

Fall 2008 COURSE ANNOUNCEMENT

MATHEMATICS 542D: NONPARAMETRIC STATISTICS

Times: Monday & Wednesday 4:00-5:15 PM

Instructor: Mark Schilling SN 415

-- Do men and women have different levels of fear when jumping out of an airplane? To help answer this question, researchers measured the maximum heart rates of ten male and ten female skydivers just prior to their first jump and compared them to their resting heart rates.

-- An advertising agency believes that annoying advertisements will be more effective than more pleasant ones due to their stronger impact. To test their theory, the agency persuades their client to experiment by using annoying ads in several randomly selected cities and more pleasant ads in others. The agency then compared the product sales in the two sets of cities to see if the annoying advertisements produced more sales.

-- Data collected on the times between consecutive eruptions of the Old Faithful Geyser in Yellowstone National Park show that despite its reputation of erupting regularly every hour, there are two general clusters of times, one shorter and one longer. Can we fit a density curve to this data that successfully models these inter-eruption times without assuming a functional form in advance?

A standard approach to problems such as these is to assume that the data has arisen from a specific parametric model according a known probability distribution, such as the normal, exponential or poisson. In many cases, however, the distribution that has generated the observations is unknown; in these cases, analyses based on an assumed distributional form may be inefficient, at best, or worse, improper and erroneous.

This course will illustrate and analyze a body of procedures that differ from classical statistical procedures in that they do not require the distribution generating the data to be known or assumed. We will discover that several of these "distribution free" methods, while often simpler in form, compete favorably with and in fact often substantially outperform traditional "parametric" methods.

Course Structure: Class time will be comprised of lectures supplemented by group activities involving the analysis of illuminating data sets. Some use of statistical software will be involved.

Prerequisites: Introductory probability theory and mathematical statistics (Math 340 and 440A or equivalent), or instructor approval.