Andrew Peters

2/26/15

Dr. Song

Assignment 3

PSY5013 Spring 2014

**This assignment is due by** **Feb 26, 2015 in class. Relevant syntax and output from using statistical packages needs to be included or attached.**

**Analysis I: Hierarchical regression**

This data set contains a subset of cases from Aiken, West, Woodward, & Reno, 1994, Health Psychology, 13(2), 122-129. The data set is **health.dat**.

Women provided information on the following sets of variables:

DEMOGRAPHICS:

AGE: age category (7 values);

INCOME: category into which total family annual income falls (8 values);

MEDICAL INPUT:

HAVEPHSY: whether they have a physician or regular place for health care (binary: 1=yes; 0 = no)

PHYSREC: whether physician had ever recommended mammogram (binary: 1=yes; 0 = no)

BREAST DISEASE RELATED HISTORY:

LUMP: woman had ever found lump in breast (binary: 1=yes; 0 = no)

WOMREL: woman has relative (mother, sister, grandmother) who has had breast cancer (binary: 1=yes; 0 = no)

HEALTH BELIEFS

SUSCEPT1: perceived susceptibility to breast cancer

SEVERE1: perceived severity of breast cancer

BENEFIT1: perceived benefits of mammography screening

BARRIER1: perceived barriers to mammography screening

Criteria:

INTENT1: intention to obtain mammogram in next 3 months (only applies if woman not in compliance with guidelines at point of data collection)

COMPLY: compliance status with regard to mammography (1 = woman in compliance with screening guidelines 0 = woman not in compliance with screening guidelines)7. Read the following paper and summarize at least 8 ways to look at the correlation. Please use one or two sentences to describe each way you choose.

The variables were recorded in the order of

case age income havephys physrec womrel lump suscept1 severe1 benefit1 barrier1 intent1 comply

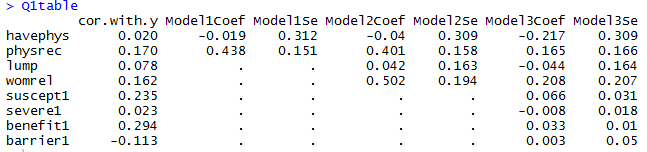
**Run a series of regression analyses to test for the extent to which intention to obtain a mammogram is predicted by the following:**

First set: Medical input only (HAVEPHYS and PHYSREC)

Second set: Medical input plus Breast Disease Related History (LUMP and WOMREL)

Third set: Medical input plus Breast Disease Related History plus the four Health Belief Construct Measures (SUSCEPT1, SEVERE1, BENEFIT1, BARRIER1)

**1. Summarize the results in a table following the format of the table on slide 14 of Lecture 10 (Lecture 11).**

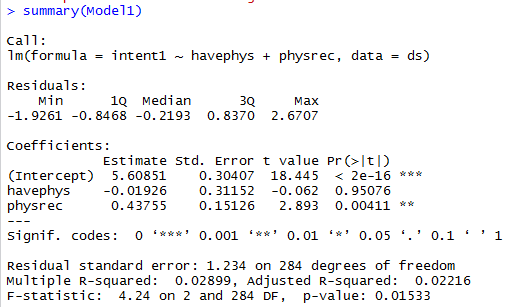


**2. Answer the following question about the analysis:**

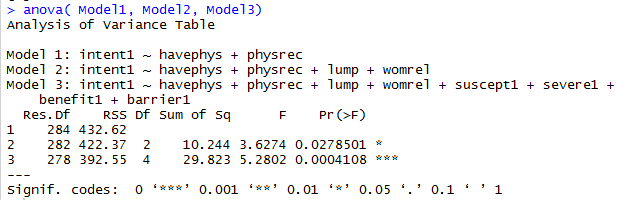
(1). What proportion of variation in the criterion is accounted for by the two medical input variables (HAVEPHYS, PHYSREC)? Is the amount of explained variation significant? Conduct a partial F test by hand and also paste the output from using computer programs for this test (only). Report the F test statistic, dfs, and p value. Which predictor(s) of this set, if any, (is/are) significant?

