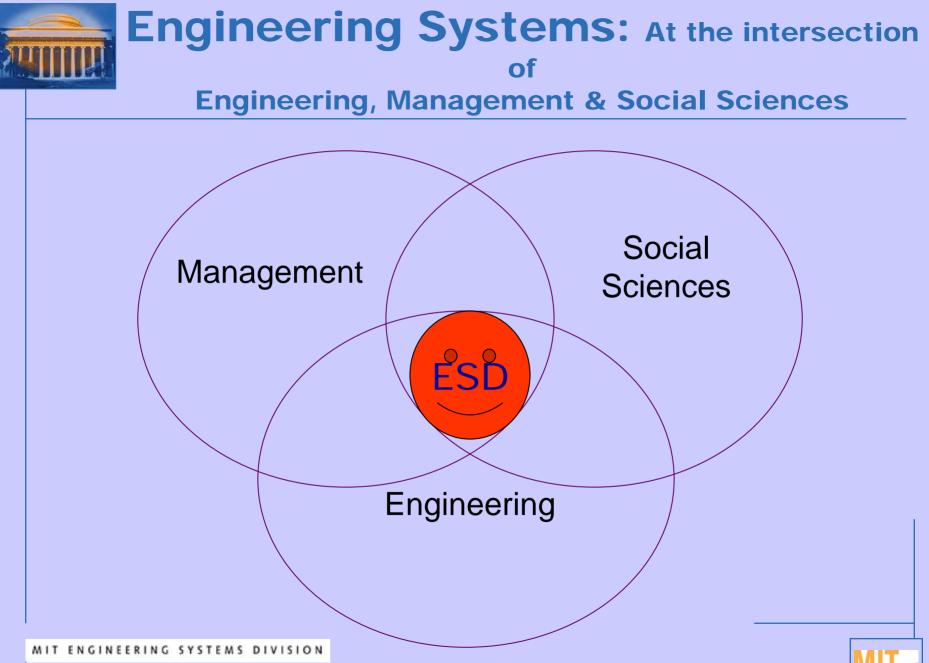


Design of an ESD Core Methodology Subject

Dick Larson, Dan Frey, with Roy Welsch









For the new Methods subject

We want to educate doctoral students to conduct research on these types of large scale engineering projects, which fall at the intersection of traditional engineering, management and social sciences





ESD.86 Models, Data, Inference for Socio-Technical Systems (New) Prereq: ESD.83, 6.041 G (Spring) 3-0-9

Use data and systems knowledge to build models of complex sociotechnical systems for improved system design and decision-making. Enhance model-building skills, including: review and extension of functions of random variables, Poisson processes, and Markov processes. Move from applied probability to statistics via Chi-squared t and f tests, derived as functions of random variables. Review classical statistics, hypothesis tests, regression, correlation and causation, simple data mining techniques, and Bayesian vs. classical statistics. Class project.

Enrollment limited to 25 students. Preference given to ESD Ph.D. Students.

Richard C. Larson, Daniel D. Frey





Overview

- > A new QUANTITATIVE methods subject, with aspects of each of three disciplinary areas: Engineering, Management and Social Science.
- > To be required of 1st year ESD Ph.D. students, spring semester
- > There will be another new subject on QUALITATIVE methods.





More Overview

- In ESD tradition, the 'math part' would be augmented with reading assignments and discussions tracing the history and application of each of the major concepts discussed and developed.
- > There would be a term project for each student.
- There would be computer-based as well as paper-based homework assignments. The subject would be rigorous.
- > While a byproduct would be continued 'class bonding' of the first year doctoral students, the primary focus is on intellectual content.





This is a Knowledge Requirement

- For students whose academic plan is to take MIT subjects that go much deeper than this subject (in statistics, probability, quantitative research methods), the requirement for taking this subject can be waived.
- > This subject represents a 'knowledge requirement' that will be assumed on doctoral general exams.





Building from...

