Review Questions for the Final Exam  
**Vartanian: SW 540**

1. You are examining the effects of adolescent variables on income as an adult over the age of 25. In this first example, you only have a single independent variable, the number of kids in the household during adolescence.

AAVGKIDS = number of kids in the household during adolescence.  
AALLINC = income as an adult.

**Model Summary**

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.193</td>
<td>.037</td>
<td>.037</td>
<td>38250.4572</td>
</tr>
</tbody>
</table>

a Predictors: (Constant), AAVGKIDS

**ANOVA**

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>187844314078.610</td>
<td>1</td>
<td>187844314078.610</td>
<td>128.388</td>
<td>.000</td>
</tr>
<tr>
<td>Residual</td>
<td>4856020514347.840</td>
<td>331</td>
<td>1463097473.442</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5043864828426.450</td>
<td>3320</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Predictors: (Constant), AAVGKIDS  
b Dependent Variable: AALLINC

**Coefficients**

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B Std. Error Beta</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>(Constant) 67992.569 922.168 73.731 .000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAVGKIDS -8967.831 791.453 -1.193 -11.331 .000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Dependent Variable: AALLINC

Questions:

A. Is this model any good? How can you tell?
B. What are the two ways of determining the F value? Is the F value significant at the .001 level?
C. What does the standardized b coefficient tell us? Is it statistically significant?
D. If the standard deviation for kids is 2 and the standard deviation for income is $1,000, how much would income change if the number of kids increased by 1 standard deviation?

2. What are the consequences of failing to control for the effects of relevant variables in multiple regression models?

3. You test for the effects of a number of independent variables on a dependent variable but find that none of the coefficient estimates is statistically significant. Give some reasons for these results.

4. You would like to determine if there is a significant relationship of two nominal scale
variables. What type of test will you use?

5. You are examining the relationship between being living in a big city and being in poverty. You hypothesize that there is a relationship between these variables. You get the following output.

Crosstabs

<table>
<thead>
<tr>
<th>BIGCIT * INPOV Crosstabulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>INPOV</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>.00</td>
</tr>
<tr>
<td>Count</td>
</tr>
<tr>
<td>5417</td>
</tr>
<tr>
<td>Expected Count</td>
</tr>
<tr>
<td>5329.8</td>
</tr>
<tr>
<td>1.00</td>
</tr>
<tr>
<td>Count</td>
</tr>
<tr>
<td>1358</td>
</tr>
<tr>
<td>Expected Count</td>
</tr>
<tr>
<td>1445.2</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>6775</td>
</tr>
<tr>
<td>6775.0</td>
</tr>
</tbody>
</table>

Chi-Square Tests

<table>
<thead>
<tr>
<th>Value</th>
<th>df</th>
<th>Asymp. Sig. (2-sided)</th>
<th>Exact Sig. (2-sided)</th>
<th>Exact Sig. (1-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>35.259b</td>
<td>1</td>
<td>.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Continuity Correction</td>
<td>34.856</td>
<td>1</td>
<td>.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>33.794</td>
<td>1</td>
<td>.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Fisher's Exact Test</td>
<td>35.255</td>
<td>1</td>
<td>.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Linear-by-Linear Association</td>
<td>8363</td>
<td>1</td>
<td>.000</td>
<td>1.000</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>8363</td>
<td>1</td>
<td>.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

a. Computed only for a 2x2 table
b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 338.75.

Questions:

A. In what direction is the relationship?
B. How many DF's are there in this test?
C. What is the critical value for this test for a 2-tailed, .05 test?
D. At what level will you reject the null hypothesis?

6. You’re examining the relationship between the economic situation of a child and whether they are in poverty as an adult. The three categories for economic situation as a child are growing up poor, growing up middle income, and growing up rich. The two categories for poverty status are in poverty (a value of 1) or out of poverty (a value of 0). You get the following output.

Crosstabs
ECSIT * INPOV Crosstabulation

<table>
<thead>
<tr>
<th></th>
<th>INPOV</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.00</td>
<td>1.00</td>
<td>Total</td>
</tr>
<tr>
<td>ECSIT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grew Up Poor</td>
<td>Count</td>
<td>2443</td>
<td>809</td>
</tr>
<tr>
<td></td>
<td>Expected Count</td>
<td>2649.9</td>
<td>602.1</td>
</tr>
<tr>
<td>Grew Up Middle Class</td>
<td>Count</td>
<td>2626</td>
<td>349</td>
</tr>
<tr>
<td></td>
<td>Expected Count</td>
<td>2424.2</td>
<td>550.8</td>
</tr>
<tr>
<td>Grew Up Rich</td>
<td>Count</td>
<td>1453</td>
<td>324</td>
</tr>
<tr>
<td></td>
<td>Expected Count</td>
<td>1448.0</td>
<td>329.0</td>
</tr>
<tr>
<td>Total</td>
<td>Count</td>
<td>6522</td>
<td>1482</td>
</tr>
<tr>
<td></td>
<td>Expected Count</td>
<td>6522.0</td>
<td>1482.0</td>
</tr>
</tbody>
</table>

Chi-Square Tests

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>df</th>
<th>Asymp. Sig. (2-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>178.082</td>
<td>2</td>
<td>.000</td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>182.270</td>
<td>2</td>
<td>.000</td>
</tr>
<tr>
<td>Linear-by-Linear</td>
<td>62.548</td>
<td>1</td>
<td>.000</td>
</tr>
<tr>
<td>Association</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>8004</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 329.02.

