Statistics 712: Applied Statistical Decision Theory
Spring 1999 Syllabus

Robert Stine
stine@wharton.upenn.edu
Department of Statistics
3014 Steinburg Hall-Dietrich Hall / 6302

Overview

This course describes the use of statistical methods and value of statistical perspectives in practical decision making. Some decisions are made with the help of hard, quantitative information and standard statistical methods. We might extrapolate the track record of some investment before deciding on how much to add to our portfolio, for example. Similarly, a company might develop a regression model to predict the sales of a new product line before deciding how much to grow its sales force. Other decisions are made on the fly, relying on instincts, hunches or downright guesses. If we are launching a brand new product, we may not have the data that traditional analytic tools require. The costs and benefits offered by the alternatives in such cases are not so clear, and often hard to trade off. Even with hard data, it can be hard to make a choice. What level of risk should we be willing to accept? With or without precise data, the ideas of statistics can lead to better decisions.

We begin with methods that leverage very precise quantitative data, and gradually move to situations more characterized by guesses and intuition. We will start with an overview of statistical modeling, with emphasis on how to avoid the pitfalls of coincidence and confounding when using these tools. Hopefully, much of this — at least the tools themselves — will be review. We next turn to problems of risk assessment and utility. Many of the motivating examples in this part of the course, including the capital assets pricing model and hedging, will come from finance. We will then move from methods that rely on precise data to those that use less quantitative data, beginning with the use of subjective confidence intervals. Finally, we will revisit some of the financial problems treated using precise data.

To this point in the course, we will have treated decisions as though they have no effect upon the environment of the decision maker. This is an artificial setting, and most decisions have important strategic implications. A decision is likely to affect what rivals and competitors will choose to do. We will start our treatment of these aspects of decision making with a survey of key results from game theory, then move to problems in determining the value of information and distinguishing uncertainty from randomness. Auctions provide an interesting setting to discuss many of these ideas. Many of the lessons learned from our study of quantitative methods will resurface in less obvious ways.

Statistical methodology for precise, quantitative data

We need to review the foundations of statistical inference and prediction so that we share a common point of reference. We will emphasize the notions of comparison and forecasting, since these are key in many decisions. We'll review classical testing methods to see when how and if they are appropriate.

- Head-to-head comparisons (setting an α level with two groups)
- Judging the goodness of an interval estimate: length versus width
- Comparisons of many things (issues of multiplicity, Bonferroni)
- Bayesian decision theory
- Regression modeling
  - selecting the model and dealing with specification error
  - dealing with sparse data: myth of interpolation and extrapolation (setting the right width of uncertainty)
- Methods using time series
  - autoregression and the use of lagged variables
  - exponential smoothing
- Probability forecasts and calibration
  - logistic regression (challenger)
  - calibrating others' forecasts
- Technology-driven methods, data mining
  - stepwise regression
  - neural networks
  - CART, MARS and other partitioning models

Risk and utility

Traditional statistical prediction intervals have a rigorous mathematical foundation, yet are often ill-suited to the problem at hand. Under-predicting sales usually has very different consequences from over-predicting sales. The notion of utility provides a unified way to look at some of these issues, particularly those that can be assigned a dollar value.

- Utility and certainty equivalent value
- Judging the goodness of an interval estimate, revisited
- Predicting the level of risk in financial markets
  implied volatility, Black-Scholes, and time series methods
- Judging an appropriate level of risk aversion
- CAPM and investing in equity markets (dice)
- The principal agent problem and incentive structures (stock ownership plans)
- Using of hedging and options to reduce risk

Subjective confidence intervals

In contrast to traditional statistical intervals, we have all formed our own estimates of what might happen in a given situation. We might expect a stock to return between 5% and 10% over the next month, or predict the spread in a football game to be 7-10 points. In either case, we have formed our own subjective estimate of what is plausible. We will see that though less formal that the usual statistical prediction interval, subjective intervals are very important and form an important tool for making decisions.

- Properties of subjective intervals
- Forming better subjective intervals: Social psychology, group dynamics
- Judging the goodness of an interval estimate, revisited
- Combining multiple subjective intervals
- Applications: project evaluation and budget allocation

Game theory and the role of strategy

Game theory explores the strategic side of decision making, recognizing that our actions are often in competition with those of others for limited resources. In making a decision, one needs to recognize that the environment will change because of our decision. The best choice in an isolated world is often different from what works best in competition with others.

