Scope of Inference Writing Examples

From *The Statistical Sleuth*

- Since this was a randomized experiment, one may infer that the difference in creativity scores was *caused* by the difference in motivational questionnaires. Because the subjects were not selected randomly from any population, extending this inference to any other group is speculative. This deficiency, however, is minor; the causal conclusion is strong even if it applies only to the recruited subjects.

- Although there is convincing evidence that the males, as a group, received larger starting salaries than the females, the statistics alone cannot address whether this difference is attributable to sex discrimination. The evidence is consistent with discrimination, but other possible explanations cannot be ruled out; for example, the males may have had more years of previous experience. [No randomization to groups (obviously!) and no random sampling was used]

- These twins were not randomly selected from general populations of schizophrenic and non-schizophrenic individuals. Tempting as it is to draw inferences to these wider populations, such inferences must be based on the assumption that these individuals are as representative as random samples are. Furthermore, the study is observational, so no causal connection between the left hippocampus volume and schizophrenia can be established from the statistics alone. In fact, the researchers had no theories about whether the abnormalities preceded the disease or resulted from it.

- First, as this was an observational study, one cannot infer a causal relationship—that the longer humerus lengths among survivors enabled them to survive. Second, the living sparrows brought to Bumpus were found in a specific area and were so stressed that they were easily collected. Inference to populations of similarly stressed sparrows is risky. Such populations are hypothetical, and there is no chance model.

- Since the samples were not random, inference to the populations is speculative. We must ask whether the participating veterans were representative of their respective groups. For example, nonparticipating Vietnam veterans may have failed to participate because of dioxin-related illnesses. If so, statistical statements about the populations of interest could be seriously biased.

- These observational data cannot be used to establish causality, nor is there any broader population of which they are a sample. But, the association between temperature and O-ring failure in these particular 24 launches is consistent with the theory that lower temperatures impair the functioning of the O-rings. (At one point in public hearings into the causes of the disaster, Feynman asked for a glass of ice water, placed a small O-ring in it for a time, removed it, and then proceeded to demonstrate that the rubber failed to spring back to its original form.)