A. Does economic situation while growing up have a significant effect on the likelihood of being in poverty as an adult?
B. What is the direction of the relationship between economic situation while growing up and poverty as an adult?
C. Is this significant effect (if there is one) driven by those who grow up rich?

7. You know that out of a sample of 300 people, with 100 men and 200 women, 50 of the men wear rings on their fingers and 50 of the men do not. For women, 80 of the women wear rings on their fingers and 120 of the women do not. Is there a significant difference in ring preference between the men and women?

<table>
<thead>
<tr>
<th></th>
<th>Women</th>
<th>Men</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rings</td>
<td>80</td>
<td>50</td>
<td>130</td>
</tr>
<tr>
<td>No Rings</td>
<td>120</td>
<td>50</td>
<td>170</td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td>100</td>
<td>300</td>
</tr>
</tbody>
</table>
8. You have determined non-standardized b coefficients and standard deviations for a group of independent variables and the standard deviation for the dependent variable number of shoes:

<table>
<thead>
<tr>
<th>Variable</th>
<th>SD</th>
<th>b Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoes</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>income</td>
<td>200</td>
<td>0.01</td>
</tr>
<tr>
<td>Books</td>
<td>2.00</td>
<td>-0.50</td>
</tr>
<tr>
<td>Zoo visits</td>
<td>1.00</td>
<td>1.25</td>
</tr>
</tbody>
</table>

Find the standardized b coefficients for each of the independent variables. Interpret the standardized b coefficients.

9. Determine the adjusted $R^2$ value given the following information:

The amount of explained variance in the model is 0.90, with

A. 25 observations and 3 independent variables.
B. 10 observations and 3 independent variables.
C. 10 observations and 6 independent variables.
D. Given your results, what do these results indicate about the effects of sample size and number of independent variables on the adjusted $R^2$ value?

10. Given the information below, determine:

<table>
<thead>
<tr>
<th>Obs</th>
<th>X₁</th>
<th>X₂</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Sum</td>
<td>10</td>
<td>8</td>
<td>6</td>
</tr>
</tbody>
</table>

A. Using the error terms, determine the partial correlation coefficients for Y and $x_1$ controlling for $x_2$ (you don’t need to plug the numbers into the bivariate $r$ formula – just know what numbers to plug in).
B. The percentage of variance explained, with $y$ as the DV and $x_1$ and $x_2$ as the IVs.
C. If the model is significant.
D. Using the error terms, determine the partial b coefficients for Y and x1, controlling for x2 (again, just know what numbers to plug into the bivariate b formula).

E. Same as part D but partial out x1, and determine the partial b for x2, with Y as the DV.

You are given the following bivariate relationships:

\[
\begin{align*}
 b_{x1x2} &= -0.500 \quad a=3.50 \\
 b_{x1y} &= 1.900 \quad a=0.40 \\
 b_{x2x1} &= -0.200 \quad a=2.50 \\
 b_{x2y} &= -0.600 \quad a=2.90 \\
 b_{yx1} &= 0.400 \quad a=0.50 \\
 b_{yx2} &= -1.500 \quad a=4.50 \\
 r_{x1y} &= 0.400 \\
 r_{x2y} &= -0.948 \\
 r_{x1x2} &= -0.316
\end{align*}
\]

11. Describe the meaning of the unstandardized b coefficient estimates in the following model, where the dependent variable is income, and the independent variables are whether or not the head of household is in jail, whether or not the head of household smokes, region of country where the family lives (north central and west included in the model, south and north east excluded from the model), age of the head of household, and number of kids.

### Regression

#### Model Summary

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.174&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.030</td>
<td>.029</td>
<td>74225.35059</td>
</tr>
</tbody>
</table>

<sup>a</sup> Predictors: (Constant), AGEHD, WEST, JAILHD, FEMALE, HDSMOKES, NC, KIDS

#### ANOVA<sup>b</sup>

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Regression</td>
<td>8.49E+11</td>
<td>7</td>
<td>1.214E+11</td>
<td>22.026</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td>2.72E+13</td>
<td>4930</td>
<td>5509402671</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>2.80E+13</td>
<td>4937</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Predictors: (Constant), AGEHD, WEST, JAILHD, FEMALE, HDSMOKES, NC, KID