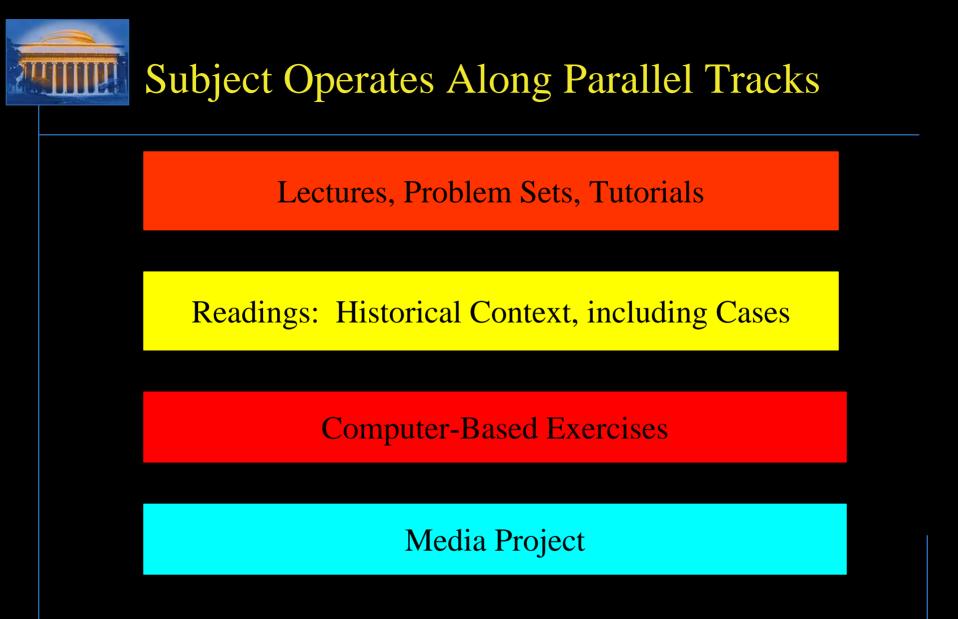
- Subject will have an enforceable prerequisite: 6.041 or equivalent.
- It will leverage all the fine work that Dan Frey has done with SOE curriculum development grant support -- funded in response to the oft-cited 'Odoni report' on the lack of a good solid engineering-focused statistics subject in the SOE.
- > But, this is NOT a statistics subject!





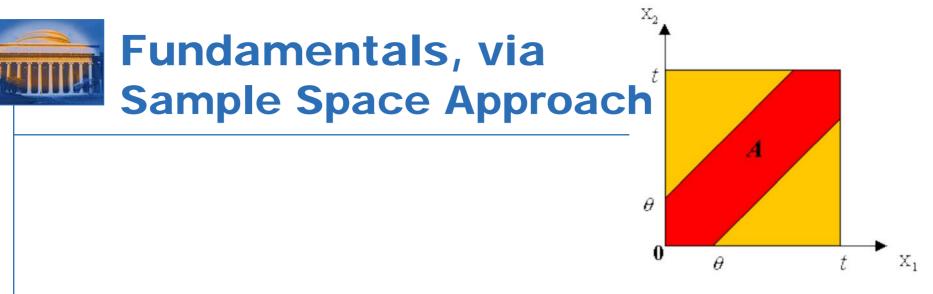


- If we are successful, we should attract students from elsewhere in the SOE who are not associated with ESD.
- > This should be ESD's first 'service subject.'
- > Tentatively, the 2007 spring semester new subject will be teamtaught by Dan Frey and yours truly, with a cameo by Roy Welsch.



Term Project





- Start with constructing probabilistic models using functions of random variables with a sample space approach.
- > Then we slide into statistics via experimental design with threats to validity, saying in essence that all designs are compromised by one or more of these.
- > Then, the statistical part should be a continuum of the sample space applied probability treatment so everything is fundamental -- no memorization of weird stat formulas, just for the sake of memorization.
- > We will do more with less in the stats area.
- > We cover Bayesian as well as classical statistics, highlighting the philosophy, strengths and weaknesses of each.
- > If they want a true stats course, that would follow this course.





Want students to be able to work with 'blank sheets of paper.'

They know fundamentals and can derive results.

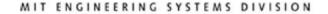
They are not just users of computer routines.





Go Deep, Use all Available Subjects

> We cannot think that ESD is so unique that no other MIT subjects can contribute to ESD students' knowledge of 'methods.' In the course of an ESD doctoral student's studies, she/he will rely most often on existing subjects at MIT or perhaps Harvard to go deep in the required methods.







