STAT 400/ MATH 463

- Stat 400 introduce the basic concept about probability models and illustrate the usage of the probability models in the real world.
- Stat 400 will teach you about some basic statistics and how to use it to make a scientific inference.
- Stat 400 will illustrate how to understand some basic statistics and how to draw scientific conclusions based on those statistics
- Stat 400 will teach you how to model the random phenomenon using statistic models

Why probability and statistic are so important?

- Data is everywhere: the latest poll of the president's popularity, the average GPA and SAT scores of UI students…
- Inference is needed: is drug A better than drug B? How would economists and financial advisors make decisions from unemployment and inflation information…
- More facts: Politicians rely on data from polls of public opinion, business decisions are based on market research data, probabilistic weather forecasting…

Importance (Continued)

Two Famous Examples:

- Simpson's Paradox
- The case of Sally Clark

Simpson's Paradox

- The admission figures for fall 1973 showed that men were more likely than women to be admitted.

	Applicants	% admitted
Men	8442	**44%**
Women	4321	35%

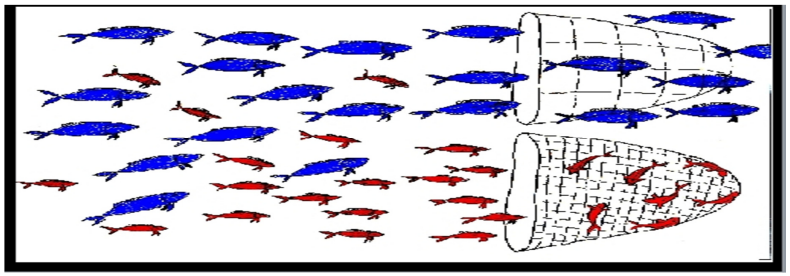
- P.J. Bickel, E.A. Hammel and J.W. O'Connell (1975). Sex Bias in Graduate Admissions: Data From Berkeley. *Science* 187: 398-40

Simpson's Paradox (Cont'd)

- Take a closer look

Major	Men		Women	
	Applicants	% admitted	Applicants	% admitted
A	825	62%	108	**82%**
B	560	63%	25	**68%**
C	325	**37%**	593	34%
D	417	33%	375	**35%**
E	191	**28%**	393	24%
F	272	6%	341	**7%**

Fishnet



- To illustrate this point, imagine a fishing boat with two different nets, a large mesh and a small net. A school of fish swim towards the boat and seek to pass it. The female fish try for the small-mesh challenge, while the male fish try for the easy route. The males go through and only females are caught. Judging by the final catch, preference toward female is clearly evident. However, if analyzed separately, each individual net would surely trap males more easily than females.

The case of Sally Clark

Sally Clark, a British woman who was accused in 1998 of having killed her first child at 11 weeks of age, then conceived another child and allegedly killed it at 8 weeks of age.
According to the expert witness, for an affluent non-smoking family like the Clarks, the probability of a single cot death was 1 in 8,543, so the probability of two cot deaths in the same family was around "1 in 73 million" (8543 × 8543).

Sally Clark (1964 –2007).
The case was widely criticized because of the way statistical evidence was misrepresented in the original trial.

What does Stat400 cover?

- Chapter 1-5: Probability.
  - probability,
  - mean, variance,
  - distributions (discrete, continuous),
  - independence,
  - Normal distribution,
  - central limit theorem etc
- Chapter 6-8: Statistics.
  - confidence interval,
  - hypothesis testing

Probability

• Probability evolve in the 16th-17th centuries from gambling and games of chance.

• Tossing a fair die or pulling a card from a well shuffled deck would have been the “experiments” that originally led to the formal theory of probability.

• Two basic concepts of probability:

– “Experiment” with uncertain outcomes

– “Sample space”, a collection of all the possible outcomes for an experiment.

Random Variable

Experiment can be described by random variable.

– Notation: X, Y, Z

– A random variable can only take numerical values with each value representing a possible outcome in the sample space.

– For each individual experiment, its value is uncertain

Two type of random variables:

- Discrete: The sample space has countable number of outcomes
- Continuous: The sample space is a interval with uncountable number of outcomes

Readings

- Reading: Chapter 1.1