Random Samples / Randomization

http://www.tufts.edu/gdallal/rand.htm

Random Samples and Randomization are two different things, but they have something in common as the presence of random in both names suggests — both involve the use of a probability device. With *random samples*, chance determines who will be in the sample. With *randomization* [aka, random assignment], chance determines the assignment of treatments.

A random sample is drawn from a population by using a probability device. We might put everyone's name on a slip of paper, mix thoroughly, and select the number of names we need. The use of a probability device to select the subjects allows us to make valid generalizations from the sample to the population.

In an intervention trial, randomization refers to the use of a probability device to assign subjects to treatment. *This allows us to use statistical methods to make valid statements about the difference between treatments for this set of subjects. The subjects who are randomized may or may not be a random sample from some larger population. If they are a random sample, then statistical theory lets us generalize from this trial to the population from which the sample was drawn. If they are not a random sample from some larger population, then generalizing beyond the trial is a matter of nonstatistical judgment* [italics added].

Randomization models: Why statistical methods work for intervention trials involving volunteers.

Statistical methods are usually presented and justified in terms of random samples. Almost all intervention trials involve volunteers, usually recruited locally. If *convenience samples* are inappropriate for surveys, how can they be appropriate for intervention trials? There are two distinct issues to address — validity and generalizability. First, whether the experiment is valid [internal validity], that is, whether observed differences in this group indicate a real difference in treatments insofar as these subjects are concerned. Second, whether the results can be generalized to any other group of individuals.

The reason volunteers can be used to make valid comparisons comes from the use of randomization in the assignment of treatments. It is beyond the scope of this course to give mathematical proofs, but the common statistical methods that are appropriate to compare simple random samples are also valid for deciding whether the observed difference between the two treatments is greater than would be expected when subjects are assigned to treatments at random and the treatments are equivalent. The random sampling model and the randomization model lead to the same statistical methods.

Within broad limits, genralizability is possible because human beings are made out of the same stuff. While this justification cannot be applied blindly, it may be comforting to know that many of the surgical advances of the mid 20-th century were developed in VA hospitals on middle-age white males. However, the ability to generalize results does not immediately follow from the use of particular numerical methods. Rather, it comes from the subject matter specialist's knowledge of the subjects who were studied and the group to whom the generalization will be made.

It follows that evidence regarding factors under the control of the investigator is different from evidence regarding factors that cannot be subjected to randomization. For example, consider an intervention trial that compares the effects of two diets in smoking and nonsmoking pregnant women. The use of statistical methods to compare diets can be justified by the random assignment of subjects to treatment. However, the comparison between smokers and nonsmokers depends on an enrollment procedure that would not recruit smokers whose response to the diets would different from the nonsmokers'.

The randomization model sometimes allows standard statistical techniques to be used when there has been no random sampling and no random assignment to treatment. From Dallal (1988):

> Consider, for example, a situation in which a company's workers are assigned in haphazard fashion to work in one of two buildings. After yearly physicals are administered, it appears that workers in one building have higher lead levels in their blood. Standard sampling theory techniques are inappropriate because the workers do not represent samples from a large population - there is no large population [nor have the workers been assigned to buildings at random]. The randomization model, however, provides a means for carrying out statistical tests in such circumstances. The model states that if there were no influence exerted by the buildings, the lead levels of the workers in each building should be no different from what one would observe after combining all of the lead values into a single data set and dividing it in two, at random, according to the number of workers in each building. The [probability] model, then, exists only in the analyst's head; it is not the result of some physical process, except insofar as the haphazard assignment of workers to buildings is truly random.

Of course, randomization tests cannot be applied blindly . . . (Perhaps, in the lead levels example, one building's workers tend to live in urban settings while the other building's workers live in rural settings. Then the randomization model would be inappropriate.) . . . In the context of randomization models, randomization tests are the ONLY legitimate tests; standard parametric test are valid only as approximations to randomization tests.