

Psych 218 Introduction to Behavioral Research Methods

Week 6: Lecture 11

Outline of Today's Lecture

- Last lecture we discussed
 - Reactivity and Demand Characteristics
 - Types of non-experimental Research
- Today we will discuss
 - Establishing reliable measures for non-experimental research
 - Sampling

Establishing Reliable Measures for Non-Experimental Research

- Non-experimental research requires the definition of specific **behavioral categories** or **recording units**
- Categories should be
 - based on hypotheses, informal observations, literature search
 - developed before the behavior is observed or the archival data is analyzed
 - simple and focused on specific behaviors or archival content
 - exhaustive, mutually exclusive, independent

Establishing Reliable Measures for Non-Experimental Research

- Behaviors within each category should be quantified using one or more of the following methods
 - Frequency method: number of times behavior occurs
 - Duration method: how long a behavior lasts Intervals method: does behavior occur within discrete time intervals?
 - **Behavior sequences**: keep track of order of behavior in addition to frequency

Establishing Reliable Measures for Non-Experimental Research

- Problem: Often behavior is complex and occurs too quickly to both observe and record at the same time
- Sampling methods
 - Time: alternate observing and recording periods
 - Individual: observe and record only one individual at a time
 - Event: observe and record only one behavior at a time
 - Recording devices (video, audio)

Establishing Reliable Measures for Non-Experimental Research

- Problem: any single observer (or content analyzer) might be biased or their observations might be idiosyncratic or unreliable
- Solution: use multiple observers and quantify their differences by computing interrater reliability

Establishing Reliable Measures

- Statistical methods for computing interrater reliability
 - Percent agreement = $100 * \text{Agreements} / \text{Observations}$
 - Cohen's Kappa (K) = $(P_o - P_c) / (1 - P_c)$
 - P_o is the actual agreement and P_c is the agreement you would expect by chance

Confusion Matrix

		observer 1		
		angry	loving	
observer 2	angry	10	5	15
	loving	3	8	11
		13	13	26

$$P_o = (\text{cell11} + \text{cell22}) / N = (10 + 8) / 26$$

$$P_c = (\text{row1} * \text{col1} + \text{row2} * \text{col2}) / (N * N) = (15 * 13 + 11 * 13) / (26 * 26)$$

Establishing Reliable Measures

- Statistical methods for computing interrater reliability
 - Pearson's product-moment correlation
- Observer bias
 - Blind observers
 - Objective vs. interpretive recording

Sampling

- Why sample? We cannot usually measure the entire population of interest so we must rely on measuring a sample of the population
- Goal of Sampling: to be able to generalize to everyone in the population of interest – sample must represent the population to insure external validity
- Terms
 - population: any group with size greater than 1
 - element: one member of a sample, e.g., person, family, city, country, etc.
 - strata: sub-group of sample which is homogeneous with respect to some variable - e.g., male/female

Random Sampling Techniques

- Simple random sample
 - need an entire list (or access) to all elements of population
 - draw sample using names in drum, random number table, etc.
 - given a big enough sample it will be representative each member of population has an equal chance of being sampled.
- Systematic random sample (short cut)
 - still need list of every element
 - take every nth element, where $n = \text{pop size} / \text{sample size}$
 - pick first element randomly

Stratified Sampling Techniques

- Stratified (Homogeneous Subgroup) Sample
 - Proportional Stratified Sample
 - sample elements are in the same proportion as they occur in the population
 - allows inferences from sample strata to population strata
 - allows inferences from entire sample to entire population
 - problem: small strata may not give enough detail
 - Equal Stratified Sample
 - equal proportion of sample comes from each strata of population
 - different size of strata populations - insures stability of sample from smaller strata
 - each strata is equally representative of its target population
 - allows comparisons between strata - internally valid
 - EXAMPLE: views of political parties in America

Other Sampling Techniques

- Purposeful Sample
 - identify cluster of sample that is representative of entire population with respect to the variable of interest
 - randomly select from cluster
- Incidental (convenience) Sample
 - sample from convenient or available population e.g., subject pool! most psychological research does this
 - phone surveys – sample only people in phonebook
 - external validity is limited