Some Aspects of Reasoning about Variability

Bernard Harris
Department of Statistics
University of Wisconsin
1300 University Avenue
Madison, Wisconsin, 53706, USA
Bernard.harris@pacc-eng.com

1. Abstract

Statistics, as a discipline, exists because of variability. If there were no variability, the universe would be deterministic and there would be no need for statistical methods. Therefore, it is with some dismay that in the populace little attention is paid to variability and consequently students in the higher grades of secondary schools and in elementary classes in universities have so little prior exposure to this concept. I am inclined to put much of the blame for this on the media. The evidence is primarily anecdotal. I will provide a number of exampl es to support this hypothesis.

A second aspect arises when "outliers" are present. "Outliers" are one or more observation in a data set that is "substantially different" from the other members of the data set. Its relevance to this topic is due to the fact that the presence of outliers results in an increase in the measaured variability of the data set. Almost every textbook and almost every elementary course in statistics has a brief discussion of dealing with outliers. In the examples to be presented, I will show that many of these suggestions are sometimes incorrect, leading to false reasoning about variability.

A third aspect that I will mention in this discussion occurs in theory of point estimation as well as in other areas of statistical inference .. If one envisions repeated experimentation, there is variability in the value of the estimates. Typically, this variability is measured by the variance of the sampling distribution of the estimator. Thus, the precision of the estimator is often measured by expected squared error. I propose to exhibit a class of problems for which squared error is inappropriate, contrary to the usual beliefs given in elemntary statistics classes and elementary statistics textbooks.