Teaching Notes for
Dragon Lesson Plan 5B: Sampling Variation
Curriculum Level 5

Overview

Achievement Objectives

S5-1: … comparing sample distributions visually, using measures of centre, spread, and proportion …

Purpose

This lesson is designed to illustrate one important statistical concept: sampling variation. This lesson can be taught in conjunction with the other Statistics Learning Centre lessons covering other big ideas of statistics introduced at curriculum level 5.

Specific Learning Outcomes

- Appreciate that different samples will provide different, but generally consistent, estimates of a population parameter.

Outline

➢ Dragon lifespan [Full class, 5 min]
➢ Sampling to estimate the population parameter - [Groups, 15 min]
➢ Discussion comparing sample medians – [Class, 10 min]
➢ More sample medians - [Groups, 20 min]
➢ Further discussion - [Full class, 5 min]

Equipment

- Two or more class packs of dragon cards (more cards are better)
- Paper for students to draw box plots on
- Some way to attach a number of box plots to the board

Key Vocabulary

Inference, median, population, population parameter, random sample, sample, sample parameter, sampling variation

Teacher Notes

Timings are approximate.

The statistics strand of the New Zealand mathematics curriculum is made up of three threads: Statistical Investigation, Statistical Literacy and Probability at all curriculum levels. This lesson is a part of the Statistical Investigation thread. The Statistical Enquiry Cycle underlies that thread. While it is used to structure this lesson it is not the main focus of the lesson. Instead the clear focus of the lesson is to introduce students to sampling variation. That is, when using sampling to estimate population parameters the sample estimates will be different for every different sample. But, the variation doesn't stop them from being good estimates.

Dragon Lifespan

The story is intended to provide context and purpose.

It is important to have a context to give the statistical investigation its purpose. This will inform the statistical questions to ask, the appropriate analyses and give meaning to the conclusions.

Sampling to estimate the population parameter

Explain that sampling will provide an estimate, not an exact value. So long as the sample is a random sample and the sample is big enough, the median of the sample will be a good estimate of the population median. A random sample is one in which every card has the same chance to be chosen. By thoroughly mixing the cards you are ensuring this.

The sample size for this lesson should be chosen to make determining the median and quartiles easy. Preferably the median and quartiles should each correspond to the age on a single card. There are different methods for calculating quartiles, be consistent with previous teaching.

The sample size suggested in the lesson is 29 cards. This corresponds to choosing the middle card for the median, the middle of the youngest 15 cards for the lower quartile and the middle of the oldest 15 cards for the upper quartiles. Here the youngest 15 cards and oldest 15 cards both include the median.

If an alternative method is used to calculate the quartiles, the sample size can be adjusted.

It is also important that the whole population is not used for the samples. It is best if there is half of the population or more left after the samples are taken. Given the population size and number of students you may need to adjust the group size or the sample size. However, if the sample size gets too small the variation between samples is likely to get too big.

To compare box plots between groups it is important that they have a consistent scale. To facilitate this it is a good idea to have axis templates with the same scale for the students to use. A grid or guidelines on the templates is also helpful.

Discussion comparing sample medians

The teacher should make clear that each group has done a statistical analysis called inference. They have estimated (or inferred) a population parameter from a sample. Importantly, in such an analysis it is not possible to know the population parameter exactly. Re-emphasise that the sample median will be a reasonable estimate of the population median if the sample was big enough because it is a random sample.

Since this is a lesson and each group analysed a different sample, the class can look at how the estimates compare across different samples. Emphasise that this is not a part of statistical analysis but a learning opportunity.

When displaying the box plots together ensure the axes are aligned. The important value to compare across the different samples is the median. It may be useful to emphasise this with a distinct colour on the box plots.

Possible points to bring out in the discussion:

- There is variation amongst samples – every sample median is different.
- The different samples generally have similar medians.
- The variation between samples is generally less than the variation within a sample.
  - For example, the range of sample medians is generally less than the range (or even the inter-quartile range) from any one sample.

More sample medians

The more sample medians students can compare the better idea they will have of sample variation. Analysing further samples also allows students to practice statistical skills.

It is important to resample from the whole population. Collect all cards and thoroughly mix them each time new samples are taken.

Further discussion

Re-emphasise the points made in the previous discussion. Also, remind students that for a statistical analysis just one sample is used. The multiple samples in this lesson were for learning only, so that they could get an appreciation of sample variation.

Further activities and resources are provided on http://shop.StatsLC.com
Last updated: 25 August, 2016