<sup>b</sup> Dependent Variable: TOTAL FAMILY INCOME
### Coefficients

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(Constant)</td>
<td>81309.875</td>
<td>4136.854</td>
<td>19.655</td>
</tr>
<tr>
<td></td>
<td>HDSMOKES</td>
<td>-16891.8</td>
<td>2567.797</td>
<td>-6.578</td>
</tr>
<tr>
<td></td>
<td>FEMALE</td>
<td>-12989.6</td>
<td>2152.913</td>
<td>-6.033</td>
</tr>
<tr>
<td></td>
<td>JAILHD</td>
<td>-28705.3</td>
<td>22421.993</td>
<td>-1.280</td>
</tr>
<tr>
<td></td>
<td>NC</td>
<td>602.148</td>
<td>2468.860</td>
<td>.244</td>
</tr>
<tr>
<td></td>
<td>WEST</td>
<td>16272.432</td>
<td>3050.122</td>
<td>5.335</td>
</tr>
<tr>
<td></td>
<td>KIDS</td>
<td>3270.567</td>
<td>965.383</td>
<td>3.388</td>
</tr>
<tr>
<td></td>
<td>AGEHD</td>
<td>-342.181</td>
<td>73.107</td>
<td>-4.681</td>
</tr>
</tbody>
</table>

a. Dependent Variable: TOTAL FAMILY INCOME

You’ve also been given descriptive statistics for these variables:

### Descriptive Statistics

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL FAMILY INCOME</td>
<td>4938</td>
<td>-73500.00</td>
<td>1991000</td>
<td>59924.03</td>
<td>75323.63340</td>
</tr>
<tr>
<td>HDSMOKES</td>
<td>4938</td>
<td>0.00</td>
<td>1.00</td>
<td>0.228</td>
<td>0.42013</td>
</tr>
<tr>
<td>JAILHD</td>
<td>4938</td>
<td>0.00</td>
<td>1.00</td>
<td>0.0022</td>
<td>0.04715</td>
</tr>
<tr>
<td>KIDS</td>
<td>4938</td>
<td>0.00</td>
<td>8.00</td>
<td>0.8838</td>
<td>1.18795</td>
</tr>
<tr>
<td>NC</td>
<td>4938</td>
<td>0.00</td>
<td>1.00</td>
<td>0.2691</td>
<td>0.1521</td>
</tr>
<tr>
<td>WEST</td>
<td>4938</td>
<td>0.00</td>
<td>1.00</td>
<td>0.1521</td>
<td>0.35914</td>
</tr>
<tr>
<td>AGEHD</td>
<td>4938</td>
<td>17.00</td>
<td>94.00</td>
<td>45.3433</td>
<td>15.89668</td>
</tr>
</tbody>
</table>

### Questions:

A. Discuss the coefficient estimates for the independent variables.

B. Head in jail has a relatively high b coefficient estimate. Why do you believe that it is not statistically significant?

C. What is your predicted level of income for someone who has a head of household who smokes, is a female, is not in jail, lives in the south, has a head who is 0 years old, and has no kids?

D. If examining these coefficients in standard deviation units, which of the interval/ratio scale variables has a larger effect on the dependent variable?

12. You have been given the following output, where the dependent variable is income and the independent variable is female (female=1 if the person is a female). Are
males and females significantly different (at the 5% level) from each other? By how much? How much is the mean income level for each group? N=500.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>6.00</td>
<td>1.000</td>
</tr>
<tr>
<td>Constant</td>
<td>9.00</td>
<td></td>
</tr>
</tbody>
</table>

13. You have been given the following output, where the dependent variable is income and the independent variable is male (male=1 if the person is male). Are males and females significantly (at the 5% level) different from each other? By how much? How much is the mean income level for each group? N=500.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>-5.0</td>
<td>0.693</td>
</tr>
<tr>
<td>Constant</td>
<td>8.2</td>
<td></td>
</tr>
</tbody>
</table>

14. You are examining the effect of area of residence (big city, urban, suburban or rural) and age on fat intake (in grams). You run a regression model on a sample of 500 people and have the following output:

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big City</td>
<td>6.3</td>
<td>2.50</td>
</tr>
<tr>
<td>Urban</td>
<td>4.2</td>
<td>1.10</td>
</tr>
<tr>
<td>Suburban</td>
<td>1.2</td>
<td>1.00</td>
</tr>
<tr>
<td>Age</td>
<td>0.5</td>
<td>0.20</td>
</tr>
<tr>
<td>Constant</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

A. What are the intercept and slopes for each of the four groups?
B. What is the difference in fat intake for each of the groups relative to the excluded category?
C. Are the coefficients significant at the 5% level?

15. You are examining a group of students for their ranking in socioeconomic status and their ranking in reading ability. You wish to determine if there is a relationship between these two rankings. Determine if there is a significant relationship between these two variables using either method taught in class.

<table>
<thead>
<tr>
<th>Ranking in Reading Ability</th>
<th>Ranking in Socioeconomic Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student 1</td>
<td>5</td>
</tr>
<tr>
<td>Student 2</td>
<td>2</td>
</tr>
<tr>
<td>Student 3</td>
<td>1</td>
</tr>
</tbody>
</table>
16. In an OLS multiple regression model, you find that your explained sums of squares is 100, your unexplained sums of squares is 50, with an n=1000, and a k=4. Is there a significant relationship between the set of independent variables and the dependent variable?
#17. You have the following set of 4 observations, with y as your dependent variable and x₁ and x₂ as your independent variables. You are given the following information.

