Making Claims & Arguments with Statistics

Claudia J. Stanny
PSY 6217 – Research Design

University of West Florida

Importance of Explicit Comparisons

- Isolated data are difficult to interpret
- Explicit comparisons create a context in which to interpret data
- Effects of one's choice of comparison
  - *Directs attention to specific types of explanation*
  - *What is the appropriate comparison to evaluate the observation of a long life expectancy for conductors?*
    - *Adults?*
    - *Adults who have lived to at least age 30?*
    - *Adults who are musicians?*
    - *Adults who have lead active lives?*

Cognitive Biases in Interpretation of Findings

- Underestimate the amount of variability created by random variation
  - *Bias to discover ordered patterns*
  - *Exaggerated perception of systematic effects*
  - *Overconfidence in our ability to predict future outcomes*
- Failure to appreciate the increased variability of findings based on small samples
  - *False belief in the "law of small numbers"*

Rival Explanations Proposed During Hypothesis Testing

- All variation is systematic variation
  - *Predicts differences in group means*
  - *Predicts no variability in observations within groups*
  - *Validity of this explanation for observations in biology, chemistry, and physics*
- All variation is random variation
  - *Predicts variability in observations within groups*
  - *Predicts no differences between group means*
- Variation is a mixture of random and systematic variation

Problems in Statistical Decision Making

- Explanation based on random variation alone is almost never (technically) true
  - *The means of any two groups are almost never* ***exactly*** *identical*
  - *How large must the differences between means be to be considered different?*
- Burden of proof that the difference is "large enough" to discredit the "randomness alone" explanation falls on the researcher
- Statistical analysis supports arguments that are essentially logical and rhetorical in nature

Statistical Decisions are Ambiguous

- No decision is guaranteed to be true
- Decision to reject the explanation based on random variation alone can be made in error
  - *Type I Error*
  - *Estimated by level of significance (α) or the p value of the test*
- Decision to *not* reject the explanation based on random variation alone can be made in error
  - *Type II Error*
  - *Estimated by β*

Limited Value of an Isolated Finding

- *One swallow does not make a summer.*
- One statistically significant finding is not enough to build a persuasive argument
  - *Doubters can attack the methodology*
  - *Doubters can suggest other sources of systematic variation to explain the finding*
- Abelson's MAGIC criteria for strong arguments
  - *Magnitude*
  - *Articulation*
  - *Generality*
  - *Interestingness*
  - *Credibility & Coherence*

Explanations Based on Randomness

- Random Generation of Data
- Random Sampling of Data from a Distribution
- Random Assignment of Sampled Individuals to Groups
  - *Relevant when the individuals have not been randomly sampled*

Random Processes for Data Generation

- Binomial process
  - *2 outcomes*
    - *Success*
    - *Failure*
  - *Each outcome has a constant probability*
  - *Trials are independent*
- Patterns of outcomes of a binomial process
  - *P(success) for a large set of trials equals P(success) for a single trial*
  - *Independent trials allow us to estimate the probability of runs of various lengths for a given outcome*

Biases in Thinking about Random Generation Processes

- **Gambler's fallacy**
  - *Assumes that random processes are self-correcting*
  - *Error is that processes have no "memory" for prior trials*
  - *Trials are independent*
- **Representativeness Heuristic**
  - H H T H T T H T
  - H H H H T T T T
  - *Long runs of identical outcomes violate our stereotype of a "typical" random process*
  - *Random processes are lumpy and create long runs*
  - *These are mistakenly attributed to systematic patterns*

Random Processes for Sampling from a Distribution (Population)

- **Need for a representative sample from a population**
- **Characteristics of Random Samples**
  - *Each element has an equal probability of being selected*
  - *All samples of a given size are equally likely to be selected*
- **People underestimate the variability of populations and of samples**
- **People are overconfident in the representativeness of small samples**

Effect of Sample Size on Variability of the Sample Mean

	A			B	
A	A			B	B
A	A			B	B
1	2	3	4	5	6

- **Samples of size 5**
- **Mean for Sample A = 1.6**
- **Mean for Sample B = 5.4**

Samples of Size 10

		A	B		
	A	A	B	B	
A	A	A	B	B	B
A	A	AB	AB	B	B
1	2	3	4	5	6

- Mean of Sample A = 2.4
- Mean of Sample B = 4.6

Samples of Size 10

		AB	AB		
	A	AB	AB		
A	AB	AB	AB	AB	B
A	AB	AB	AB	AB	B
1	2	3	4	5	6

- Mean for Sample A = 3.07
- Mean for Sample B = 3.90

Explaining Differences between Sampled Groups

- Explanation based on random error
  - *Sampling error*
  - *Random processes can generate non-representative samples (especially when sample size is small)*
- Explanations based on systematic error
  - *Sampling bias*
  - *Non-random processes in selection*
    - *Attrition, non-volunteers, different parts of the population are available for sampling at different times, etc.*
- Explanations based on systematic manipulation of variables

Non-Random Selection of Samples

- Group might not be representative of the population sampled
- Random assignment to groups enables tests of random versus systematic explanations for differences between the groups
  - *Random assignment to groups*
  - *Differences explained by random assignment process*
  - *Differences explained by systematic differences in treatment of the groups created*

Dealing with Ambiguous Data

- What happens when groups differ but the statistical test has a p value > .05?
- Consider the sensitivity of your experiment to detection of effects (power)
  - *Could this be a Type II Error?*
- Acknowledge that the effect, if it exists at all, is a weak effect
  - *Is this effect worth pursuing with additional research?*
- Delay publication until additional research that resolves the ambiguity can be completed