Topic 4

**Random Sampling**

To obtain a sample that can be used in a study, *members of a population should be selected randomly*. Note that this requires a careful selection or design of a population frame. If members were not selected randomly, then biases will exist in the sample population, and results from the study cannot be generalized to the larger population.

Note that bias can come into a sampling in many ways, and not just if the population was randomly sampled. An improperly designed data frame, an improperly designed survey or improperly worded questions on the survey can create biases. Even differently worded questions, but asking essentially the same thing, generate different data.

**Obtaining a random sample is not an easy task.**

 A random process of selection may not actually be random.

 A sample that *appears* random may actually not be random.

 Converse is true: A sample that appears non-random is actually random

Examples:

1. In the coin flipping example there were several long runs of Tails. Most people would expect only to see three or 4 tails at most.
2. Two-dimension randomly plotted points. Note the spaces and clumping. This appears to have areas of non-randomness.



So *don’t trust your senses*. **Follow random selection protocols**. Selecting a Random sample can be “extremely difficult.” Can be. Not all the time, but can be.

**Simple Random Sampling**[[1]](#footnote-1)

* 1. Every member of the population has the same probability of being selected.
	2. Every sample of size n has the same probability of being selected as every other sample of size n.

**Random Selection Methods**

 **Physical Mixing** - make sure probability is the same for all objects

**Table of random digits -** number the sample OUs. Tables can be found on the internet or in many books

**Computer generated randomness –** computation method. Very fast, and computer generated randomness complies with randomized tests.

**Coin Tossing**  - generally accepted for making binary decisions

**Sampling variability** – the values of sample statistics: $\overbar{x},\hat{p}$,s vary from sample to sample.

Here are several graphs of the distribution of the means of multiple samples. There are multiple samples of sizes, n=4, n=16, and n=64 (each line represents what a histogram of multiple samples of size n would look like). Notice that for larger n, the variability of the value of the mean is smaller.



**Accuracy**[[1]](http://en.wikipedia.org/wiki/Accuracy_and_precision%22%20%5Cl%20%22cite_note-metrology_terms-1) of a [measurement](http://en.wikipedia.org/wiki/Measurement) system is the degree of closeness of measurements of a [quantity](http://en.wikipedia.org/wiki/Quantity) to that quantity's actual (true) [value](http://en.wikipedia.org/wiki/Value_%28mathematics%29).

**Precision -** of a measurement system, also called [reproducibility](http://en.wikipedia.org/wiki/Reproducibility) or [repeatability](http://en.wikipedia.org/wiki/Repeatability), is the degree to which repeated measurements under unchanged conditions show the same [results](http://en.wikipedia.org/wiki/Result)[[2]](#footnote-2).

**Precision and accuracy are dependent on sample size**

Small samples are less accurate and less precise

Large samples are more accurate and more precise

**Some samples may be very different from the population.**

 This is statistically ‘unlikely.’ This event has a low probability.

 A sample will never be an outlier.

**Population Size does not affect sampling variability.**

 (population must be ~10x as sample)

Note:

Don’t get biased and unbiased *samples*, and biased and unbiased *statistics* confused, we have not yet covered unbiased statistics.

1. The book uses the term “equal chance.” “Same probability” and “equal chance” are synonymous [↑](#footnote-ref-1)
2. [**^**](http://en.wikipedia.org/wiki/Accuracy_and_precision#cite_ref-Taylor_2-0) John Robert Taylor (1999). [*An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*](http://books.google.com/books?id=giFQcZub80oC&pg=PA128). University Science Books. pp. 128–129. [ISBN](http://en.wikipedia.org/wiki/International_Standard_Book_Number) [0-935702-75-X](http://en.wikipedia.org/wiki/Special%3ABookSources/0-935702-75-X). [↑](#footnote-ref-2)