<table>
<thead>
<tr>
<th>Obs</th>
<th>X₁</th>
<th>X₂</th>
<th>Y</th>
<th>Yₚ</th>
<th>e=y-yₚ</th>
<th>e²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0.50</td>
<td>-0.50</td>
<td>0.25</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0.50</td>
<td>0.50</td>
<td>0.25</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>2.50</td>
<td>0.50</td>
<td>0.25</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2.50</td>
<td>-0.50</td>
<td>0.25</td>
</tr>
<tr>
<td>Sum</td>
<td>10</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>X₁</td>
<td>2.5</td>
<td>1.29</td>
</tr>
<tr>
<td>X₂</td>
<td>2.0</td>
<td>0.82</td>
</tr>
<tr>
<td>Y</td>
<td>1.5</td>
<td>1.29</td>
</tr>
</tbody>
</table>

The intercept and slope coefficients on the same line are from the same model.

bₓᵧ₁ = .80   a=-.5
bₓᵧ₂ = -1.0  a=3.5
bₓₓ₁ₓ₂ = -.5 a=3.5
bx₂ₓ₁ = -.2  a=2.5
bₓₓ₁ₓ₂ = .667
bₓₓ₂ₓ₁ = -.667

SEₜₓᵧ₁ = .471
(Yᵢ - Ȳ) = 5
(Yᵢ - Yₚ) = 1
rₓ₁ₓ₂ = .316
rₓᵧ₁ = .80
rₓᵧ₂ = -.632

Questions:
A. What is the multiple regression equation for this model (Yₚ=…).
B. What is the standard error for bₓₓ₂ₓ₁?
C. How much of the variance of the dependent variable is explained in this model?
D. Are either of the coefficient estimates significant when using both independent variables in the statistical model?
E. Is the model significant?
Question 18: Overall, is there a difference in the mean values for the unemployment rate in the area in which people live for the difference groups examined? Which groups are significantly different from each other (at the .05 level)? How different are they?

Oneway

ANOVA

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td>759.258</td>
<td>6</td>
<td>126.543</td>
<td>25.206</td>
<td>.000</td>
</tr>
<tr>
<td>Within</td>
<td>78037.573</td>
<td>15544</td>
<td>5.020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>78796.831</td>
<td>15550</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Post Hoc Tests
## Multiple Comparisons

