THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

DATA to insight
An Introduction to Data Analysis

Ways of obtaining Confidence Intervals

Chris Wild, University of Auckland

Methods for calculating confidence intervals come via two main routes:

- **mathematical theory** obtained under idealised assumptions (e.g. normal distributions);
- **computer intensive methods** of which the most generally useful is the **bootstrap**.

The first is the historical route - necessitated by a historical lack of computational power. The second route provides solutions that are easier to understand, more powerful, and more generally applicable but the mathematical approaches "got there first".

Consequently, for the most common situations and long-standing problems, the default methods in statistical software, and the methods most researchers are familiar with, are the ones from the historical, mathematical tradition. The places where computationally-intensive methods are most firmly established are for newer and more complicated problems.

In many common situations, 95% confidence intervals based upon normal theory have the nice simple form:

estimate ± 2 standard errors.

(The "2" here is approximate. The multiplier from large-sample normal theory is 1.96. Larger multipliers obtained from something called Student's *t*-distribution are used when we have small samples.)

The "2 standard errors" is the *margin of error*. The formula used to calculate the standard error depends on the type of estimate; the standard-error formula to be used for a mean is different from that for a proportion, or a difference in means, and so on.

Confidence intervals can be formed in the same way from bootstrap resampling. In that case we use the standard deviation of the set of bootstrap estimates to give us the standard error. For means and proportions (or percentages) the standard error we get from the normal theory is virtually identical to that obtained from a very large number of bootstrap resamples.

Theoretically derived mathematical methods are deduced from idealised assumptions that will never be satisfied exactly, so deriving a method is just a first step. We then have to check the extent to which the method will work under departures from its assumptions.

This is the idea of **robustness** and we investigate robustness using computer simulations. That is how we know that confidence intervals for means from normal-distribution theory are fairly robust, but normal-theory confidence intervals for standard deviations are so sensitive to even small departures from normality that they are not fit for practical use.

(*The paragraph to follow is only for those who have previously learned some statistical theory*) We can also relax the strict normality assumption and show theoretically that the traditional CIs for means work "asymptotically", as a consequence of a mathematical result called the central limit theorem. "Asymptotically" means 'as the sample size tends to infinity'. (The n=30 rule of high school statistics has no theoretical basis.) The speed at which asymptotic behaviour starts to work depends on the parent distribution being sampled from. This behaviour has been investigated, once again, by using simulation. The results of these simulations are how we know that asymptotic behaviour for means works fairly fast for distributions not "too far" from normal. It takes large samples to get reliable coverage rates for a proportion – very large as that proportion gets closer to 0 or 1.

Review Questions:

- *What are the two main ways of approaching the problem of obtaining confidence intervals?*
- *Why are methods based on mathematical theory the default methods in most packages for long-standing problems?*

- *What is the nice simple form that many intervals based on normality theory have?*
- *Why is it necessary that the fitness for practical use of any method (whether theoretically derived or from computer-intensive methods) be investigated using computer simulation?*

(updated March 2017 by adding Review Questions)