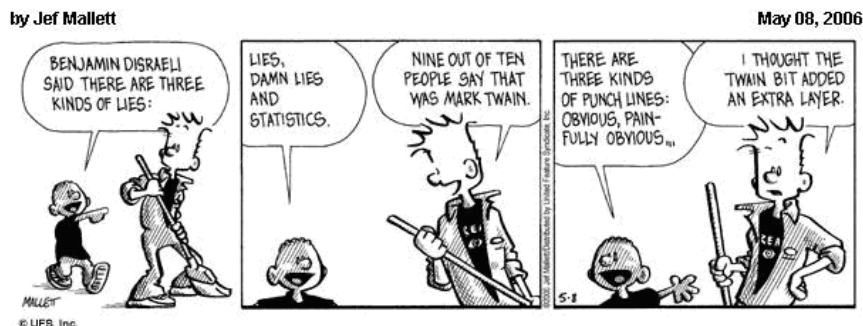


NCEA Level 2 Mathematics

19. The Statistical Enquiry Process

Statistics is one of the most important tools in experimental science, and medicine. Even in the everyday world we are bombarded with statistics given to us by the media, politicians, and the internet. Because of this, being able to interpret statistics and judging whether or not they support a given argument, result, or point of view is an important skill in the modern world.



The statistical inquiry cycle is something you should already know about from as far back as intermediate school. Here it is, in a slightly expanded form:

1. We begin with a question we want to answer, and perhaps a hypothesis backed up by some theory.
2. We decide what data we need to collect to answer our question.
3. We plan how we will collect our data, and how we will ensure it is accurate and precise.
4. We plan how we will process our data.
5. We collect our data. (This is intentionally *after* we plan how we will use it.)
6. We process our data in accordance with our plan.
7. We answer our question.
8. We decide how reliable our process was.
9. We write a report, detailing *all* of the above steps, such that another person could pick it up and try to replicate our findings.

Writing a question, and deciding on data to collect

A question should be written clearly and precisely, and there should be a straightforward way to answer it.

Some good questions:

- Are high school students in Wellington City more likely to cycle to school than students in Upper Hutt?
- Do male Y12 students tend to be taller than female Y12 students?
- Does this drug lower the risk of heart failure after a stroke?
- With what speed does a 0.5 kg weight hit the ground after being dropped 10 m?

Some bad questions (why?):

- Does this drug work?
- Will I crash if I drink and drive?
- Does the average person support the attempt by the USA to bring freedom and democracy to other places in the world?
- Any question you write after you've already collected your data!*

When you write your question, you should come up with a good idea as to what kind of data you will need to gather to answer it. This could be something like a set of measurements, or responses to a questionnaire, or something else.

Planning how to collect data

Usually, we want to collect data to answer a question about a whole population. Unfortunately, it's not cheap or easy to measure or survey every individual in a population, so we have to make do with examining a smaller sample; we then need to process our data to make an educated inference about the state of the full population. In order to do this, we need a reliable sample that reflects the trends in the population.

In general, we need two things from a sample:

- A sample needs to be *large*.
- A sample needs to be *random*.

Both of these statements are harder to unpack than they might seem right now:- they are what we will discuss in the next topic.

At this stage, we also decide how we are going to actually collect the data: will we need to write a questionnaire? Do we need any specialised equipment? How are we going to ensure the accuracy and precision of our data? Do we need a control group?

Planning how to process data

We then need to decide what we are going to do to our massive spreadsheet of data after we collect it. If we want to work out the average value of some measurement, we need to decide how to do this (median, or mean, or something else?). If we want to work out whether there is a difference in some value between two populations (whether one group is taller than another group, for example), then we need to decide how big a difference will be 'significant' and not likely to be due to random chance.

Collecting and processing the data

This is where we apply the plan we've come up with so far. If we deviate from the plan in any way, we carefully specify why and how we are doing so, and we make a careful note that this needs to be stated in our report. Further, we *only* perform the analysis we planned to do. (This is for two reasons: firstly, as I mentioned above, if you measure a bunch of variables then some of them will end up significant simply due to random chance; and secondly, the data we planned to collect is more likely to be reliable than the data we collected along the way with no planning.)

*Try googling 'data dredging': "Data dredging (also data fishing, data snooping, data butchery, and p-hacking) is the misuse of data analysis to find patterns in data that can be presented as statistically significant when in fact there is no real underlying effect. This is done by performing many statistical tests on the data and only paying attention to those that come back with significant results, instead of stating a single hypothesis about an underlying effect before the analysis and then conducting a single test for it.

"The process of data dredging involves automatically testing huge numbers of hypotheses about a single data set by exhaustively searching—perhaps for combinations of variables that might show a correlation, and perhaps for groups of cases or observations that show differences in their mean or in their breakdown by some other variable." (Wikipedia contributors. (2019, January 13). Data dredging. In Wikipedia, The Free Encyclopedia. Retrieved 22:29, January 20, 2019, from https://en.wikipedia.org/w/index.php?title=Data_dredging&oldid=878260765) See also <https://xkcd.com/882/>.

Answering the question

Now, we look at our data processing and we give a simple answer to our question together with a rough idea of how likely we are to be correct:

According to our analysis, our hypothesis was incorrect: in our sample, girls tended to be taller than boys. We found that the mean height of the girls in our sample was 1.8 m, while the mean height of the boys was 1.6 m. The difference between these means is large enough that it is unlikely to be due to random chance, and so it is likely that in the general population of the class the average height of girls is greater than the average height of boys.

Evaluating our process

The final step we perform before writing down our report is to evaluate our findings.

- What difficulties did we encounter when trying to measure our results?
- Did we have to change any of our planning while performing our experiment?
- Did anything go wrong that might affect our results?
- Is there anything we would do differently? Why?

Why do we do this? Because it allows people who read our report to decide for themselves how reliable our process and results are, and it allows people who want to reproduce our findings to be aware of why we chose to do things a particular way and what challenges we faced when performing our experiment.

Questions

This week, we will work through the planning for a sample task provided by the Ministry of Education (resource 2.10A).

It has been claimed that teenagers are becoming addicted to energy drinks (which contain high amounts of caffeine) without knowledge of the short-term or long-term effects. People are also concerned that teenagers feel the need for mind or body-altering substances like energy drinks (see the web links below).

- <http://www.stuff.co.nz/the-press/news/2820097/Caffeine-drinks-may-be-hurting-teenagers>
- http://kidshealth.org/teen/food_fitness/nutrition/caffeine.html
- http://www.foodsafety.govt.nz/elibrary/industry/Caffeine_Intake-Confirms_Advice.htm

Select an experimental situation, based on the effects of drinking caffeine, to investigate. Identify the variables you think are important, and write a question to investigate.

The following situations are suggested as a basis for your experiment:

- the effect of drinking caffeine on heart rate;
- the effect of drinking caffeine on reaction times;
- the effect of drinking caffeine on memory.

Write a plan for the experiment. The plan should:

- describe the variables and measures you have chosen and why you have chosen them;
- explain how you will collect your data and record your results;
- link to relevant knowledge about the situation;
- describe any related variables and the possible effects of these;
- describe the experimental method.