**Dependent Variable: UNCY**

**Schefte**

<table>
<thead>
<tr>
<th>(I) RACE</th>
<th>(J) RACE</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interv</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>White</td>
<td>-.3575*</td>
<td>.052</td>
<td>.000</td>
<td>-.5413 - .17</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>-.4680</td>
<td>.265</td>
<td>.794</td>
<td>-1.4081 .47</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>American Indian</td>
<td>-.4679</td>
<td>.297</td>
<td>.870</td>
<td>-.5852 1.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Latino/Latina</td>
<td>-.6693</td>
<td>.376</td>
<td>.787</td>
<td>-2.0029 .66</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Color other than b/w</td>
<td>-1.3729</td>
<td>1.179</td>
<td>.968</td>
<td>-5.5573 2.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>-2.1632*</td>
<td>.218</td>
<td>.000</td>
<td>-2.9360 -1.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>White</td>
<td>.3575*</td>
<td>.052</td>
<td>.000</td>
<td>.1737 .54</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>-.1105</td>
<td>.269</td>
<td>1.000</td>
<td>-1.0634 .84</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>American Indian</td>
<td>.8254</td>
<td>.300</td>
<td>.271</td>
<td>-.2391 1.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asian, Pacific Islander</td>
<td>-.3118</td>
<td>.378</td>
<td>.995</td>
<td>-1.6544 1.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Latino/Latina</td>
<td>-.9050</td>
<td>1.180</td>
<td>.994</td>
<td>-5.2027 3.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Color other than b/w</td>
<td>-1.0154</td>
<td>.222</td>
<td>.000</td>
<td>-2.5939 -1.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>-1.8057*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian</td>
<td>White</td>
<td>.4680</td>
<td>.265</td>
<td>.794</td>
<td>-.4722 1.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>.1105</td>
<td>.269</td>
<td>1.000</td>
<td>-.8425 1.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>American Indian</td>
<td>.9358</td>
<td>.397</td>
<td>.474</td>
<td>-.4725 2.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asian, Pacific Islander</td>
<td>-.2014</td>
<td>.459</td>
<td>1.000</td>
<td>-1.8301 1.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Latino/Latina</td>
<td>-.9050</td>
<td>1.208</td>
<td>.997</td>
<td>-5.1925 3.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Color other than b/w</td>
<td>-1.8408</td>
<td>.367</td>
<td>.000</td>
<td>-2.9803 -1.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>-2.6311*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian, Pacific Islander</td>
<td>White</td>
<td>-.4679</td>
<td>.297</td>
<td>.870</td>
<td>-1.5210 .58</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>-.8254</td>
<td>.300</td>
<td>.271</td>
<td>-1.8899 .23</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>American Indian</td>
<td>-.9358</td>
<td>.397</td>
<td>.474</td>
<td>-2.3442 .47</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asian, Pacific Islander</td>
<td>-.11372</td>
<td>.478</td>
<td>.462</td>
<td>-2.8336 .58</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Latino/Latina</td>
<td>-.9050</td>
<td>1.215</td>
<td>.891</td>
<td>-6.1545 2.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Color other than b/w</td>
<td>-2.6311</td>
<td>.367</td>
<td>.000</td>
<td>-3.9336 -1.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>-2.6311*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latino/Latina</td>
<td>White</td>
<td>.6693</td>
<td>.376</td>
<td>.787</td>
<td>-.6643 2.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>.3118</td>
<td>.378</td>
<td>.995</td>
<td>-1.0308 1.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>American Indian</td>
<td>.2014</td>
<td>.459</td>
<td>1.000</td>
<td>-1.4274 1.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asian, Pacific Islander</td>
<td>1.1372</td>
<td>.478</td>
<td>.462</td>
<td>1.5592 2.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Color other than b/w</td>
<td>-.7036</td>
<td>1.237</td>
<td>.999</td>
<td>-5.0942 3.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>-1.4939</td>
<td>.433</td>
<td>.065</td>
<td>-3.0321 4.43E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Color other than b/w</td>
<td>White</td>
<td>1.3729</td>
<td>1.179</td>
<td>.968</td>
<td>-2.8114 5.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>1.0154</td>
<td>1.180</td>
<td>.994</td>
<td>-3.1718 5.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>American Indian</td>
<td>.9050</td>
<td>1.208</td>
<td>.997</td>
<td>-3.3826 5.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asian, Pacific Islander</td>
<td>1.8408</td>
<td>1.215</td>
<td>.891</td>
<td>-2.4729 6.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Latino/Latina</td>
<td>.7036</td>
<td>1.237</td>
<td>.999</td>
<td>-3.6870 5.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>-.7903</td>
<td>1.199</td>
<td>.999</td>
<td>-5.0442 3.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>White</td>
<td>2.1632*</td>
<td>.218</td>
<td>.000</td>
<td>1.3904 2.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>1.8057*</td>
<td>.222</td>
<td>.000</td>
<td>1.0175 2.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>American Indian</td>
<td>1.6953*</td>
<td>.342</td>
<td>.000</td>
<td>.4822 2.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asian, Pacific Islander</td>
<td>2.6311*</td>
<td>.367</td>
<td>.000</td>
<td>1.3286 3.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Latino/Latina</td>
<td>1.4939</td>
<td>.433</td>
<td>.065</td>
<td>-4.426E-02 3.03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Question 19: Is there an overall difference in the wage levels for wives of different races? Which groups are significantly different from one another (at the .05 level)? How different are they?

Oneway

ANOVA

<table>
<thead>
<tr>
<th>WAGEWF</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>16976.600</td>
<td>6</td>
<td>2829.433</td>
<td>38.935</td>
<td>.000</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1137360.7</td>
<td>15651</td>
<td>72.670</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1154337.3</td>
<td>15657</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Post Hoc Tests
### Multiple Comparisons

