Political Science Research Methods – Practicum Study Guide

Equations for the Exam (you'll be given these on the exam itself)

Mean

$$\overline{X} = \frac{\Sigma X}{n}$$

Standard Deviation

$$s = \sqrt{\frac{\Sigma(X - \overline{X})^2}{n-1}}$$

Z-score

$$z = \frac{(X - \overline{X})}{s}$$

For each of these hypotheses:
1. "Republicans are more likely to vote in presidential elections."
2. "Individuals with higher incomes are more likely to trust government institutions."
3. "There is a difference between men and women in their levels of political efficacy."
4. "There is a difference between freshmen and juniors on their quiz scores."
   A) Identify the independent variable.
   B) Identify the dependent variable.
   C) Identify if the hypothesis is directional or non-directional.
   D) State the null hypothesis.
   E) Specify the causal mechanism.

Consider the following dataset:

	Australia	Croatia	Mauritania	Poland	Switzerland
GDP (Measured in 1000s of dollars)	30.5	14.5	5.8	15.5	19.3
Life expectancy (Measured in average years lived)	75	75	65	80	80
Landlocked (0=yes; 1=no)	1	1	1	1	0
Happiness (Number of times "happy" appears in national newspaper in 2015)	4401	78	239	1003	851

For this dataset (you should be able to do the math by hand):
1. What are your variables? How many variables do you have?
2. What are your observations?
3. What is the level of measurement for each of your variables?
4. What is the most appropriate measure of central tendency to report for each of these variables?
5. Calculate the most appropriate measure of central tendency for **life expectancy**.
6. Calculate the most appropriate measure of central tendency for **landlocked**.
7. Which variable has the most trouble with validity? Explain your reasoning.

	Ben Carson	Chris Christie	Ted Cruz	Carly Fiorina	Mike Huckabee	John Kasich	Rand Paul	Marco Rubio	Rick Santorum	Donald Trump
Percentage of voters who said they would vote for candidate in Iowa, 1/5-19	10	5	20	1	2	3	3	10	1	25
Years served in Congress	0	0	3	0	0	16	5	5	16	0
Gender (0=male, 1=female)	0	0	0	1	0	0	0	0	0	0

For this dataset (you should be able to do the math by hand):
1. What are your variables? How many variables do you have?
2. What are your observations?
3. What is the level of measurement for each of your variables?
4. What is the most appropriate measure of central tendency to report for each of these variables?
5. Calculate the most appropriate measure of central tendency for **all your variables**.
6. Assuming that the standard deviation for *percentage of voters* is 2.00 and the standard deviation for *years served in Congress* is 1.00, calculate the z-scores for each of these two variables for **Marco Rubio**. On which variable is Marco Rubio more average?

SPSS Review (know how to do the following things in SPSS):
- Identify your observations
- Identify your variables
- Determine the level of measurement for variables in the dataset
- Look up the values of your categorical variables in your dataset
- Calculate and read frequency distributions for your variables
- Write your own hypotheses from the variables available in a SPSS dataset
- Calculate the appropriate descriptive statistics (both central tendency and variability) of your variables, dependent on level of measurement
- Be able to draw the appropriate graph of your variables, dependent on level of measurement
- Calculate z-scores for each observation of your variables and interpret its meaning

Reading
- Use Szafran as a reference guide to support any material you are unclear about
- Steve Lohr: For today's graduate, one word: Statistics
- Geoffrey Lean: Using mobile phones while pregnant

Key Terms

political science
statistics
data
qualitative
quantitative
inferential statistics
population
sample
research Methods
science
scientific method
question
hypothesis
testing
analysis
directional hypothesis
nondirectional hypothesis
null hypothesis
variables
independent variables
dependent variables
intervening variables
extraneous variables
spurious variables

causality
correlation
causal mechanism
endogeneity
spurious variable
operationalization
concept
measurement
levels of measurement
nominal variables
ordinal variables
categorical variables
interval-ratio variables
dichotomous/binary
variables
reliability
test-retest reliability
Validity
face validity

descriptive statistics
central tendency
mean
median
mode
variability
standard deviation
range
variance
histograms
bar graphs
Probability
Frequency curves
The normal curve
Peak
Tails
68-95-99 rule
z-scores
characteristics of normal
curve
statistical significance