Developing Statistical Thinking (ME626)

This module focuses on statistics and probability, and how people learn and develop statistical thinking.

Statistics is an important branch of mathematics which helps us understand about the collection, organisation, analysis, interpretation and presentation of data. It also involves the study of probability, the measure of likelihood that an event will occur. The learning of statistics enables learners to explore relationships and ask questions about the world around us. As learners develop their understanding of statistics, they are able to interrogate data sets to form conclusions and apply statistical thinking to question the sources and validity of claims made, such as in newspapers and marketing campaigns.

The main study resource is the module course book, with extra resources on the module webpage. Ideas, such as the PCAI framework for statistical investigations (Pose, Collect, Analyse, Interpret), Summarising and Visual Representations of data, will be important for the way you think about and reflect on your learning whilst studying ME626. As you study you will be directed to ‘have a go’ at the activities in the book and to reflect on yourself as a learner whilst undertaking the activities.

Why not have a go at the activity below?

Average student

Open the attached Excel spreadsheet “Census at school 2009 subset”. This data set contains information submitted by 11-year-old school students. (You can find this complete data set, and other data sets, online at: http://new.censusatschool.org.nz/)

Spend some time looking at the data presented in the spreadsheet.

What questions do you have?

Posing questions is the first stage of a statistical investigation cycle using the PCAI framework.

Here are some examples of the sorts of questions you might: Are boys taller than girls? Do taller students have larger arm-spans? Is there a relationship between foot and wrist size?
Choose which question you are going to investigate. You can return to your other questions later if you want to.

**Collecting data and Sampling**

In this task, the data has been collected for you. This means the data set is secondary. Information you have collected yourself is called primary data.

Since there is a large amount of data provided, you will probably need to take a sample of the data in order to enable you to investigate your question.

First, you should decide how many students you want in your sample. If you are going to calculate statistical summaries and draw graphs by hand, you might want to limit your sample size to about 20 students. If you are going to use formulae in Excel, you might prefer a larger sample size, of say 100 students.

Next you need to decide how you will sample the data. For example, you could take a *convenience sample* of the first 50 students in the list. Or you might like to ensure you have an equal number of male and female students in your sample. Alternatively, you might like to use *random sampling* using a random number generator.

A random sample of 20 students has been provided for you to use. Click on the second tab of the Excel spreadsheet to use this sample.

**Summarising and presenting the data**

Depending on the question you have posed, you will need to calculate some summary statistics. This could involve finding the average (mean, median and mode) and measures of spread, such as the range or interquartile range.

If comparing bi-variate data, such as the relationship between arm-span and height, you may like to plot scatter graph to see if there is any correlation between the two variables. If you are comparing the heights of boys and girls, you might like to draw two box plots to compare the distribution of heights for each gender.

You should choose which data summaries and graphs are most appropriate for your investigation. Try to reason why you have chosen the graph or summary data for your investigation.

You will learn more about the advantages and disadvantages of different summaries and visual representations of data when you study ME626.

If you need some help with how to calculate averages or plot graphs, there is lots of free support on Open Learn.

**Analysing the data**

Now that you have summarised your data using averages or represented the data visually using a graph or diagram, you need to analyse the data. This means identifying any trends or highlighting which group has a higher average or wider measure of spread.
**Interpreting the data**

From your analysis you should be able to draw some conclusions by relating back to the context of the data and returning to your question posed. For example, if you have found a positive correlation between height and arm-span you may draw a conclusion that *the taller a student is, the larger their arm-span*. If you have found that the mean height was higher for girls than boys in your sample, you might conclude that *11-year-old girls are taller on average than 11-year-old boys*.

It is important to relate your analysis to the context in order that your claims are valid. It is also important to realise that a sample of 20 students does not necessarily accurately represent the wider community of students. Realising the limitations of your conclusions is a key part of thinking statistically.

**Continuing the PCAI cycle**

Your investigation may have sparked further questions you would like to investigate or you might like to take a second sample of students in order to validate your claims.

The PCAI framework is cyclical so investigations can be evaluated and continued, with modified questions.

**Reflections**

When you were first looked at the data set, what were you drawn to? What questions did you pose and how did you choose which question to investigate?

In the collecting stage of the PCAI framework, how did you choose to sample the data set? Why did you choose this sampling method for your investigation? Did you think about how big your sample should be?

When analysing the data, how did you decide which summary measures to calculate? Which techniques and methods did you use? Can you identify any advantages or disadvantages of using your chosen measures?

How did you choose to present the data? Did you use tables, graphs or charts? Why did you choose to represent the data in this way? What did these visual representations of data allow you to see that the original list of data did not?

When you interpreted the data, what conclusions did you form? Were you able to answer the question you originally posed?

Did you identify any limitations from your investigation? What would you do next if you were to continue this investigation to carry out another cycle of the PCAI framework?

You may reflect that you used Summarising or Visual Representations, and the PCAI cycle, whilst you investigated the average student task.
Using this task with learners

This task can be used with learners. If you choose to use this task with school-aged learners, you might ask them to collect their own primary data by carrying out a survey at school. If working with adults or more experienced learners, you might like to give them the full data set from the Census at schools’ website.

Tutorials and assessment

Students on ME626 have regular contact with their tutors, usually by phone, email or Skype. In addition, online tutorials are held regularly where module ideas are discussed and students can ask for advice on the assessments.

There are three tutor marked assessments plus one final assessment. These are typically based on activities which you undertook during your studies.

For example, an assessment based on the Average student investigation might look like this:

**Question 1 (maximum 1000 words) 50 marks**

This question is about you as a learner.

Explain your own work on the average student investigation. Include annotated samples of your work as supporting evidence.

You need to include the following:

a) An account of what you did, how you did it and why you did it that way.
b) An explanation with evidence of how you used the PCAI cycle framework and at least one other idea from the module.
c) Your personal reflection on your role as a learner making use of any ME626 module ideas. This section of the assessment gives you the opportunity to reflect on how your engagement with the tasks affected/enhanced your learning and to comment on any learning issues that arose.

**Question 2 (maximum 1000 words) 50 marks**

This is about you as the teacher. Include examples of your learners’ work as supporting evidence.

a. Adapt the average student investigation you worked on in Question 1 to use with your learners. Explain how, and where, you expect your learners to make use of the PCAI cycle when they are working on the task.

b. Try the task with your learner(s). Explain what actually happened, including how they used the PCAI cycle, and possibly other module ideas from the module, as they engaged with the task. These should be annotated with ME626 module ideas, in particular the PCAI cycle. Review the session and explain where you noticed that your learners were using ME626 module ideas, in particular the PCAI cycle.