NUMBERS IN EVERYDAY LIFE

Union College Academy for Lifelong Learning
April 2008

Necip Doganaksoy
Gerry Hahn
Ricki Lewis
Jane Oppenlander
Josef Schmee
Coordinator: Jim Comly

CLASS 1: SOME EXAMPLES AND BASIC CONCEPTS
Gerry Hahn and Jane Oppenlander

OUR GOAL

- **Provide insights into numbers and demonstrate their value**
- **Highlight the abuses**

CLASS MOTTO: Numbers are highly useful, but can be readily abused—assess with care!

Copies of today's slides available from gerryhahn@yahoo.com

COURSE PLAN

- **April 3: Some examples and basic concepts (Gerry and Jane)**
- **April 10: Public opinion polls and election forecasts (Josef)**
- **April 17: Medical and health studies (Ricki)**
- **April 24: Business and industrial applications (Necip)**
- **May 1: Some further examples and wrap-up (Gerry and Jane)**




WHAT COURSE WILL NOT COVER

- **Accounting numbers**
- **Various government statistics (e.g., U.S. census)**
- **Statistical methods**

INVITATION

- **SEEK INTERESTING NUMBERS FROM THE MEDIA**
- **SEND US (OR BRING ALONG) FOR POSSIBLE CLASS DISCUSSION**



TODAY'S TALK TOPICS

Impact of NY Indoor Clean Air Act	Before and after comparisons
Impact of Marriage on Longevity	Observational Study, Correlation vs. Cause and Effect
Impact of Prayer on Outcome of Bypass Surgery	Controlled Study, Relative Risk, and Statistical Confidence
Birthday Exercise	When bell-shaped curve applies
Impact of 2005 Medicare Drug Plan	On the average: Mean vs. Median
Advanced Scout: An NBA Coach's Assistant	Data Mining

EXAMPLE: IMPACT OF NY STATE CLEAN INDOOR ACT


- WAMC September 27, 2007: "Fewer New Yorkers have been treated for heart attacks since the State's wide-ranging no smoking law took effect four years ago. That's according to the State Health Department which says that since the Clean Indoor Air Act began in July 2003 the number of New Yorkers admitted to the hospital for a heart attack dropped by 8%."
- Implication: Act reduced heart attack rate by 8%
- Some similar examples:
  - Impact on accident rate of changing speed limits
  - Impact of 1996 welfare reform law (Best 2004, Chapter 6)
  - Impact of change in Daylight Savings Time

BASIC CONCEPTS

- Be wary of advocates with numbers
- Watch out for "data cherry picking"
- Beware of simple before/after comparisons over time
- Difficult to separate impact of action from other factors because
  - Typically "other factors" not known
  - Even if known, not measured
- When known and measured, might be controlled or removed "statistically", but
  - Statistical methods complex and imperfect
  - All "confounding" factors may not have been found
- Possible starting point for further research

EXAMPLE: IMPACT OF MARRIAGE ON LONGEVITY


- Claim: Marriage increases longevity
- Implication: Get (and stay) married to live longer
- Source: Studies reported in Waite and Gallagher (2000): The Case for Marriage: Why Married People are Happier, Healthier and Better off Financially

SOME OF THE NUMBERS (from various studies)


- "The non-married…have higher rates of mortality than the married: about 50% higher among women and 250% higher among men."
- "Divorce seemed to be as dangerous to a man's health as picking up a pack-a-day cigarettes habit."
- "For men, staying married boosts the chances of surviving to age 65 from about two out of three, to almost nine out of ten; for women, wedlock ups the likelihood …from about 80% to more than 90%."
- "Being unmarried can actually be a greater risk to one's life than having heart disease or cancer. For example, having heart disease shortens the average life span by slightly less than six years. But being married chops almost ten years off a man's life."
- "Not being married will shorten a woman's life span by more years than would being married and living in poverty."
- "Almost nine out of ten married men alive at age 48 would still be alive at age 65. By contrast, just six out of every ten married men alive at age 48 would make it…;divorced and widowed men were almost as likely as confirmed bachelors to die before age 65."
- "Nine out of ten married women at age 48 reached age 65, compared to about eight out of ten never-married and divorced women."

THE BIG QUESTION

- "Is really marriage (the reason)…rather than a product of…selection? Perhaps healthier people are more likely to marry…"
- Problem with observational studies: Other factors may be "confounding" results.
- Some "counteracting" measures
  - Compare otherwise similar people
  - Try to adjust for other factors statistically
- Search for explanation
- Authors conclude "marriage itself moves people to a healthier way of life" (e.g. "the virtues of nagging")
- Are they right?

