or minus 3.1?

2. A history teacher tells you that the 3.1 margin of error takes care of errors that might cause bias in the results. Explain what is wrong with this statement, and give a better explanation of the relationship between margin of error and bias.

3. Explain why the interviewer will call back repeatedly.

# Pre-election polling/ answer types

1. Read the scenario and think about the questions.

2. Discussion – questions assigned.

3. You write 2 different answers to the same question.
   Work in 2s.
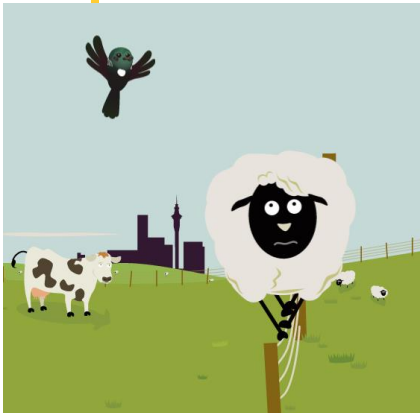
**For one of the questions, write two answers:**

**1 correct, but not really scholarship level**

**1 outstandingly good**

## Q 1 Answers (from Waikato Schol day)

What does it mean to say that the maximum margin of error is plus or minus 3.1?

As the sample size is 1007 when we use the rule of thum $1/\sqrt{n}$ we get the answer of 3.1. This means that there is a 95% chance nationals (NZ) percentage support will be 3.1% higher or lower than the result shown.

Margins of error are calculated using the standard rule of thumb, being MoE $= 1/\sqrt{n}$ where n is population size. The population size (n) is 1007 giving a MoE of 3.1%. This means that a parties poll percentage can vary from 3.1% of the original percentage. It is important to recognise that the poll percentages could differ more from the population than 3.1% however we can safely estimate the actual percentage to be within 3.1% of poll percentages.

**Q 2** A history teacher tells you that the 3.1 margin of error takes care of errors that might cause bias in the results. Explain what is wrong with this statement, and give a better explanation of the relationship between margin of error and bias.

The teacher is wrong as the margin of error only relates to sampling error – it deals with the fact that the poll is only a sample of the population. Bias is caused by non-sampling error and can't be taken care of by margin of error as it has nothing to do with it.

If we take the entire population there will be no bias as the whole population has been picked. If we pick a sample that is perfectly random then there will be no bias as we don't know who we are picking and everyone is equally likely to be picked. This is not perfectly random as the sample was taken over the phone. Not everyone has a phone for example students who go flatting and are trying to save money. The history teacher's statement is wrong because margin of error only takes into account people who were missed.

## Q 3 Explain why the interviewer will call back repeatedly.

They call back to reduce bias. Otherwise, the people who stay at home and those who are more eager to answer the phone are far more likely to have their opinions taken. The people who go out more are more exposed to signboard advertising, so can be influenced more. This adds a great level of bias to the results

To ensure that they collect sufficient data to carry out the poll. If the interviewer did not call back repeatedly then they would not collect data from that household which would result in insufficient distribution of results.

The interviewer will call back repeatedly so it ensures that the sample is random and potential bias is eliminated.

# Teaching implications

**Show how to "craft" answers**

**Maybe use OSEM to encourage context**

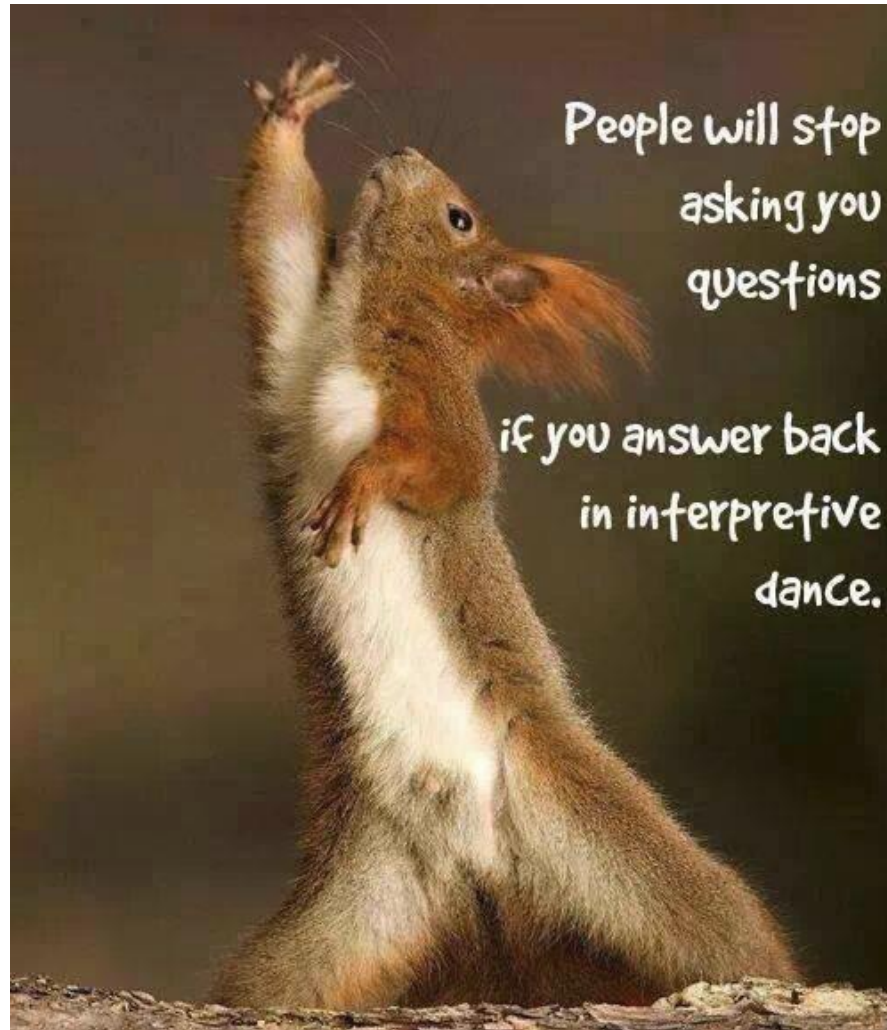**Work in pairs to write answers – communication and using language is part of fluency.**

# **What have we done today?**

■ Statistical insight can only be displayed if statistical principles are understood deeply across multiple contexts.

■ Courage is important – growth mindset

■ Need underlying "fitness" – Quick T/F quizzes.

■ Get students to practise writing answers and identifying levels.

And…

■ StatsLC resources can help with Scholarship.

# Questions and comments

# Scholarship Comparison for 2013

| Stats | Calc |
|---|---|
| **Cohort 12606** | **8392** |
| **Entries 1462** | **1421** |
| **Assessed 1065** | **1086** |
| **Scholarships 370 (34.7%)** | **212 (20%)** |
| **Outstanding 36 (3.4%)** | **31 (3%)** |
| **Grade range 22-30** | **21-31** |