SAN JOSÉ STATE UNIVERSITY **College of Social Work**
S. W. 242
Spring 2008
Edward Cohen

Week 6: Hypothesis Testing continued; Chi Square Test of Association

Concepts you should know:

- Statistical inference
- Theoretical sampling distribution
- Null hypothesis
- Research, or alternative, hypothesis
- One-tailed vs. two-tailed hypotheses
- Type I and Type II errors
- Statistical significance
- Critical region
- Alpha level
- *p* value
- Parametric statistics
- Non-parametric statistics
- Binomial distribution
- Chi-squared test of association
- Expected frequencies
- Observed frequencies
- Degrees of freedom

I. The Logic of Statistical Inference--Testing Hypotheses

A. Confirming your research hypothesis (relationship between 2 variables) is dependent on ruling out

1. Rival hypotheses

2. Research design problems (e.g. measurement error, non-representative sample), and/or

3. Chance—*sampling error*--the natural tendency of any sample to differ from the population from which it was drawn. *This one is what statistical inference deals with.*

B. Statistical Inference--The use of theoretical sampling distributions to test hypotheses

1. Theoretical sampling distribution: the normal curve showing all possible sample statistics produced from an infinite number of randomly drawn samples

2. Theoretical sampling is based on the premise that no relationship exists between the two variables (e.g. that the statistical outcome is only obtained by chance)

3. Statistical inference tests the *null hypothesis*--a statement that there is no relationship between two variables of interest. Another way of saying it: Any relationship between these variables is only due to *chance*, not a real relationship that exists *in the population*

4. The *research hypothesis*, or alternative hypothesis, states that there is a relationship between the variables

    a) Hypotheses: one-tailed or two-tailed?

    b) One-tailed (directional) hypothesis—implies the direction (either positive or negative) in which one variable affects or is related to another

        (1) "Female social workers are more satisfied with their jobs than male social workers."

        (2) "Adolescents receiving Cognitive Behavioral Therapy will have decreased anxiety scores compared to adolescents receiving treatment as usual."

    c) Two-tailed (non-directional) hypothesis—does not imply a direction either way

        (1) "Gender is related to job satisfaction."

        (2) "Cognitive Behavioral Therapy is related to changes in adolescent anxiety scores."

C. Type I and Type II Errors. Errors in drawing conclusions about relationships.

		Our Decision	
		Reject Null Hypothesis	Accept Null Hypothesis
TRULY in the Real World	Null hypothesis is false (The alternative hypothesis is true and there is truly an effect)	No error	Type II error (We have wrongly rejected the alternative hypothesis)
	Null hypothesis is true (The alternative hypothesis is false and there is truly no effect)	Type I Error (We have wrongly accepted the alternative hypothesis)	No error

A Type II error is preferred if being conservative.

For example, if you do a clinical trial for Drug X, the null hypothesis is that Drug X is not related to improvement. If Drug X were TRULY

EFFECTIVE but you made a Type II error, you would be accepting the null hypothesis and rejecting the alternative hypothesis, thus delaying or preventing Drug X from use.

BUT, if Drug X were TRULY INEFFECTIVE and you made a Type I error, you would be rejecting the null hypothesis and accepting the alternative hypothesis, thus approving an ineffective (or dangerous) drug for use!

D. Statistical Significance

1. The theoretical sampling distribution shows, out of all the "area under the curve," the tiny probability of rejecting the Null hypothesis in favor of the research hypothesis. The steps are:

   a) By calculating (or computing) a test statistic

   b) Then locating where the statistic falls in the theoretical sampling distribution, and from that

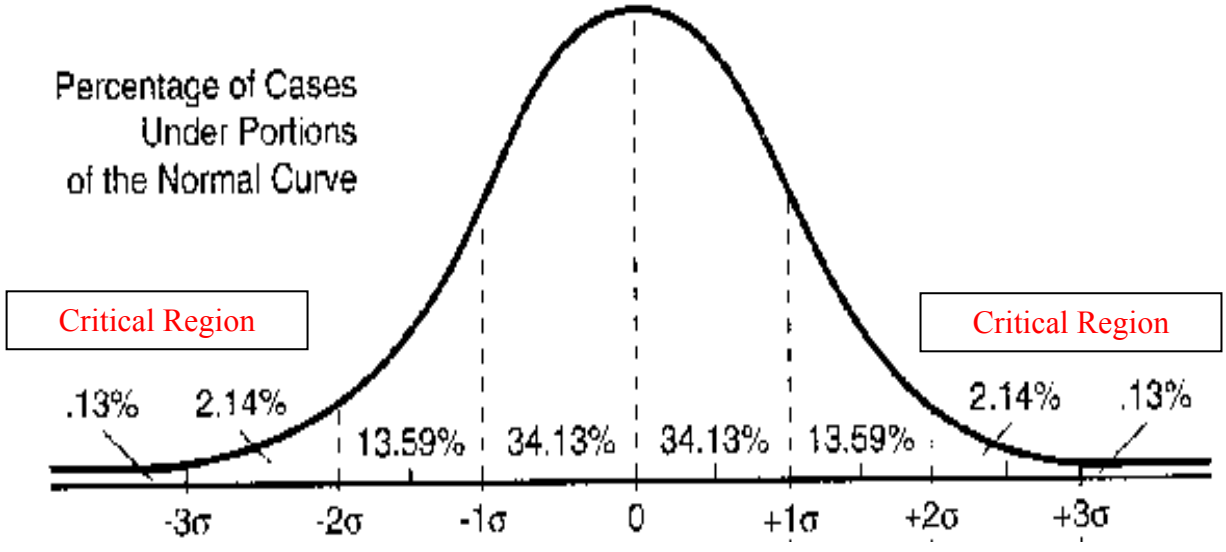
   c) Determining the likelihood (probability) that the statistical result you found is due to chance alone (sampling error)