SOME SIMILAR EXAMPLES

- An old favorite: Houses with storks nesting have more babies

- Children with bigger feet spell better (Paulos, 1991)
- Dog owners have lower cholesterol
- Tall people earn more money

BASIC CONCEPTS

- Don't confuse statistical relationship (correlation) from an observational study with cause and effect
- Observational studies can be improved by
  - Case-control: Matching otherwise similar subjects
  - Statistical adjustment for other (identified and measured) subjects

  and search for explanation
- Still unlikely to yield definitive results

GOLD STANDARD IS CONTROLLED, RANDOMIZED STUDY

BUT—OBSERVATIONAL STUDIES MIGHT STILL BE USEFUL

- Provide insights and hypotheses
- Example: Framingham Heart Study

  - Ongoing since 1948
  - Suggested link between smoking and lung cancer
- Suggest further research
- Main interest is in prediction (rather than understanding), e.g.,
  - Epidemiological studies
  - Weather forecasting
  - Credit scores

EXAMPLE: IMPACT OF PRAYER ON OUTCOME OF BYPASS SURGERY

- Headline, Daily Gazette, March 31, 2006: "Study finds prayer may make patients worse."
- Basis of results: "$2.8 million study…in the American Heart Journal (AHJ)."
- Text: "Doctors who followed 1,800 heart bypass patients at six medical centers found that those who *knew* they were being prayed for suffered higher rates of complications than others who weren't sure."
- Implications:
  - Based on headline: Prayer may make patients worse.
  - Basis of text: *Knowing* someone is praying for you may make patients worse.

QUESTIONS AND ANSWERS

- What patients? Coronary artery bypass graft surgery (CABG)
- How was study conducted? 600 patients randomly assigned to
  - Group 1: Told they *might be* prayed for and *were*
  - Group 2: Told they *might be* prayed for and *were not*
  - Group 3: *Promised* they would be prayed for and *were*
- What were some study characteristics/limitations (per Internet news release)?
  - Six medical centers
  - "Complications" after CABG surgery
  - Distant prayer by two Catholic and one Protestant prayer group
  - Only patient's first name and last initial provided to person praying
  - Same standardized prayer
  - $2.8 million study supported by the John Templeton Foundation ("philanthropic catalyst for research on concepts and realities such as love, gratitude, forgiveness and creativity")
- Results: "Complications" following surgery
  - Group 1 (told they *might be* prayed for and *were*): 52%
  - Group 2 (told they *might be* prayed for and *were not*) : 51%
  - Group 3 (*promised* they would be prayed for and *were*): 59%

RELATIVE RISKS

- Definition: Ratio of probability of an event occurring in a "treatment group" versus a "control group"
- *Estimated relative risk* (of complications after CABG surgery) *of being prayed for* (after being told "might be prayed for"): 0.52/0.51 = 1.02 (Group 1 vs. Group 2)
- *Estimated relative risk* (of complications after CABG surgery) of being *informed of being prayed for* versus being told *might be* prayed for (and then prayed for): 0.59/0.52 = 1.13 (Group 3 vs. Group 1)
- What is statistical "margin of error?"
- Are results statistically significant?
- Technical time out

95% CONFIDENCE INTERVAL
(Statistical "margin of error")

- Working definition: Range calculated from random sample to include quantity of interest (e.g., relative risk) with 95% probability.
- 95% confidence interval on relative risk *of being prayed for*: 0.92 to 1.15 (from AHJ article)
- 95% confidence interval on relative risk of *being promised prayer*: 1.02 to 1.28 (from AHJ article)
- Result is deemed "statistically significant" (versus "statistical dead heat") if confidence interval *excludes* 1.0
- Want to be *surer*?
  - Calculate 99% confidence interval (but increase interval length)
  - Increase sample size (and decrease interval length)
- Note: Confidence interval deals only with *statistical* uncertainty

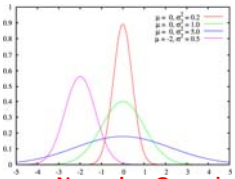
REVISITING ARTICLE HEADLINE

- Headline "Study finds prayers may make patients worse" is misleading.
- More precise headline: "Study finds that prayer may make patients better or worse or not have any impact."
- More realistic headline: "Study finds promising patients prayer may make them worse."

