

Week 4: A Short Course in Statistics Handout

This information was prepared to call your attention to some basic concepts underlying statistical procedures and to illustrate what types of research questions can be addressed by different statistical tests. You may not fully understand these tests without further study. However, you are strongly encouraged to note distinctions related to type of measurement used in gathering data and the choice of statistical tests. Feel free to post questions in the “Contact the Instructor” section of the course.

Statistical symbols:

μ mu (population mean)

α alpha (degree of error acceptable for incorrectly rejecting the null hypothesis, probability that results are unlikely to occur by chance)

$\neq$ (not equal)

$\geq$ (greater than or equal to)

$\leq$ less than or equal to)

r (sample correlation)

ρ rho (population correlation)

t r (t score)

z (standard score based on standard deviation)

χ^2 Chi square (statistical test for variables that are not interval or ratio scale, (i.e. nominal or ordinal))

p (*probability that results are due to chance*)

Descriptives:

Descriptives are statistical tests that summarize a data set.

They include calculations of measures of central tendency (mean, median, and mode), and dispersion (e.g., standard deviation and range).

Note: The measures of central tendency depend on the measurement level of the variable (nominal, ordinal, interval, or ratio). If you do not recall the definitions for these levels of measurement, see

http://www.ats.ucla.edu/stat/mult_pkg/whatstat/nominal_ordinal_interval.htm

You can only calculate a mean and standard deviation for interval or ratio scale variables.

For nominal or ordinal variables, you can examine the frequency of responses. For example, you can calculate the percentage of participants who are male and female; or the percentage of survey respondents who are in favor, against, or undecided.

Often nominal data is recorded with numbers, e.g. male=1, female=2. Sometimes people are tempted to calculate a mean using these coding numbers. But that would be

meaningless. Many questionnaires (even course evaluations) use a likert scale to represent attitudes along a continuum (e.g. Strongly like ... Strongly dislike). These too are often assigned a number for data entry, e.g. 1–5. Suppose that most of the responses were in the middle of a scale (3 on a scale of 1–5). A researcher could observe that the mode is 3, but it would not be reasonable to say that the average (mean) is 3 unless there were exact differences between 1 and 2, 2 and 3 etc. The numbers on a scale such as this are ordered from low to high or high to low, but there is no way to say that there is a quantifiably equal difference between each of the choices. In other words, the responses are ordered, but not necessarily equal. Strongly agree is not five times as large as strongly disagree. (See the textbook for differences between ordinal and interval scale measures.)

Inferential Statistics:

Statistical tests for analysis of differences or relationships are Inferential, allowing a researcher to infer relationships between variables.

All statistical tests have what are called assumptions. These are essentially rules that indicate that the analysis is appropriate for the type of data. Two key types of assumptions relate to whether the samples are random and the measurement levels.

Other assumptions have to do with whether the variables are normally distributed. The determination of statistical significance is based on the assumption of the normal distribution. A full course in statistics would be needed to explain this fully. The key point for our purposes is that some statistical procedures require a normal distribution and others do not.

Understanding Statistical Significance

Regardless of what statistical test you use to test hypotheses, you will be looking to see whether the results are statistically significant. The statistic p is the probability that the results of a study would occur simply by chance. Essentially, a p that is less than or equal to a predetermined (α) alpha level (commonly .05) means that we can reject a null hypothesis. A null hypothesis always states that there is no difference or no relationship between the groups or variables. When we reject the null hypothesis, we conclude (but don't prove) that there is a difference or a relationship. This is what we generally want to know.

Parametric Tests:

Parametric tests are tests that require variables to be measured at interval or ratio scale and for the variables to be normally distributed.

These tests compare the means between groups. That is why they require the data to be at an interval or ratio scale. They make use of the standard deviation to determine whether the results are likely to occur or very unlikely in a normal distribution. If they are very unlikely to occur, then they are considered statistically significant. This means that the results are unlikely to occur simply by chance.

The T test

Common uses:

- To compare mean from a sample group to a known mean from a population
- To compare the mean between two samples
 - The research question for a t test comparing the mean scores between two samples is: Is there a difference in scores between group 1 and group 2? The hypotheses tested would be:
 $H_0: \mu_{\text{group1}} = \mu_{\text{group2}}$
 $H_1: \mu_{\text{group1}} \neq \mu_{\text{group2}}$
- To compare pre- and post-test scores for one sample
 - The research question for a t test comparing the mean scores for a sample with pre and posttests is: Is there a difference in scores between time 1 and time 2? The hypotheses tested would be :
 $H_0: \mu_{\text{pre}} = \mu_{\text{post}}$
 $H_1: \mu_{\text{pre}} \neq \mu_{\text{post}}$

Example of the form for reporting results: The results of the test were not statistically significant, $t(57) = .282$, $p = .779$, thus the null hypothesis is not rejected. There is not a difference in between pre and post scores for participants in terms of a measure of knowledge (for example).

An explanation: The t is a value calculated using means and standard deviations and a relationship to a normal distribution. If you calculated the t using a formula, you would compare the obtained t to a table of t values that is based on one less than the number of participants ($n-1$). $n-1$ represents the degrees of freedom. The obtained t must be greater than a critical value of t in order to be significant. For example, if statistical analysis software calculated that $p = .779$, this result is much greater than .05, the usual alpha-level which most researchers use to establish significance. In order for the t test to be significant, it would need to have a $p \leq .05$.

ANOVA (Analysis of variance)

Common uses: Similar to the t test. However, it can be used when there are more than two groups.

The hypotheses would be

- $H_0: \mu_{\text{group1}} = \mu_{\text{group2}} = \mu_{\text{group3}} = \mu_{\text{group4}}$
- $H_1: \text{The means are not all equal (some may be equal)}$

Correlation

Common use: to examine whether two variables are related, that is, they vary together. The calculation of a correlation coefficient (r or ρ) is based on means and standard deviations. This requires that both (or all) variables are measured at an interval or ratio level.

The coefficient can range from -1 to +1. An r of 1 is a perfect correlation. A + means that as one variable increases, so does the other. A – means that as one variable increases, the other decreases.

The research question for correlation is: “Is there a relationship between variable 1 and one or more other variables?”

The hypotheses for a Pearson correlation:

$H_0: \rho = 0$ (there is no correlation)

$H_1: \rho \neq 0$ (there is a real correlation)

Non-parametric Tests

Nonparametric tests are tests that do not require variable to be measured at interval or ratio scale and do not require the variables to be normally distributed.

Chi Square

Common uses: Chi square tests of independence and measures of association and agreement for nominal and ordinal data.

The research question for a chi square test for independence is: Is there a relationship between the independent variable and a dependent variable?

The hypotheses are:

H_0 (The null hypothesis) There is no difference in the proportions in each category of one variable between the groups (defined as categories of another variable).

Or:

The frequency distribution for variable 2 has the same proportions for both categories of variable 1.

H_1 (The alternative hypothesis) There is a difference in the proportions in each category of one variable between the groups (defined as categories of another variable).

The calculations are based on comparing the observed frequency in each category to what would be expected if the proportions were equal. (If the proportions between observed and expected frequencies are equal, then there is no difference.)

See the SOCW 6311: Week 4 Working With Data Assignment Handout to explore the Crosstabs procedure for chi square analysis.

Other non-parametric tests:

Spearman rho: A correlation test for rank ordered (ordinal scale) variables.