E. *p*-value (*p*) is the mathematical probability that a relationship between variables found within a sample may have been produced by chance or error.

1. Each statistic result is accompanied by a *p* value

2. SPSS gives you the actual *p* value by using the statistic's computation formula and the distribution tables for the statistical test you've chosen

3. If a statistic result falls in the tiny "critical region" then there is low probability that our results are due to chance alone meaning that there is a good chance there is a positive relationship between variables, and we can reject the Null Hypothesis



Example:

F. Alpha level--we set a threshold for the critical region ahead of time, called the "alpha level"

a) $\alpha = .05$ is typical in social sciences research

b) In some cases higher, $\alpha = .10$

Example:

An experimental group's outcome improved by 10 points, the control group by only 2. Let's say the statistic comparing the groups' difference in post-treatment scores has a *p* value of .046. So,

"The probability is less than 5 in 100 (*p= .046)* that the difference between the groups is due to chance alone. We can reject the Null (that there is no difference) in favor of the alternative (one-tailed) hypothesis, that treatment outcomes will improve more for the experimental group."

II. Parametric and Non-Parametric Tests

A. Parametric tests require:

1. At least one variable (usually the dependent) is either at the interval level or ratio level of measurement;

2. The dependent variable is normally distributed (within the population); if samples are drawn from different populations and then compared, the distributions of the variable within these different populations are required to have near equal variances.

3. Cases have been selected independently, that is, they are randomly selected or, if an experimental design was used, randomly assigned to experimental and control groups.

B. Nonparametric tests are for research situations in which one or more of the conditions for the use of parametric tests do not exist, that is, their assumptions are not met.

III. Chi-squared Test of Association – a non-parametric statistical test.

A. Suitable for two or more nominal level variables (gender; satisfied vs. not; ethnicity; receiving services vs. not)

1. By definition the distribution of a nominal variable is not normal, but binomial

2. But the basic logic of inference testing is the same.

3. The chi-squared test can be used for both nominal and ordinal variables

B. How is the chi-squared test of association typically used?

1. As a bivariate analysis (two variables) to describe your sample, e.g.

Baseline demographics (Excerpted table)

	Condition (Randomly Assigned)		
	Experimental	Control	Difference
Age	43.5 SD = 7.8	43.5 SD 4.7	No diff
Gender			
Male	61%	60%	No diff
Race			No diff
White	37%	38%	
Black	61%	62%	
Hispanic	2%	0%	
Employed (at least one day per month)	25%	19%	No diff
Has marketable skill or trade	66%	66%	No diff
Current driver's license	26%	11%	$p < .05$

Excerpted from: Zanis, D. A., Coviello D., Alterman, A. I., & Appling, S. E. (2001). A community-based trial of vocational problem-solving to increase employment among methadone patients. *Journal of Substance Abuse Treatment, 21*, 19-26.

2. To compare two or more groups' results

Gender Differences in Social Integration: African Americans and Whites

	% women ≥ once per week		% men ≥ once per week	
	African American	White	African American	White
Visits by friends	46.76	44.33*	51.27	42.32*
Visits to friends	37.43	37.36	50.33	38.89*
Phone close friends or relatives	82.33	89.21	68.34	75.32*
Church, other groups	60.76	53.25*	43.91	43.91
Someone to share private feelings, concerns	82.69	88.23*	81.65	84.96*

Note: Differences tested via chi square: (1) African American women vs. White women; (2) African American men vs. White men

*$p < .05$

From: Snowden, L. R. (2001). Social embeddednessand psychological well-being among African Americans and Whites. *American Journal of Community Psychology, 29* , 519-536.