BASIC CONCEPTS

- Read beyond headline
- Find out
  - How study was conducted (e.g., controlled versus observational)
  - Limitations
- Refereed journal adds credibility
- Relative risk compares "treatment" versus "control."
- Margin of error (confidence interval) measures uncertainty due to sampling (only)

THAT MAGICAL BELL-SHAPED CURVE



- Also known as Normal or Gaussian distribution
- But how "normal" is normal distribution?
- Justification: Sum (or arithmetic mean) of "many" small independent factors
- Class exercises:
  - Counts (or histogram) of three birthdates within month
  - Arithmetic mean of three birthdates

BASIC CONCEPTS

- Justification for normal distribution: Sum (or arithmetic mean) of "many" small independent factors
- Some examples:
  - Points scored by a basketball team
  - Heights of U.S. adults
  - IQ scores of 4th graders

- But *not everything* is sum of many small independent factors; e.g.,
  - Life of humans (and many products)
  - Household income

BEWARE OF "ON THE AVERAGE"

- Consider five families with yearly incomes of $30K, $30K, $50K, $100K and $2 million
- What is their average income?
  - Mean (arithmetic average) = $(30+30+50+100+2,000)K =$442K
  - Median (middle ranked value) = $50K
  - Mode (most frequent value) = $30K

EXAMPLE: IMPACT OF MEDICARE DRUG PLAN

(H.Wainer, *Visual Revelations,* Chance Magazine, Spring 2006)

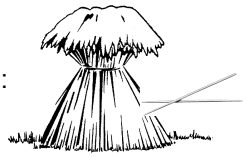
- President Bush, June 2005: "On the average the folks who sign up for (the new) prescription drug are going to save $1,300 a year"
- Looking at the numbers: Need $5,137 yearly drug costs to save $1,300
- President was referring to *mean savings*
- *Median savings* much lower

BASIC CONCEPTS

- **Make sure that average is defined**
- **Interpret results accordingly**
- **Median often more useful than mean**

WHAT IS DATA MINING?

- **Automated** process of collecting and **analyzing large volumes** of data to find **hidden patterns**.



- Three basic steps:
  - Data Preparation
  - Data Processing
  - Interpretation and Knowledge Discovery

EXAMPLE: DATA MINING IN THE NBA

- **Advanced Scout: Data mining software to help coaches analyze game data.**
- **Helps coaches formulate game strategies and assess effectiveness of decisions.**
- **First used in 1995-96 season.**
- **Available to ALL National Basketball Association teams.**

Play-by-play Sheet

Houston Rockets		Los Angeles Lakers
(12:00) Jump Ball Brown vs Yao		
	11:42	Walton Jump Shot: Missed
	11:41	Brown Rebound (Off:1 Def:0)
	11:29	Bryant Jump Shot: Missed
Alston Rebound (Off:0 Def:1)	11:28	
Yao Jump Shot: Missed	11:06	
	11:05	Turiaf Rebound (Off:0 Def:1)
	10:42	Turiaf Jump Shot: Missed
Yao Rebound (Off:0 Def:1)	10:41	
Battier 3pt Shot: Missed	10:25	
	10:24	Turiaf Rebound (Off:0 Def:2)
	10:08	Walton Jump Shot: Missed

How big is NBA data mine?

For 1 season:

- **30 teams in the NBA**
- **15 players per team**
- **82 games/team (not counting playoffs)**
- **~ 100 points/team/game**
- **48 minutes/game (not counting overtime)**

For one team's season approximately 41,000 plays

"Statistics always speak in a loud tone"

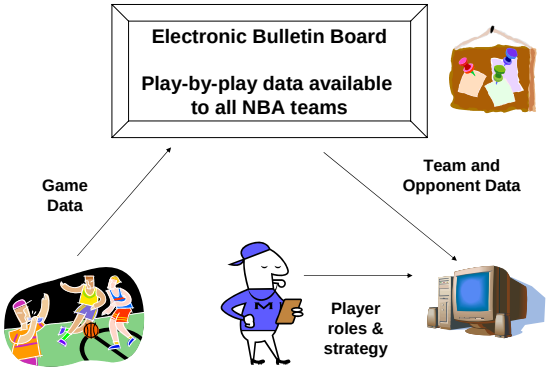
Bo Ryan (coach of the University of Wisconsin men's basketball team)

- **Approximately 65 basketball statistics**:
  - AST = Assists
  - BLK = Blocks
  - FG = Field Goals
  - PER = Player Efficiency Rating
- **Formula for unadjusted Player Efficiency Rating (overall measure of a player's performance)**