**Dependent Variable: WAGEWF**

**Scheffe**

<table>
<thead>
<tr>
<th>(I) RACE</th>
<th>(J) RACE</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval Lower Bound</th>
<th>95% Confidence Interval Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>African American</td>
<td>2.8849*</td>
<td>.197</td>
<td>.000</td>
<td>2.1870</td>
<td>3.58</td>
</tr>
<tr>
<td></td>
<td>American Indian</td>
<td>2.7703</td>
<td>1.008</td>
<td>.272</td>
<td>-8.066</td>
<td>6.34</td>
</tr>
<tr>
<td></td>
<td>Asian, Pacific Islander</td>
<td>-.4553</td>
<td>1.113</td>
<td>1.000</td>
<td>-4.4057</td>
<td>3.45</td>
</tr>
<tr>
<td></td>
<td>Latino/Latina</td>
<td>.2509</td>
<td>1.421</td>
<td>1.000</td>
<td>-4.7925</td>
<td>5.26</td>
</tr>
<tr>
<td></td>
<td>Color other than b/w</td>
<td>4.8527</td>
<td>4.486</td>
<td>.978</td>
<td>-11.0669</td>
<td>20.77</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>-2.3113</td>
<td>.811</td>
<td>.230</td>
<td>-.51905</td>
<td>.5384</td>
</tr>
<tr>
<td>African American</td>
<td>White</td>
<td>-2.8849*</td>
<td>.197</td>
<td>.000</td>
<td>-3.5829</td>
<td>-2.13</td>
</tr>
<tr>
<td></td>
<td>American Indian</td>
<td>-1.1146</td>
<td>1.022</td>
<td>1.000</td>
<td>-3.7399</td>
<td>3.51</td>
</tr>
<tr>
<td></td>
<td>Asian, Pacific Islander</td>
<td>-3.3402</td>
<td>1.125</td>
<td>.185</td>
<td>-7.3345</td>
<td>6.64</td>
</tr>
<tr>
<td></td>
<td>Latino/Latina</td>
<td>-2.6341</td>
<td>1.431</td>
<td>.759</td>
<td>-7.7118</td>
<td>2.44</td>
</tr>
<tr>
<td></td>
<td>Color other than b/w</td>
<td>1.9678</td>
<td>4.489</td>
<td>1.000</td>
<td>-13.9628</td>
<td>17.81</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>-5.1962*</td>
<td>.828</td>
<td>.000</td>
<td>-8.1353</td>
<td>-2.22</td>
</tr>
<tr>
<td>American Indian</td>
<td>White</td>
<td>-2.7703</td>
<td>1.008</td>
<td>.272</td>
<td>-6.3472</td>
<td>.84</td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>.1146</td>
<td>1.022</td>
<td>1.000</td>
<td>-3.5106</td>
<td>3.73</td>
</tr>
<tr>
<td></td>
<td>Asian, Pacific Islander</td>
<td>-3.2256</td>
<td>1.498</td>
<td>.591</td>
<td>-8.5417</td>
<td>2.07</td>
</tr>
<tr>
<td></td>
<td>Latino/Latina</td>
<td>-2.5194</td>
<td>1.739</td>
<td>.910</td>
<td>-8.6912</td>
<td>3.65</td>
</tr>
<tr>
<td></td>
<td>Color other than b/w</td>
<td>2.0825</td>
<td>4.596</td>
<td>1.000</td>
<td>-14.2298</td>
<td>18.35</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>-5.0816*</td>
<td>.828</td>
<td>.000</td>
<td>-9.6582</td>
<td>-.52</td>
</tr>
<tr>
<td>Asian, Pacific Islander</td>
<td>White</td>
<td>.4553</td>
<td>1.113</td>
<td>1.000</td>
<td>-3.4952</td>
<td>4.44</td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>3.3402</td>
<td>1.125</td>
<td>.185</td>
<td>-6.541</td>
<td>7.34</td>
</tr>
<tr>
<td></td>
<td>American Indian</td>
<td>3.2256</td>
<td>1.498</td>
<td>.591</td>
<td>-2.0906</td>
<td>8.54</td>
</tr>
<tr>
<td></td>
<td>Latino/Latina</td>
<td>.7062</td>
<td>1.802</td>
<td>1.000</td>
<td>-5.6894</td>
<td>7.13</td>
</tr>
<tr>
<td></td>
<td>Color other than b/w</td>
<td>5.3080</td>
<td>4.621</td>
<td>.971</td>
<td>-11.0902</td>
<td>21.72</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>-1.8560</td>
<td>1.373</td>
<td>.935</td>
<td>-6.7301</td>
<td>3.01</td>
</tr>
<tr>
<td>Latino/Latina</td>
<td>White</td>
<td>-.2509</td>
<td>1.421</td>
<td>1.000</td>
<td>-5.2943</td>
<td>4.75</td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>2.6341</td>
<td>1.431</td>
<td>.759</td>
<td>-2.4437</td>
<td>7.71</td>
</tr>
<tr>
<td></td>
<td>American Indian</td>
<td>2.5194</td>
<td>1.739</td>
<td>.910</td>
<td>-3.6524</td>
<td>8.65</td>
</tr>
<tr>
<td></td>
<td>Asian, Pacific Islander</td>
<td>-.7062</td>
<td>1.802</td>
<td>1.000</td>
<td>-7.1017</td>
<td>5.63</td>
</tr>
<tr>
<td></td>
<td>Color other than b/w</td>
<td>4.6019</td>
<td>4.704</td>
<td>.987</td>
<td>-12.0934</td>
<td>21.25</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>-2.5622</td>
<td>1.633</td>
<td>.873</td>
<td>-8.3576</td>
<td>3.22</td>
</tr>
<tr>
<td>Color other than b/w</td>
<td>White</td>
<td>-4.8527</td>
<td>4.486</td>
<td>.978</td>
<td>-20.7724</td>
<td>11.06</td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>-1.9678</td>
<td>4.489</td>
<td>1.000</td>
<td>-17.8984</td>
<td>13.96</td>
</tr>
<tr>
<td></td>
<td>American Indian</td>
<td>-2.0825</td>
<td>4.596</td>
<td>1.000</td>
<td>-18.3947</td>
<td>14.22</td>
</tr>
<tr>
<td></td>
<td>Asian, Pacific Islander</td>
<td>-5.3080</td>
<td>4.621</td>
<td>.971</td>
<td>-21.7063</td>
<td>11.06</td>
</tr>
<tr>
<td></td>
<td>Latino/Latina</td>
<td>-4.6019</td>
<td>4.704</td>
<td>.987</td>
<td>-21.2971</td>
<td>12.02</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>-7.1640</td>
<td>4.557</td>
<td>.872</td>
<td>-23.3773</td>
<td>9.00</td>
</tr>
<tr>
<td>Other</td>
<td>White</td>
<td>2.3113</td>
<td>.811</td>
<td>.230</td>
<td>-.5680</td>
<td>5.18</td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>5.1962*</td>
<td>.828</td>
<td>.000</td>
<td>2.2571</td>
<td>8.13</td>
</tr>
<tr>
<td></td>
<td>American Indian</td>
<td>5.0816*</td>
<td>1.290</td>
<td>.017</td>
<td>.5049</td>
<td>9.64</td>
</tr>
<tr>
<td></td>
<td>Asian, Pacific Islander</td>
<td>1.8560</td>
<td>1.373</td>
<td>.935</td>
<td>-3.0181</td>
<td>6.73</td>
</tr>
<tr>
<td></td>
<td>Latino/Latina</td>
<td>2.5622</td>
<td>1.633</td>
<td>.873</td>
<td>-3.2333</td>
<td>8.33</td>
</tr>
</tbody>
</table>
Question 20.