- Two-person games
- Tit-for-tat and the rewards of cooperation
- Prisoner's dilemma
- Private knowledge

Decision making with no data

Without firm data and even the knowledge to build a subjective interval, can anything be done when confronted with a pressing question? Of course — get some information! Determining how much and recognizing its value is essential for good decision making.

- Relative value of information
  - Decision trees and valuing information
- Separating randomness from uncertainty (calibration, revisited)
  - "I don't know" vs "It's between .3 and .5."
  - Bayesian decision theory, revisited
- Is more information better?
  - Yes, in the non-strategic world, but not always when strategies emerge.
  - "No good deed goes unpunished." (working late)
  - Incentive structures
- Winner's curse and auctions (Bonferroni, revisited)
  - Types of auctions
  - MBA course auction, on-line auctions (eBay)

Office Hours

Office hours *follow* class, running from 4:30 to about 5:30 p.m. For other times, you will need to make an appointment. Other than coming by my office and leaving a note, you can best reach me via e-mail at `stine@wharton.upenn.edu`. As needed we'll use a web page that you can access as `http://www-stat.wharton.upenn.edu/~bob/stat712`. Handouts and assignments will be distributed using this web page, so if you miss something, you can retrieve it there (as an Acrobat *pdf* file).

Textbooks and Software

There are several books that contain the majority of reading material for this course. For those of you that have not used this text, much of our initial material comes from the following casebook which many of you already have:

Foster, D. P., R. A. Stine, and R. Waterman (1998).
Business Analysis using Regression, Springer, New York.

We will also use the following inexpensive books to supplement this text. We may discover some other useful ones along the way. You can obtain these most easily from amazon. For our treatment of game theory, the key ideas are well-explained in

Axelrod, R. (1985). *The Evolution of Cooperation*. Basic Books.

For further discussion (from a bit different point of view) of subjective intervals and the importance of taking all the relevant factors into account, we'll rely on

Russo, J. E., and P. J. H. Schoemaker, and E. J. Russo (1990). *Decision Traps*.
Fireside.

We will use the JMP software package which is available from SAS and is installed in the Wharton computing labs. You may find it useful to install on your own computer at some point. The software with a book is available as

J. Sall and A. Lehman (1996). *JMP Start Statistics*. Duxbury, Belmont CA.

Grading

Grading for the course will come from a combination of in-class participation, a final exam, and several assignments.

Assignments	40%
Class participation	20%
Final exam	40%

Proposed Calendar

The following list gives, at best, a rough approximation of time that will be devoted to topics in the course. Typically, I distribute lecture notes *just before* the indicated lecture, or at the lecture. In either case, the notes will be available via the class Web page.

	Date	Topics	Reading
1	Jan 12	Overview of the course.	—
	14	Hypothesis testing, confidence intervals	Notes (or see FSW, *BBS*)
2	19	Combining confidence intervals	Notes
	21	Combining intervals without data	Notes
3	26	Multiple and dependent sources	Notes
	28	Bias in information sources	Notes
4	2	Picking the winner when comparing many groups	FSW, C 9
	4	Separating coincidence from systematic pattern	Dawkins
5	9	Regression: model specification and confounding	FSW, C 4,5
	11	Data mining	FSW, C 8
6	16	Regression: extrapolation and myths of interpolation	FSW, C 3
	18	Time series methods for forecasting (volatility)	FSW, C 12
7	23	Exponential smoothing and forecasting	Notes
	25	Probability forecasts and calibration	FSW, C 11
8	Mar 2	Utility, risk and certainty equivalent value	Notes
	4	Utility: principal agent and incentivizing	Notes
—	9	*Spring Break*	
—	11	*Spring Break*	
9	Mar 16	CAPM and beta	Notes
	18	”	
10	23	Hedging risk	Notes
	25	”	
11	30	Implied volatility and options	Notes
	Apr 1	”	
12	6	Subjective confidence intervals	Dec Traps
	8	Working with subjective intervals	
13	13	Foundations of game theory	Axelrod
	15	Cooperation and tit-for-tat	
14	20	Auctions	Notes
	22	Summary and closing review	