

Interpreting Your Questionnaire Data

Soc 357
Fall 2006

Guide to Your Statistical Analysis

- **3 Parts:**
 1. **Frequencies**
 2. **Reliability Analysis**
 3. **Hypothesis Testing**
 1. **Difference of Means**
 2. **Correlations**

Frequencies (1)

- The number of people who fall into each category of a variable
 - Gender – # of men, # of women
 - “Ordinary Americans are losing jobs to illegal immigrants”
 - # of Strongly Agree; # of Agree; # of Disagree; # of Strongly Disagree

Frequencies (2)

Notes:

1. Sometimes not all categories will appear on the frequency table for a variable. If one category/value is missing, that means frequency for that category is **zero**.

Frequencies (3)

- Three things to look for
 1. How much variability in the answers
 2. Comparisons across different answers measuring the dependent variable
 3. Information about the distribution of attitudes in the population.

Interpreting Variation

- Low variability:
 - IF 80% or more cases are in one category: statistical results are problematic; question usually thrown out
 - IF 60% in one category; 80% in 2 adjacent categories (eg. Agree & Strongly Agree): interpret with caution; could be **representative of population**; could be **biased sample**, or **bad question** – You have to decide!

Variability – Example 1

*How frequently do you use other kinds of tobacco products?
(Lowest Frequency = 1; Highest Frequency = 7)*

4tobacco

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1	26	86.7	86.7	86.7
2	4	13.3	13.3	100.0
Total	30	100.0	100.0	

Variability – Example 2

Regardless of sexual orientation, I believe all people have the right to life, liberty, and the pursuit of happiness.

(Strongly disagree = 1; Disagree = 2; Agree = 3; Strongly Agree=4)

Rights2

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 3.000	5	10.4	10.4	10.4
4.000	43	89.6	89.6	100.0
Total	48	100.0	100.0	

Comparing Answers Across Questions

- Which items in your Index elicit most favorable response? Which items elicit least favorable?

Example: Comparing Questions

SMEAN(Closed3R14)

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 2.000	18	90.0	90.0	90.0
2.053	1	5.0	5.0	95.0
3.000	1	5.0	5.0	100.0
Total	20	100.0	100.0	

How challenging do you find UW classes? (Low values = difficult)

SMEAN(Closed4R15)

How many hours do you study per day? (Low values = few hours)

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1.000	2	10.0	10.0	10.0
2.000	6	30.0	30.0	40.0
3.000	4	20.0	20.0	60.0
4.000	2	10.0	10.0	70.0
5.000	6	30.0	30.0	100.0
Total	20	100.0	100.0	

Distribution of Attitudes in General Population

- For the sake of finding out what people said:
 - How many agreed with your questions?
 - How many disagreed?
 - Are the results surprising to you?

Guide to Reliability Section

- Case Processing Summary
- Alpha Coefficient
- Item statistics
- Inter-item correlations
- Item's relation to the total scale

Reliability Analysis

- All of the items in your index should be **positively correlated** to each other, since they are supposed to be measures of the same concept
- **Coefficient Alpha** is a measure of how correlated all the items are to each other. An Alpha Coefficient of **.7 or higher** is a **good** indicator that your index is valid and reliable (measuring the same underlying concept; likely to get consistent results)

Example: Correlation Matrix

Inter-Item Correlation Matrix

	3cigarettes	5alcohol	6marijuana
3cigarettes	1.000	.519	.652
5alcohol	.519	1.000	.291
6marijuana	.652	.291	1.000

Interpreting the Correlation Matrix

- Direction and strength of correlations:
 - A few negative correlations: these should not be here. Could be 1-2 bad questions. Usually discarded to form new index.
 - Mostly weak correlations – your items are not as closely related as you thought.
 - Most correlations are moderately or strongly positive. This means that your index is good.

Interpreting the Item-Index Table

- Look to see if dropping a question would result in a higher or lower Coefficient Alpha

Immigration Questionnaire Index Statistics (Alpha= .898)

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
jobs2	24.88	49.583	.617	.624	.891
labor3	25.38	53.317	.221	.708	.906
punish4	24.88	47.717	.575	.715	.892
benefits5	24.81	42.563	.871	.958	.874
protectn6	24.69	40.629	.901	.947	.872
health7	24.75	42.333	.826	.954	.877
english8	25.50	49.067	.636	.843	.890
amcultre9	24.69	46.363	.808	.854	.881
hcountry10	24.19	50.429	.452	.526	.897
troops11	24.50	47.600	.565	.861	.893
taxes12	24.56	47.863	.662	.881	.888
law13	25.56	52.263	.227	.731	.909

Hypothesis Testing

- Dependent Variable is your Index (a sum of the answers to your questions)
- Independent variables can be **Nominal** (Eg. Gender), **Ordinal** (Eg. Strongly disagree-Strongly Agree OR Age groups) or **Scale** (Eg. Age in years)
- **Nominal**: use Difference of Means test
- **Ordinal & Scale**: use Correlations

Difference of Means

Does the mean DV Index score differ from category to category?

Report

Eating_Index

sex 18	Mean	N	Std. Deviation
1	32.2667	15	6.26175
2	40.4000	15	3.69942
Total	36.3333	30	6.53021

Correlations

How big is the correlation? What direction is it in? Is it likely to have happened by chance?

Correlations

		DV_Index	Age 15
DV_Index	Pearson Correlation	1	.452*
	Sig. (2-tailed)		.045
	N	20	20
Age 15	Pearson Correlation	.452*	1
	Sig. (2-tailed)	.045	
	N	20	20

*. Correlation is significant at the 0.05 level (2-tailed).