You are examining the relationship between economic situation while growing up and poverty status as an adult. You get the following cross tabulation.

<table>
<thead>
<tr>
<th></th>
<th>Poor as an adult</th>
<th>Mid Income as an Adult</th>
<th>Rich as an Adult</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grew Up Poor</td>
<td>50</td>
<td>30</td>
<td>10</td>
</tr>
<tr>
<td>Grew Up Mid Income</td>
<td>40</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>Grew Up Rich</td>
<td>10</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

A. Is there a relationship between the two variables?
B. What is the direction of the relationship?
C. Do the adult middle income group add much to the significance of the model? Examine cell b (with an observed frequency of 30) with a residuals test.

21. You want to determine the effects of parental education (x1) on child’s income later in life (Y), controlling for the effects of number of siblings (x2) and good health status (x3), a dummy variable. You will use a regression analysis to determine these effects, using residuals. How would you set this up?

22. Using the Child Development Supplement to the PSID, you run an OLS regression with the natural log of income of the child’s family as your dependent variable. You use age of the child, whether or not the head of household is a college graduate, family size, and whether or not the primary caregiver of the child has emotional problems as your independent variables.

A. Why use log dependent variables?
B. How do you interpret the b coefficients for each of the variables?
C. What is the predicted level of income for someone who had the following characteristics: the head of household was a high school graduate, had a family size of 4, the child was 10, and the primary caregiver had no emotional problems.

Regression

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.420</td>
<td>.177</td>
<td>.175</td>
<td>1.09994</td>
</tr>
</tbody>
</table>

a. Predictors: (Constant), Emot st PCG:have emot conditions, number of persons in family unit, age of child in 1997, head: college grad
### ANOVA

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Regression</td>
<td>4</td>
<td>139.476</td>
<td>115.283</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td>2150</td>
<td>1.210</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>2154</td>
<td>1.210</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. Predictors: (Constant), Emot st PCG:have emot conditions, number of persons in family unit, age of child in 1997, head: college grad

b. Dependent Variable: LOGINC

### Coefficients

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(Constant)</td>
<td>9.888</td>
<td>.090</td>
<td>109.399</td>
</tr>
<tr>
<td></td>
<td>age of child in 1997</td>
<td>1.713E-02</td>
<td>.007</td>
<td>.052</td>
</tr>
<tr>
<td></td>
<td>head: college grad</td>
<td>1.019</td>
<td>.058</td>
<td>.348</td>
</tr>
<tr>
<td></td>
<td>number of persons in family unit</td>
<td>5.179E-02</td>
<td>.019</td>
<td>.055</td>
</tr>
<tr>
<td></td>
<td>Emot st PCG:have emot conditions</td>
<td>-.501</td>
<td>.058</td>
<td>-.171</td>
</tr>
</tbody>
</table>

a. Dependent Variable: LOGINC
Answers:

#1.
A. Look to the $R^2$, the f and t tests to determine whether the model is any good. The independent variable in this model explains 3.7% of the variation in the dependent variable. This is a statistically significant effect. You can determine statistical significance by examining both the F and t tests.

B. You can use the $R^2$ value to determine the F value or you can use the mean squared regression value and the mean squared error to determine the value for F. The F value is statistically significant at the .000 level.

C. The standardized b indicates that for a one standard deviation unit increase in the number of kids, income will decrease by .193 standard deviation units.

D. If the number of kids increased by 1 SD unit, we would predict that the level of income would decrease by (.193) standard deviations. Because the standard deviation for income is $1000, the total change in income is (.193)*1000 = 193.

