Kiwi Kapers 2

Note: In this activity in an attempt to capture some of the teaching ideas common concepts have been used to help make the language more student friendly. For some of the concepts that are common to both everyday language and statistical technical language the intention was not for the technical definition to be used. For example using reasonably sure and fairly confident in this introduction.

Follows Kiwi Kapers 1. This activity explores what sample size we want to take to be reasonably sure that the inference(s) we make are representative of the population. This can be done either manually using the data cards and drawing different size samples or, more efficiently, using a random sample generator in Fathom (or similar technology).

Part One (two lessons): Take multiple samples of size 15, 30, 50 and 100 to decide what is the smallest sample size we can take to be fairly confident of the results. The students have already explored samples of size 15 and 30 using the manual method. They will do these again and also do samples of 50 and 100.

Part Two (one lesson): Take 100 samples of each sample size and look at the distribution of sample medians. The two parts together provide learning experiences towards making a decision about what is a suitable sample size to use.

Problem

We are going to use the same question as Kiwi Kapers 1.

You may decide to use any summary question from any data set that uses continuous data.

Investigative question 1:
I wonder what are typical weights of kiwis?

What is a sensible and reliable sample size to use for making inferences about a population?

Feedback from trialling in bold.

Examples of questions
I wonder…
What are typical weights of kiwis?
What are typical heights of kiwis?
If female kiwis tend to be lighter than male kiwis?
If there is a difference in weight of kiwis depending on their species?
If Great Spotted kiwis tend to be taller than Southern Tokoeka kiwis?

Plan

To answer our question what variable(s) do we need to use from our data set?

We also want to know how big a sample we need to take to get a reliable answer.

Use Fathom (or similar) to generate random samples. Use sample sizes 15, 30, 50 and 100.

Prior thinking needed before starting this activity. We want to use the median of the sample to make an estimate of the median of the population. Any given sample will develop a slightly different median, so we want to use this sample median to find an interval that is likely to captrue the population median. We want to estimate the population median with some level of confidence. How might we go about doing that?

Refer back to the outcomes from the previous activity Kiwi Kapers 1. The samples of size 15 had median weights that ranged from 2.2-3.0kg and the samples of size 30 had median weights that ranged from 2.3-2.8kg.

We need to use weight.

Have explored 15 and 30 previously, want to look at these again and also explore 50 and 100.

Data

PART ONE:
The data for the first part is available in the Fathom files **kiwipopsample.ftm**; for the second part **kiwipopsample2.ftm**.

The student instructions for using Fathom are in the file **kiwistudentfathom.doc**.
Preparation: PART ONE: print pages 1 and 2 backed, print page 3 separate
PART TWO: print pages 4-6, backed and collated.

Students need to take 5 samples of size 15, 30, 50 and 100. These need to be recorded in their books by copying the box plot from the screen and recording the median for each sample. Details on page 3 of student instructions.

See **sampleskiwi2part1.doc** for an example of what this might look like.

Lesson one involved familiarisation with what is in the Fathom file and working out how to take samples. Most groups managed to get 5 samples of 15 recorded.

Lesson two involved getting the remaining samples done and moving onto the written analysis.

Analysis

From the box plots, I notice statements need to be made for each set of 5 graphs from each sample size. They should comment on medians, IQR and range.

Once they have commented on the individual samples sizes, they should make I notice statements how the variation in the medians, IQR's and ranges changes as the samples size increases.

What I did was put five samples of each up on the data show and ran my finger down the medians for each group of samples (eg sample size 15), which indicated quite nicely how they varied.

I notice that the median weights of kiwis in samples of size 15 are from ___ to ___.
I notice that the IQR of weights of kiwis in samples of size 15 are from ___ to ___.
I notice that the range of weights of kiwis in samples of size 15 are from ___ to ___.

I notice that the variation of the median weights of kiwis decreases as the sample size increases. For samples of size 15 the median weight ranged from 2.2-3.0kg, a difference of 0.8kg, whereas for samples of size 100 the median weight ranged from 2.38-2.68kg, a difference of 0.3kg.

Data

PART TWO:
The data for the first part is available in the Fathom files **kiwipopsample.ftm**; for the second part **kiwipopsample2.ftm**.

The student instructions for using Fathom are in the file **kiwisamplefathom.doc**.
Preparation: PART ONE: print pages 1 and 2 backed, print page 3 separate
PART TWO: print pages 4-6, backed and collated.

Students generate 100 samples of size 15, 30, 50 and 100. The sample medians are recorded and graphed on a dot plot. The median and IQR of the sample medians is recorded in a table. The graphs of the sample medians are copied and pasted into a word document. The median and IQR of the sample medians are noted also. Details outlined in pages 4-6 of the task.

Actually talked the students through collecting their first 100 measures (medians) and copying this into a word document (see pages 4-6). Then immediately made an I notice statement about the range of the sample medians, they typed these into the word document. *I notice that the sample median weights of kiwis for samples of size 15 vary from 2.15kg to 3.15kg.*

Analysis

Make I notice statements about the range of the sample medians, the bulk of the sample medians and the middle of the sample medians.

The file is set up to give the median of the sample medians and the median of the sample IQR.

See **mediancollection.doc** for an example of what the fathom information might look like and what students should be able to put into a word document.

Once the students have written about their graphs, collect in the class information. Set up four axes on the board for median of the sample medians, and four axes for the median of the sample IQR.

Make a dot plot of the medians and the IQRs by collecting the values from the students. On the first axis collect all the medians for samples of size 15, on the second samples of size 30 and so on. This should

I notice that the sample median weights of kiwis for samples of size 15 vary from 2.15kg to 3.15kg. The bulk of the sample median weights of kiwis for samples of size 15 were in the range 2.34kg to 2.75kg. The median sample median weight of kiwis for samples of size 15 was 2.486kg.

help to build the argument that 15 is not very good, but the others improve bit by bit.

Repeat for the IQR's.

See **classmedians.doc** for an example of what this may look like. This has been done in Fathom, recommend just collating on board.

Conclusion

I wonder what are typical weights of kiwis?
They may add to their conclusion previously regarding this question. They may have found additional evidence.

What is a sensible and reliable sample size to use to make inferences about a population?
I would like students to identify that samples of 30 or 50 are ok. (Because there is not too much variation in the medians.) Samples of 15 are too variable and samples of 100, while even better, are not so much better that we need to spend the extra time and energy collecting 100 kiwis when 30 or 50 will do. We want the idea of the smallest sample that should give us a reliable estimate of the population.

Reflection

What have you learned about the variability of sample medians? Why?
Will you be able to use your learning to evaluate results from any sample? Why?

Extension activity

What is the typical IQR for weights of kiwis for a particular sample size?