General Strategies for Sampling a Population

ORIGIN ≡ 0

In *Fundamentals of Biostatistics 6th Edition* (Thomson Brooks/Cole publishers 2006, pp. 169-177) Bernard Rosner provides an excellent survey terminology often employed in obtaining "unbiased" and "representative" samples that may be viewed as fairly representing the population from which they come. Although statistical calculations are generally silent about issues of sampling, of course, the believability of statistical results obtained are often critically dependent on them. Rosner highlights the use of pseudo-random number tables in setting up different kinds of studies, and provides some important terminology, summarized here...

Random Selection - Use of random numbers to select uniquely identified individuals from a population, usually without replacement.

Random Assignment - Use of random numbers to a assign *fixed numbers* of individuals to each treatment or analysis category, usually without replacement.

Randomized Trial - In comparing the effect of different levels of "treatment" (clinical or otherwise), individuals from a population are assigned *at random* to specific treatment classes (or categories). This hopefully guards against some other factor biasing the sample and being responsible for observed difference in outcome between the classes, rather than the treatment themselves.

Block Randomization - Random selection placing individuals into treatment classes often involves *replicate blocks of all treatments* - each essentially a randomized trial.

Stratified Design - Treatment classes are set up explicitly regarding values observed in individuals for one or more "accessory" or "covariate" variables. The different classes defined by these variables are called *strata* (singl. *stratum).* Within strata, random selection, random assignment, or block ransomization may also be employed.

Blind Design - When *knowledge* on the part of researcher, subject, or both ("double blind') might influence behavior within strata or blocks, care is taken insulate the study from this knowledge.

Standard Statistical packages, such as SYSTAT, SAS, SPSS and R, offer the ability to partition data into strata and sub-blocks with ease. Thus, once prototyped, these programs can offer a significant time advantage in analysis of large data sets having complex design.