Populations and samples

Tuesday, 26 February 2019 12:19 PM

By the end of the lesson I would hope that you have an understanding and be able to apply to questions the following concepts:

- What it means by the term **population**
- What it means by the term **sample**
- What it means by the term **simple random sample**
- Know how to choose a simple random sample
- What it means by the term **population parameters**
- What it means by the term **sample statistic**
- What it means by the term **random allocation**
- Know how to generate samples of normal data for analysis

RECAP:

We need to remember that, for Further Mathematics, we are going to be asked to complete a number of SACs.
The first SAC is based on all the work covered to date.
You will be given data and expected to analyse it and talk about it.
Short summary reports will be required.

The biggest "problem" with this course is knowing what all the language means.
This is important and, hence, this lesson will look at some of the key language used.

By understanding the language we can ensure that no question can trick us!

Population

A population describes a whole group of people.
We generally define what the population might be.

Examples:

- All year 12 students in Australia
- All 18 year old's in the world
- Every male in Victoria
- Every female in London

These are all examples of populations.
We might choose to look at all year 12 students in Australia.
We might then choose to look at a smaller group of these students.

Sample

This is a (small) group of people selected from the **population**.
For example, if the **population** is all Year 12 students in Australia, my **sample** might be all the Year 12 students in Victoria. As Victoria, is a smaller part of Australia, then we are taking a **sample** of my **population**.

A sample MUST be taken from the **population** from which you wish to find information.

The question is … how do we choose which sample to take?
Why did I choose Victoria?
Why didn't I choose New South Wales?

It's not always practical to think of **samples** and **populations** which are so large.
Let's make it smaller.

I'm going to count my **population** as all the Year 12 students in my school.
I'm going to **sample** 50 of these Year 12 students.

200 1/12

How am I going to choose which 50 to choose?

Choosing a simple random sample

So … I want 50 students from my Year 12 … but which 50?
One way to select is to give each students an ID (or number).
This can then be placed into a hat, and 50 chosen.



It would be awesome if we had a magical hat! But we don't.
Writing all those numbers of pieces of paper and placing them into a hat seems a little inefficient!

Let's use some technology! This is a CAS course after all.

Using your CAS to select a random sample

These instructions are for the CASIO classpad.
T-inSpire will follow soon …

OK! We have our 50 people … now what?

I am assuming we have chosen these 50 people for a reason!
So, let's find out something about them.
We could look at their:

- Average Height
- Average Age
- Average Weight
- Average IQ

sample mean $\bar{x}$

population mean = μ

Any of the above would be called a **sample statistic** as it's been calculated from a sample data set. This number represents something about the sample we have chosen.

A **population parameter** deals with the whole population.

If we could measure the height of all people in Year 12, or the weights of all Year 12 students, or the IQ of all Year 12 students, each one of these values would be called a **population parameter.**

Note: The population mean for IQ would be the same in all calculations. The sample mean may not be the same for each sample which we choose.

Randomly allocating to groups

If you have ever been part of a school sports team where you had to be "picked" to be in someone's group you might know the indignity of being chosen last. You were the one who no one wanted. This was my life at school!

If only there way a better way!

There is. Using technology!

Imagine I have the following students who I want to split into two groups:

Grace
Abby
Brittany
Charlotte
Scout
Naomi
Danielle
Emily
Isabelle
Mia
Dayna
Eloise
Peyton
Audrey
Ava
Lily
Emma
Louisa
Gabrielle
Keira

I can use a random number generator to assign each student with a 1 or a 2.

Grace	1
Abby	1
Brittany	2
Charlotte	2
Scout	1
Naomi	2
Danielle	1
Emily	1
Isabelle	2
Mia	1
Dayna	2
Eloise	2
Peyton	1
Audrey	1
Ava	2
Lily	2
Emma	1
Louisa	2
Gabrielle	2
Keira	1

Note: Due to the nature of how the calculator works, there is no guarantee that you will have even numbers of people in each group. You might need to do some constructive fudging to make it even.

Generating samples of normally distributed data

I'm really struggling to think of a reason as to why you might want to do this, but the CAS has an awesome way of taking a mean and standard deviation and turning it into a pretty graph showing a normal distribution.

Example:
Extracted, with permission, from the Cambridge Further Mathematics Textbook series

The distribution of delivery times for Pizzas made by Pizza House is approximately normal with a mean of 25 minutes and a standard deviation of 5 minutes.

a) Use the **randnorm(** command on your calculator to generate a simple random sample of 200 delivery times drawn from a normal population of delivery times with a mean of 25 minutes and standard deviation of 5 minutes. Assign the data to a variable named time.

a) Use the data to calculate the mean and standard deviation of the sample of delivery times and construct a histogram to display data. Comment on the shape of the histogram