Random sampling vs. assignment

Statistics 101

Duke University

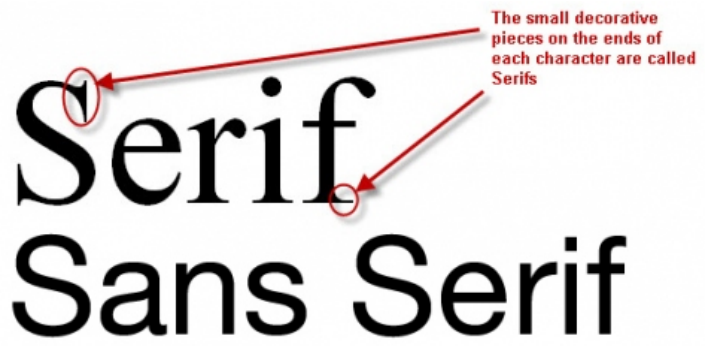
Mine Çetinkaya-Rundel

Learning objective(s):

- Classify a study as observational if the researcher merely observes the data and as an experiment if treatments are imposed on subjects.
- Assess whether the study's results can be generalized to the population based on whether or not random sampling is employed.
- Determine whether the study's results suggest causation or association association based on whether or not random assignment is employed.

Example: sampling vs. assignment

Suppose you want to conduct a study evaluating whether people read serif fonts or sans serif fonts faster.



Source: http://nicolecommawoo.wordpress.com/2010/02/08/signs-of-worcester3

- Ideally, you first randomly *sample* subjects for your experiment.
- Then, once you have a collected a sample of subjects, you randomly *assign* half of them to read text in serif font, and the other half to read text in sans serif font.
- So sampling happens first, and assignment happens second.

Why random sampling and assignment?

- Random sampling allows us to obtain a sample representative of the population.

  Therefore, results of the study can be generalized to the population.

- Random assignment allows us to make sure that the only difference between the various treatment groups is what we are studying. For example, in the serif/sans serif example, random assignment helps us create treatment groups that are similar to each other, and the only difference between them is that one group reads text in serif font and the other in sans serif font.

  Therefore, causality can be inferred.

Summary

	Random assignment	No random assignment	
Random sampling	Causal inference, generalized to the whole population.	No causal inference, correlation statement generalized to the whole population.	Conclusions generalized to population.
No random sampling	Causal inference, only for the sample.	No causal inference, correlation statement only for the sample.	Conclusions not generalized to population.
	Causation	Correlation	