Statistics and the Design of Experiments & Surveys [Guest Lecture for EPOM 409]

Thomas E. Love, Ph. D.
Director, Biostatistics & Evaluation Unit
Center for Health Care Research & Policy
CWRU-MetroHealth Medical Center
CWRU Department of Medicine and
Department of Operations, Weatherhead
778-1265 or TEL3@po.cwru.edu
www.chrp.org/love

What am I doing here?

- "How to perform experiments and analyze data as an integral part of your job regardless of how much statistics you've actually learned"
- "How to solve engineering problems that, incidentally, require statistical thinking and methods"
- Is Statistics orthogonal to Engineering?

What do engineers do?

- Develop new products
- Improve previous designs
- Build and test prototypes
- Design tools, machines and processes to make and test products
- Maintain, control, troubleshoot and improve ongoing manufacturing processes
- Maintain, service and repair products

Engineers collect and analyze data!

- Statistics includes the art and science of collecting data, through designed experiments, or through sampling.
- Whether or not engineers have learned statistics – they will do statistics.
- Those that have learned statistics are likely to be more effective engineers.
- Statistical methods are not an adjunct to, but are an inseparable part of how engineers solve engineering problems.

Role of Statistics in Industry

- Statisticians have contributed much to the design, analysis and execution of industrial experiments and sample surveys.
- Statistical tools help with efficient data acquisition, analysis, and presentation.
- These tools allow us to collect data that are free of bias and hence are useful in drawing proper conclusions.

Product Sampling

- In a manufacturing operation, we sample batches of product, inspect and measure to create a *control chart*.
- Quality characteristic plotted over time.
  - Upper & Lower Control Limits, Center Line
  - As long as chart indicates that the process is operating normally, the process continues.
  - When the chart signals trouble, referred to as a *special cause* of variation, action is taken.
- Ensures minimal interruption, while providing for timely detection of trouble.

Use of a Statistical Model

- A successful model yields understanding of the fundamental causes of process variation.
- Ideally, a model yields accurate predictions of the future behavior of the process -- allowing for process *control*.
  - Control can mean keeping process outcomes near a predetermined "target"
  - Make the process operate more efficiently, produce superior output, or both

Every process can be improved.

- Improvement comes about by:
  - elimination of unnecessary steps
  - use of better equipment
  - development of better definitions of work required
  - involvement of workers in decision making
  - redesign of products/services to better meet customer needs
- But to improve, we need information...

Methods of Data Collection

- Collecting data can be time consuming and expensive. One need not collect every possible piece of data to make decisions.
- Sampling methods are *designed* to gain maximum information at minimum cost.
  - Product Sampling/Observation [Passive]
  - Sample Surveys [Passive]
  - Designed Experiments [Active Learning]

Steps in Data Collection

- State the purpose for collecting data
- Determine sources,
- Determine data capture and presentation methods
- Train personnel
- Collect the data accurately
- Document the work

The Big Issues in Data Collection

- The extent to which a given set of data can be relied on depends on its quality:
- Data Design Quality
  - Are the data relevant to the problem we wish to solve?
- Data Production Quality
  - Were the data collected with sufficient skill and care?

Sample Surveys

- Most common data collection approach
- *sample* of elements selected from a *universe*
  - *universe*: collection of all elements of interest
  - *sample*: collection of elements actually observed and studied by collecting data
- Opinion Polls
  - Random Digit Dialing & "Non-response" Bias
- Personal Interviews

Getting A Good Sample

- Non-probability surveys do not use probability methods to select the sample from the universe
  - Questionnaire sent to entire universe (Volunteers)
  - Use of Judgment samples is also problematic.
  - Often seriously mishandled
- In Probability surveys, each element has a known probability of being in the sample
  - Specification of a Sampling Frame

It's All in the Wording

- Deliberate Bias
  1. Do you agree that abortion, the murder of innocent beings, should be outlawed?
  2. Do you agree that there are circumstances under which abortion should be legal, to protect the mother's rights?
- Unintentional Bias
  - How many doors are in your home?
- Desire to Please
- Unnecessary Complexity
- Ordering of Questions

Survey Shortcuts to Avoid

- Failure to use a proper sampling procedure
  - Should avoid use of convenience samples
- No pre-test of the field procedures
  - Only way to find out if everything "works"
- Failure to follow up with non-respondents
  - Common for initial response rate to be < 50%
- Inadequate quality control
  - Sloppy execution can seriously damage results

Common Survey Problems

- Sampling and Non-Sampling Errors
- Difficulties
  - Using the wrong sampling frame
  - Not reaching the individuals selected
  - Getting no response or a volunteer response
- Disasters
  - Getting a volunteer sample
  - Using a convenience or haphazard sample

Designing an Experiment

- Consider a cake mix, whose instructions say, "Bake at 325 degrees Fahrenheit for 60 minutes"
- Oven and timer settings vary, so the mix must bake satisfactorily at *ranges* of temperatures and times.

What has to happen?

- Need to:
  - choose various combinations of temperature (300 to 350) and time (55 to 65) settings,
  - bake cakes at each combination,
  - submit cakes to a taste test yielding numerical scores,
  - evaluate the results for evidence that the mix performs well, or fails to perform well, over the combinations of settings.

