MA 568 – Statistical Analysis of Point Process Data
Uri Eden

Data Analysis Project:
Airplane Arrival Data

The U.S. Bureau of Transportation Statistics provides access to data about scheduled and actual arrival and departure times for flights from major air carriers. While the scheduled arrival and departure times are deterministic, the actual ones are quite variable and have complicated dependencies on previous delays, time of day, and other factors. Understanding relationships associated with delayed flights can be helpful in developing new flight schedules to reduce the delays and congestion often associated with airline travel.

Each data set (In Microsoft Excel format) contains the actual arrival and departure times of all flights for a single major airline over one month. There are twelve data sets corresponding to each month of 2006. Also included are the delay times, the date of the flight, the day of the week, and the origin and destination of each flight. The scheduled time for each flight can be computed based on the actual times and the associated delays.

Questions:

1) How does flight scheduling depend on time of day and day of the week? Does a multiplicatively separable model accurately describe these dependencies?

2) Is there a significant difference between the intensity of scheduled arrivals/departures and actual ones? Does this difference depend on other variables such as time of day or location of origin?

3) How well do the scheduled departure and arrival times predict the actual ones?

4) How do delays occurring in one airport affect subsequent departure times from that airport and from other airports? Does this depend on which airports you consider? Are there specific airports that have more of an effect on the network?

5) What types of interactions are present between all of these processes? How can these interactions be modeled?

6) Do these analyses suggest any strategies for reducing the variability and delays associated with airline travel?