COLLOQUIUM

Department of Statistics and Probability
Michigan State University

Raoul LePage

Department of Statistics and Probability, Michigan State University

Randomization Transfer

Tuesday, March 23, 2010
A405 Wells Hall
10:20 a.m. - 11:10 a.m.
Refreshments: 10:00 a.m.

Abstract

The title refers to methods whereby randomization prior to data collection is afterwards mathematically transferred onto such discrepancies as may exist between a proposed model and data. The effect of randomization transfer is to transform discrepancies, which may not be random, into random variables. Among many worthwhile objectives we seek to:

1. Replace unscientific assumptions as to randomness of data with the means to make it so.
2. Replace unscientific assumptions that models are correct with the means to statistically analyze models relative to their capacity to fit data or achieve other goals.
3. Legitimize the application of simple general-purpose statistical methods such as multiple linear regression and experimental design.

How comfortable are we with assuming that the departures of sample scores from their population mean are intrinsically random regardless of how we obtain our sample units? Why then are we seemingly comfortable with likewise specious assumptions in far more complex settings?

Freedman studied bootstrap for correlation-model multiple linear regression. That is, (x, y) jointly random as may be achieved by randomly sampling x and observing y. His results accomplish (1) and (2) if we are willing to use sample least squares fit as an estimator of population least squares fit. Heteroscedastic discrepancies between the population fit and the population y-scores do not especially recommend this, even if we might be willing to accept linear models as suitable population descriptions. That is where randomization transfer comes in.

Reference, Freedman, D. (1981). BOOTSTRAPPING REGRESSION MODELS. *Annals of Statistics,* **9(6)**, 1218-1228.