Circa 1951, developed by John Hollinger, ESPN

uPER = (1 / MP) * [3P + (2/3) * AST + (2 - factor * (team_AST / team_FG)) * FG + (FT *0.5 * (1 + (1 - (team_AST / team_FG)) + (2/3) * (team_AST / team_FG))) - VOP * TOV - VOP * DRB% * (FGA - FG) - VOP * 0.44 * (0.44 + (0.56 * DRB%)) * (FTA - FT) + VOP * (1 - DRB%) * (TRB - ORB) + VOP * DRB% * ORB + VOP * STL + VOP * DRB% * BLK - PF * ((lg_FT / lg_PF) - 0.44 * (lg_FTA / lg_PF) * VOP)]

Step 1: Data Preparation



Step 2: Data Processing

- **How does it work?**
  - Finds the overall distribution of an attribute (for example, field goal attempts)
  - Searches for subsets of data that have a different pattern from the overall distribution
  - Patterns are sorted in order of their interestingness

Step 3: Knowledge Discovery

Advanced Scout Displays:

- **An Interesting Pattern**
- **Explanation of why it's interesting**

The Coaches:

- **Determine the underlying cause of the pattern (often by watching the game video)**
- **Determine how to incorporate the information into the game strategy**

1997 Eastern Conference Playoff First Round

Miami  **Orlando**

Advanced Scout used by Orlando discovered:

- **While guards Penny Hardaway and Brian Shaw were on the floor, Orlando was losing by as many as 17 points.**
- **While backup guard Darrell Armstrong replaced Shaw, the team had a 14 point advantage.**

The Coaches:

- **Increased Shaw's playing time dramatically in the next 2 games.**

Bottom Line

- **Orlando won the 3rd and 4th games before losing to Miami in the 5th game.**
- **Miami lost to Chicago 4-1 in the Eastern conference finals**

"By helping us make better decisions, Advanced Scout is playing a huge role in establishing incredible fan support and loyalty – that means millions of dollars in gate traffic, televisions sales and licensing."

-Tom Sterner, assistant coach of the Orlando Magic

ELEVATOR SPEECH: CLASS 1

- **Watch out for "data cherry-picking"**
- **Beware of simple before/after comparisons**
- **Don't confuse statistical relationship from observational study with cause and effect**
- **Controlled, randomized study is "gold standard" (but often impractical)**
- **Find out how numbers were obtained/defined and limitations of study--be wary of number advocates**
- **Margin of error quantifies statistical uncertainty**
- **Bell-shaped curve applies often-but not always**
- **Know which average you are dealing with**
- **Careful data mining can provide important leads**

NUMBERS ARE HIGHLY USEFUL, BUT CAN BE READILY ABUSED—ASSESS WITH CARE!

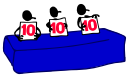
NUMBERS IN SPORTS

- **Baseball:**
  - A numbers person's dream
  - Some reflect variables beyond a player's ability
  - No "super-statistic"
- **Game strategy:**
  - Baseball: The deliberate walk
  - Football: Go for 4th down
- **Front office strategy:** Building the Oakland A's on a shoestring (per M.Lewis: Moneyball)

EXAMPLE: QUANTIFYING THE SUBJECTIVE—COLLEGE RANKINGS

- U.S. News and World Report provides yearly college rankings
- Union ranks 40th among Liberal Arts colleges
- RPI ranks 44th among 262 universities
- Similar issues arise in ranking
  - Other service providers, e.g., hospitals
  - Employees
  - Movies

CRITERIA



- Peer assessment: 25%
- Graduation and retention rate: 20%
- Student selectivity (SAT/ACT scores, high school standing, acceptance rates): 15%
- Faculty resources (Class size, compensation, top degree, %full-time, student/faculty ratio): 20%
- Financial resources: 20%
- Alumni giving: 5%
- Graduation rate: 5%

DATA SOURCES

- Questionnaire sent to colleges
- Peer reviews (college presidents, provosts and deans of admission)

ISSUES AND CONCERNS

(See Best, 2004)

- What do we mean by "best?"
- Are the right criteria and weights being used?
- Emphasis on what can be measured
- Incentive to colleges to "game" the system

BASIC CONCEPTS

- Einstein: Not everything that can be measured is important, and not everything that is important can be measured
- Numerical rankings of
  - Service providers is difficult
  - Products sometimes less difficult
- Need to ask: How were rankings developed?
- Relate to *your* value system