Module 12: Advanced Topics

12.2 Advanced Topics in Environmental Statistics

Advanced Topics

- There are several topics we have just barely touched on and that would warrant their own courses:
  - Nonparametric Methods
  - Time Series Analysis
  - Spatial Statistics
  - Risk Assessment
- You may want to look for courses in these areas if they are relevant to your work

Advanced Topics

- There are many topics in advanced statistics that we haven't even touched on
  - Bayesian Methods
  - Categorical Data Analysis
  - Biostatistics/Survival Analysis
  - Epidemiology
  - Multivariate Methods
    - Factor Analysis
    - Discriminant Analysis
    - Cluster Analysis

Bayesian Methods

Reverend Thomas Bayes (1702-1761)

Classical vs Bayesian Statistics

- The classical, or frequentist, perspective is that the probability of an event is defined as the frequency with which it occurs in a long sequence of similar trials
- The Bayesian, or subjectivist, perspective is that the probability of an event occurring is the degree of belief a person has in the occurrence

Advantages of the Bayesian Approach

- No long sequence or representative population of similar trials may exist
- Allows probabilities to be defined using expert judgement
- Allows distributions to be updated/refined if new information becomes available
- Allows data and judgment to be combined

Bayes Theorem

$$f''(Y) = \frac{f'(Y) f_e(S|Y)}{\int f'(Y) f_e(S|Y)}$$

- where:
- Y = predicted values from a model
- S = sample of new data values
- f'(Y) = prior distribution describing the probability distribution function of the model predictions prior to the
- incorporation of new data
- f''(Y) = posterior distribution describing the probability distribution function of the model predictions after the incorporation of new data
- $f_e(S|Y)$ = likelihood function of sample S given that Y is true

Categorical Data Analysis

- Sometimes data are not quantitative but are counts
  - A study of people might collect data such as sex, educational level, ethnicity, etc
- Studies of this type often look for relationships between these categorical variables
  - relationship of sex and ethnicity with educational level
- Sometimes the goal is to develop relationships between quantitative variables and the categorical variables
  - Relationship of sex, ethnicity, and educational level with income

Biostatistics/Survival Analysis

- There is an entire body of statistics that deals with biological organisms, health and survival
- These studies are complicated by many covariates and uncontrolable variables
- It's particularly difficult to study humans since many aspects of our genetics and behavior are not well understood or in the experimentor's control

Biostatistics/Survival Analysis

- Example:
  - 1000 people with heart disease are split into groups
  - One group is given a placebo
  - Others are given various treatments
  - They are studied for ten years and mortality is recorded
- Complications:
  - People drop out of the study for many reasons
  - People die of other things (car accidents for example)
  - Many of the people won't have died in 10 years

Epidemiology

- Epidemiology is the study of human disease using retrospective data (in the past)
- This type of data is used to look for hazards
  - Worker health data may be examined for health impacts from occupational exposure
- It is also used to look for disease clusters
  - May be related to local hazards

Multivariate Methods

- These methods are used when there are a large number of variables which are correlated with one another
- Often it is difficult to determine relationships when variables are all correlated
  - A suite of water chemistry data can give many variables but they have complicated cross-correlations
- These methods can be used to sort data points into groups or to develop a smaller number of independent composite variables that can then be used as new “factors”

Advanced Topics

- The most important thing is to know when you need help
- Don’t use a method that you don’t understand
- Look up assumptions and check to see if they are met
- Find someone to help you!