Cake Mix Example Design

- Possible strategy: study recommended times and temperatures, and extremes of ranges.
- Table of Treatment Combinations:

	55 minutes	60 minutes	65 minutes
300°F	X	X	X
325°F	X	X	X
350°F	X	X	X

X's mark the nine possible combinations

Implementing the Design

- Use an oven & timer that provide extremely precise and accurate settings.
  - Do not try to mimic what actual users do.
- Develop a protocol for preparing the mix and the oven so that cakes are prepared under essentially identical conditions.

Implementing the Design

- Prepare the cakes at combinations chosen in a random order.
  - If there are any carry-over effects, a systematic move through the table of combinations can be disastrous.
  - Random selection of combinations creates a sound theoretical basis for the use of statistical analysis.
    - Tables of random digits are available.

Implementing the Design

- Make sure that the taste testers don't know the experimental conditions under which the cake was baked. This is *blinding*.
- Make sure the people conducting the taste test don't know either. (*Double blinding*)
- Offer cakes to taste testers in random order.
- Recruit taste testers who can truly discriminate differences between cakes.

Factorial Experiments

- Experiments in which all possible combinations of the treatment levels are run in a replication. Several replications are done in a factorial experiment.
  - Most efficient procedure is to vary several factors (here, time and temperature) at once, instead of one at a time.
- Cake mix experiment thus uses 3 x 3 = 9 combinations in each replication.

Comparisons

- How do we separate data into categories?
- Before-and-after studies
  - Pre-test and Post-test measure of value of a new training program
  - Measure of Quality tracked before and after an innovation/intervention.
- Stratification by gender, age, region, price
- Control Groups vs. Experimental Groups

Gastric Freezing Example

- (1958) New procedure for treating ulcers
  - Paper reported results on 24 patients, but the results were not compared to a control group
  - Procedure looks very successful
- (1963) Experiment -- Random Assignment
  - 82 patients receiving gastric freezing treatment,
  - 78 receiving non-treatment (placebo) disgused so patients thought it was real
  - Procedure shown to be of little or no value

Basic Principles of Designed Experiments

- Control
  - needed to make explicit what factors are allowed to have an effect on the response
- Randomization
  - needed to protect against biases due to the effects of unsuspected/uncontrolled factors
- Replication
  - needed to reduce effects of natural variation to allow experimental factor effects to be detected

What Happens?

- Lack of Control over factors, especially the lack of a control group:
  - Desired comparisons cannot be made without additional analytical adjustments
- Lack of Randomization or lack of blinding:
  - A lurking variable may be the true cause of relationships observed in the data
- Lack of Sufficient Replication
  - Natural variation may mask real effects of experimental factors

Is statistics just applied probability?

- What is important for the science of statistics shouldn't be judged by how mathematically complex it is…
  - Ex: emphasis on acceptance sampling in QC
- Graphics and EDA need little math but are extremely important.
- Ishikawa's Seven Tools are trivial mathematically, incredibly valuable

Real Statistical Work

- Let X_1, X_2, … be independent, identically Normally distributed random variables…
- Lots of practical considerations that can only be learned by doing statistics
- Universities teach large quantities of facts – in statistics, that's techniques, esp. hypothesis testing techniques. More is less here.

Statistical Tool Box

- Data Description and Summarization
- Estimation of Effects through Confidence Intervals (testing)
- Analysis of Variance (ANOVA) -- main tool for comparing averages in experimental design
- Categorical Data Methods -- often useful in analyzing survey results

Summarizing Survey Data

- Tallies
- Two-Way (and Multi-Way) Tables
- Measures of Location (Average) and Variation
- Distribution Displays
  - Histograms
  - Stem-and-leaf plots
  - Scatterplots

Analysis of Variance

- Compares the sample means to see if differences between populations are statistically significant.
- Specification of Factors of Interest
- ANOVA Table and F test
- Interaction
- Methods for Multiple Comparisons

A Common Fallacy

- Can everything you need to know fall out of pure deductive reasoning?
  - Newton's Three Laws
  - Maxwell's Equations
  - Kirkhoff's Laws
  - Three Laws of Thermodynamics
- All originated from experiments in an iterative process of induction - deduction

Reality - Experimentation

- Once you leave, you'll have to confront reality and experiment. Make sure you're well equipped to do so.
- Experiments from which NEW knowledge is generated, not demos of already-known phenomena
- How do we construct valid and efficient experiments, collect data, analyze the results, and build models?

"All models are wrong, some are useful."

- Statistics is about inductive reasoning, learning from the real world, detective work, and experimental design.
- Statistics in the hands of a large community of engineers can have a dramatic influence on the future of modern manufacturing and engineering.

A Few References

- Moen, Nolan & Provost, *Improving Quality Through Planned Experimentation*
- Box, Hunter & Hunter *Statistics for Experimenters*
- http://www.sematech.org/public/resources/stats/
- http://www.execpc.com/~helberg/statistics.html
- http://lib.stat.cmu.edu/
- http://www.amstat.org/sections/SPES/
- http://web.utk.edu/~wparr/qplinks/qplinksalpha.html
- http://www.asq.org/