Model Building

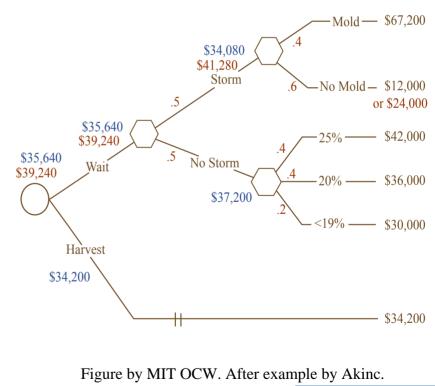
- > *Model building*, based on empirical evidence and axiomatic conjectures, should be the emphasis for the new subject.
- > We are not creating a new subject in applied probability nor are we creating one in statistics. But we use both to obtain our objective.
- > The focus is more on model synthesis, not data analysis *per se*.
- > It is an active model creating focus, not a passive critical social science focus.
- > Axiomatic models would be emphasized more than data inferred models, inferred from curve fitting -- where causation and correlation can become confused.





Introduce New Ideas in Homework

- Stochastic Dynamic Programming, Real Options, via Sequential Decision Trees
- > Shannon measure of information, Entropy (the element of 'surprise')
- Derivation of certain statistical tests (F, T, Chi-Squared)







Linkages to 'ilities..."

>Reliability

- Measures of..
- Systems designs with redundancy
- >Robustness
- > Predictability
- >Stability

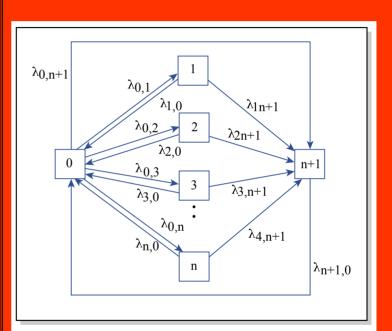
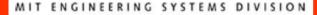


Figure by MIT OCW.

16







A Real Null Hypothesis: A Sports '.500' Team

- Each game is essentially decided by an independent flip of a fair coin
- Track the media coverage as certain expected 'streaks' during the year.
 - Wide use of derived distributions of Max and Min random variables.
 - Random incidence, potential fallacies in sampling



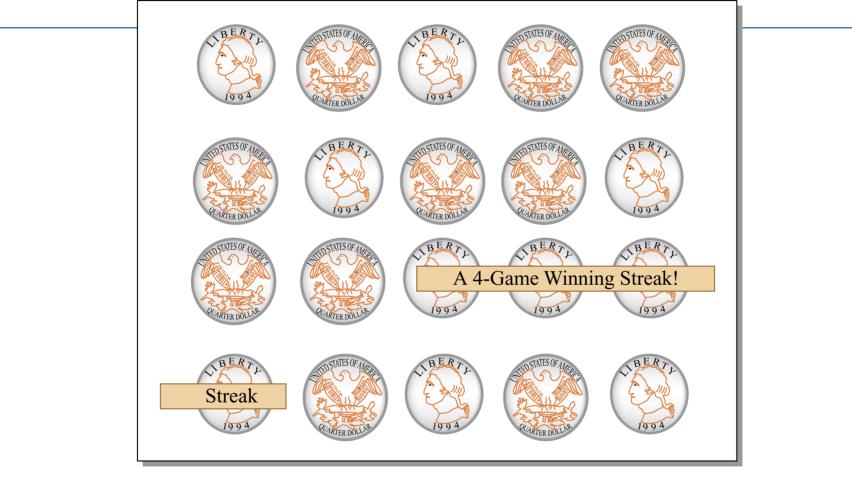


Figure by MIT OCW.







In a 162 game season, we should not be surprised to see

- > At least one 7 game loosing streak.
- > At least one 7 game winning streak. :)
- > All within the null hypothesis that each game is an independent fair coin flip.
- > But imagine the press coverage of these two events.
- > Generalize to more important topics.





On-Going Student Project

- > Track media weekly to find media mis-interpretation or misuse of data.
 - Statistical significance of the media phrase, "If it bleeds, it leads."
 - Making inferences based solely on sampling from extremes.



Class Projects





The End!

