

POS 3713: Assignment 4

Assigned: Monday, 6/5/2000

Due date: In my mailbox (561-A) by 1:00pm on Monday, 6/12/2000

Tutorial Session: Friday, 6/9/2000, 9am–10:45am and during class on 6/9/2000

The purpose of this assignment is to introduce you to Analysis of Variance (ANOVA) and measures of association. We will be using the 1996 National Election Study for this exercise. To begin the assignment, open your saved NES file from your previous assignments. Remember to fully answer all questions (typed) and include all relevant output in your final product.

Part A: One-Way ANOVA

One-way ANOVA allows you to test the differences between the means of ordinal or ratio level dependent variables for multiple groups. One-way ANOVA receives its name because the groups that you compare differ on a single independent variable of theoretical interest, such as religion, race, educational level, etc. For example, the table below presents the mean number of days reading the newspaper by educational level among 1711 respondents to the NES survey:

Mean Number of Days Reading News Paper by Education

< 8 th Grade	9-11 Grades	H.S. Grad.	Some College	Junior degree	BA/BS Degree	Advanced Degree
2.56	3.10	3.28	3.23	3.17	3.86	4.01

ANOVA tests the null hypothesis that population means for each educational level are equal against the research hypothesis that at least one of the means is different:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6 = \mu_7$$

H_1 : The mean of at least one of the groups is different from the others, or

$$\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4 \neq \mu_5 \neq \mu_6 \neq \mu_7.$$

To determine if there is a significant difference, ANOVA compares the variance within groups to the variance between groups and computes an F-ratio test statistic based on the ratio of between group variance relative to within group variance. If the between group variance is large enough relative to the within group variance, then the F-ratio will exceed the critical level for $\alpha = .05$ (95% confidence), and you can reject the null hypothesis of equal means. SPSS generates a F-ratio for the above example equal to 4.424, which exceeds the critical F value of approximately 2.09 for d.f. (within) = 1710, and d.f. (between) = 6. The degrees of freedom (within) is calculated as $N - k$, while the degrees of freedom (between) is calculated as $k - 1$, where k is the number of groups you are comparing. Thus, you can reject the null hypothesis and conclude that at least one of the educational levels has a mean significantly different from the others.

The goal of this exercise is to determine if there is a significant difference in presidential approval by educational level (V960610) and party identification (V960420). The first step is to rename V960610 = *Educate* and V960420 = *PartyID2*. Also, recode the value 9 for *Educate* and the values 8 & 9 for *PartyID2* as system-missing. The main dependent variable for this analysis will be the *ClinTher* (V960272) from previous exercises. Follow these steps to conduct the ANOVA test for *Educate* and *PartyID2*:

- Select “Analyze”, “Compare Means”, “One-way ANOVA” from the menu. This will open up the ANOVA interface.
- Click the options button, and then place a checkmark in the “Descriptive” box, and click “Continue”. This tells SPSS to produce a table of descriptive statistics by group, so you can better understand the substantive results.

- Move *ClinTher* into the dependent variable text box using the arrow button. Move *Educate* into the factor text box using the arrow button. You have now defined your dependent and independent variables. Click “OK” to execute the ANOVA.
- For *PartyID2*, repeat the same exact steps, but replace *Educate* with *PartyID* in the factor text box.

Exercise A: Use the information calculated above to answer the following questions:

- 1) Copy the descriptive statistics and ANOVA tables from each analysis into your Word file using CTRL-K (copy objects) and CTRL-V (paste).
- 2) State the null and alternative hypotheses for each case.
- 3) Interpret the output for each analysis (conduct the hypothesis test). Is there a significant difference in support for Clinton according to educational level, or according to party identification? What are the theoretical explanations for the differences or lack of differences? Which independent variable has a more powerful influence on Clinton support? Be sure to discuss the specific numbers in the tables as evidence for your answers to these questions.

Part B: Measures of Association for Nominal Variables (Lambda and Phi)

We want to determine the extent to which a person's approval of Congress is influenced by his/her party identification. We will examine the relationship between V960417 (which was renamed PartyId in previous assignments) and V960270 (Congressional approval). Both of these variables are measured at the nominal level.

- Rename V960270 as *Congapp*, and change the Variable Label to "Congress Approval". Also, recode values 8-9 as system missing for V960270. Produce a frequency distribution for both variables. *PartyID* should have three valid responses (Democrat, Republican, or Independent), while *Congapp* should have two valid responses (Approve, Disapprove).
- Click on "Analyze", "Descriptive Statistics", and "Crosstabs". Select the variable "*Congapp*" and move it into the Row(s) box. Select the variable "*PartyId*" and move it into the Column(s) box. Click on the Statistics bar at the bottom of the window. Click on the boxes for Phi, Chi-square and Lambda, and then click on Continue. Next, click on the Cells bar at the bottom of the window, and select Row and Column percentages, and then click continue. Finally, click on OK.

Exercise B: Use the information calculated above to answer the following questions:

1. Examine the relationship between party identification and approval of Congress by comparing the basic percentages of approval and disapproval for each category of party ID (make sure you refer to the percentages in your discussion). Are people who affiliate with certain parties more likely to support Congress than people who affiliate with other parties? Knowing that this survey was conducted in 1996, do these results make sense to you?
2. Next we want to determine the size and statistical significance of the relationship between party ID and approval of Congress. What is the value for lambda (use the value labeled "Congress Approval Dependent", which tells you *Congapp* is the dependent variable); how can you interpret this statistic? Is lambda significantly different from zero? Conduct a hypothesis test by comparing the p-value (Approx. Sig.) to the value of $\alpha = .05$. What conclusions can you draw about the relationship between party identification and approval of Congress?
3. Examine the value of phi and chi-square. Are these values statistically significant? Are these results consistent with what you found in questions 1 and 2?

Part C: Measures of Association for Ordinal Variables (Gamma and Kendall's tau-b)

Next, we are going to analyze the relationship between ideology and issue positions. In particular, we want to determine how a person's ideology influences their attitudes regarding the importance of defense spending.

- For this exercise, we will be using V960465 (how important is the defense spending issue to R) and V960368 (summary of R's self placement on liberal-conservative ideological spectrum). Rename V960465 as *ImpDef* and V960368 as *Lcsum*. Change the variable labels to "Importance of Defense Spending" and "Summary of Liberal-Conservative Spectrum". For *ImpDef*, recode 0, 8, and 9 as system missing. For *Lcsum*, recode 7, 8, and 9 as system missing. If you have done this procedure correctly, you will have three valid responses for *Lcsum* and five valid responses for *ImpDef*.
- Follow the same procedure for crosstabs described in Part A above (replacing the variable in the Row(s) box with *ImpDef* and the variable in the Column(s) box with *Lcsum*). When you click on the Statistics box, unclick the lambda and phi options. Then select Gamma and Kendall's tau-b.

Exercise C: Use the information calculated above to answer the following questions:

1. Examine the relationship between ideology and attitudes about defense spending by comparing the basic percentages of importance of defense spending for each category of ideology. Are people who identify as liberals more or less likely to believe that defense spending is important (refer to percentages in the tables for your response)? Are these results consistent with what you expected to find?
2. Next we want to determine the size and statistical significance of the relationship between these variables. What is the value for gamma; how can you interpret this statistic? Is gamma significantly different from zero? Conduct a hypothesis test by comparing the p-value (Approx. Sig.) to the value of $\alpha = .05$.
3. What is the value of tau-b? Is tau-b statistically significant? What conclusions can you draw about the relationship between ideology and attitudes about the importance of defense spending?