The first model, containing only the predictor variables “havephys” and “physrec” explains =0.029 = approximately 3 % (adjusted = 0.022) of the variance in dependent variable intent1, which is a significant improvement over a model with no predictor variables (F(2, 284)=4.24, p=0.01533).

The recommendation of a physician was a significant predictor of subjects’ intent to get a mammography, t(286)=2.89, p<0.001. Whether a subject had a regular physician did not seem to be a significant predictor of intent, t(286)=0.062, p>0.05.



(2). What proportion of variation in the criterion is accounted for by the set of breast related history variables, over and above the medical input variables? Is the gain in the explained variation significant? Conduct a partial F test by hand and also paste the output from using computer programs for this test (only). Report the F test statistic, dfs, and p value. Which predictor(s) of this set, if any, (is/are) significant?



The R-squared value of the second model was =0.052, giving us a delta-R-squared of ; thus, the history variables in the second model explain approximately 2.3% of the variance, over and above the medical input variables. This represents a significant improvement over a model with only the medical input variables, F(2,282)=3.42, p<0.05.

3.42

Whether the subject was related to a woman who has had breast cancer was a significant predictor of intent, t(286)=2.579, p<0.05.

Whether the subject had ever felt a lump was not a significant predictor of intent, t(286)=0.58, p>0.05.

(3). What proportion of variation in the criterion is accounted for by the set of Health belief model variables with the criterion, over and above the medical input plus breast disease related variables. Is the gain in the explained variation significant? Conduct a partial F test by hand and also paste the output from using computer programs for this test (only). Report the F test statistic, dfs, and p value. Which predictor(s) of this set, if any, (is/are) significant?

The proportion of variance explained by the belief variables, over and above the medical and history variables is = 6.7% of the variance.

This is a significant improvement over the variance explained by a model with only medical and history variables, F(4, 278)= 5.28, p<0.05.

5.28.

Perceived susceptibility was a significant predictor of intent, t(278)=2.013, p<0.05.

Perceived benefits of a mammography was a significant predictor of intent, t(278)=3.453, p<0.001.

Perceived severity of breast cancer was NOT a significant predictor, t(287)=-0.456, p>0.05.

Perceived barriers to a mammography was NOT a significant predictor, t(287)=.053, p>.0.5.

**3. Report the results in text by following the template shown in Slide 21 of Lecture**

**11.**

**Analysis II: Semi-Partial and Partial Correlations**

You will use the data set HW3\_GPA to compute semi-partial and partial correlations. The variables included are CGPA, HGPA, SAT, and SEX.

1. Compute the **semi-partial correlation** between CGPA and SAT by partialling out the effect of HGPA from CGPA. You need to use three ways to do this: (a) working on the residuals to compute this correlation using computers (refer to the lecture for this method), (b) using zero-order correlation coefficients to compute this correlation by hand, (c) using R-squares to compute this (squared) correlation by hand.
2. The correlation between SAT and the residuals of CGPA regressed onto HGPA is .
3. The semi-partial correlation between CGPA and SAT, controlling for effects of HGPA on CGPA, using correlation coefficients is
4. The semi-partial correlation between CGPA and SAT, controlling for effects of HGPA on CGPA, using coefficients of determination, is

And therefore,

1. Compute the **partial correlation** between CGPA and SAT by partialling out the effect of HGPA from both CGPA and SAT. You need to use three ways to do this: (a) working on the residuals to compute this correlation using computers (refer to the lecture for this method), (b) using zero-order correlation coefficients to compute this correlation by hand, (c) using R-squares to compute this (squared) correlation by hand.
2. The correlation between the residuals of SAT regressed onto HGPA and the residuals of CGPA regressed onto HGPA is .
3. The partial correlation between SAT and CGPA, partialling out the effects of HGPA on both, is
4. The partial correlation between SAT and CGPA, controlling for the effects of HGPA on both, using coefficients of determination is

And therefore,