2. If these relevant variables that you fail to control for are correlated with the variables you have included in your model, the b coefficients you obtain will be biased. Also, if you fail to control for relevant variables in your model, your $R^2$ value will be lower than it otherwise will be, increasing the likelihood of failing to reject the null hypothesis for the model.

3.
A. There is no relationship between the IVs and DV in the population. This may because the error sums of squares is relatively high.

B. You have made type II errors – there is a relationship between these independent variables and the dependent variable but your sample is not representative of the population.

C. You have a small n, which may cause the standard error to be high for each of the independent variables, causing the b coefficients to be insignificant.

D. There may be a high level of correlation among your independent variables, causing your standard errors to be relatively high, causing the coefficient estimates to be insignificant.

E. There may be too many variables in your model, causing the standard error to be relatively high, causing the t value to be relatively high.

F. If relevant variables you have not included in your model are correlated with variables you have included in your model, the coefficient estimates are biased, perhaps causing the t value to be lower than it otherwise would be. (It is also possible that biased b coefficients will make the b coefficients higher than they otherwise would be, causing the t value to be higher than it otherwise would be.)

These are all the reasons I can give for not finding a significant relationship between independent and dependent variables.

4. Use a chi-square test. It’s a non-parametric test that is less powerful than a parametric
5. A. There is a positive relationship between living in a big city and being in poverty. Or, there is a negative relationship between living outside of a big city and not being in poverty.
B. \((r-1)(c-1) = (2-1)(2-1) = 1\).
C. 3.84
D. at the .000 level or the .001 level.

6. A. Yes. The Person Chi-Square value of 178.082 is statistically significant at the .000 level. There are 2 DF’s. For a 2-tailed, .05 test, the CV=5.99.
B. Those who grew up poor are more likely to be in poverty than what we would expect if there was no relationship between these variables. Those who grew up in middle class families are much less likely to be poor than what we would expect if there was no relationship between these variables. And those who grew up rich are less likely to be poor than what we would expect if there was no relationship between the variables.
C. To answer this question, you will need to do a residual analysis. The residual analysis takes the form:

\[
\frac{f_o - f_e}{\sqrt{f_e(1 - \text{rowproportion})(1 - \text{Columnproportion})}}
\]

In this case, let’s look at the inpov=1 and the grew up rich=1 case.

\[
\frac{324 - 329}{\sqrt{329(1-.222)(1-.185)}} = \frac{-5}{\sqrt{329*.778*.825}} = \frac{-5}{14.53} = -.344
\]

The absolute value for the critical value for this is 1.96. We will thus fail to reject the null hypothesis. This indicates that the group that grew up rich is not the driving force behind the significant relationship between economic situation while growing up and poverty status.

7.

Answer:

The expected proportion for wearing rings is 130/300=0.433
The expected proportion for not wearing rings is 170/300=0.5667
<table>
<thead>
<tr>
<th></th>
<th>Actual</th>
<th>Expected</th>
<th>(Actual-expected)²/expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>80</td>
<td>86.6</td>
<td>0.514</td>
</tr>
<tr>
<td>b</td>
<td>50</td>
<td>43.3</td>
<td>1.027</td>
</tr>
<tr>
<td>c</td>
<td>120</td>
<td>113.34</td>
<td>0.393</td>
</tr>
<tr>
<td>d</td>
<td>50</td>
<td>56.67</td>
<td>0.785</td>
</tr>
</tbody>
</table>

chi-square value = 2.719

DFs = 1

The critical value at the 5% level is 3.841. Therefore, there is not a significant relationship between gender and ring preference.

8.
Income: \(0.01 \times (200/1) = 2\)
Books: \(-0.50 \times (2/1) = -1\)
Zoo Visits: \(1.25 \times (1/1) = 1.25\)

Income: For each 1 SD change in income, the number of shoes changes in the same direction by 2 SDs.

Books: For each 1 SD increase in books, the number of shoes changes in the opposite direction by 1 SD.

Zoo Visits: For each 1 SD increase in zoo visits, the number of changes in the same direction by 1.25 SDs.

#9. The adjusted \(R^2\) value is given by the formula:

\[
\overline{R^2} = R^2 - \frac{k}{n-k-1} \times (1 - R^2)
\]

A.

\[
\overline{R^2} = 0.90 \times \frac{3}{21} \times 10 = 0.8857
\]

B.

\[
\overline{R^2} = 0.90 \times \frac{3}{6} \times 10 = 0.85
\]

C.
D. These indicate that the higher the number of observations and the lower the number of independent variables, the higher will be the adjusted $R^2$ value.

10.

For $r_{x1y,x2}$:

- For $b_{yx2}$:
  - $b_{yx2} = -1.500$
  - $a = 4.50$
- For $b_{x1x2}$:
  - $b_{x1x2} = -0.500$
  - $a = 3.50$

You must then determine the predicted values for $y$, and subtract these from the actual $y$'s to get the error terms.

$$Y_p = a + b_{yx2}x_2$$

- $y_{p1} = 4.5 - 1.5 \times 3 = 0$
- $y_{p2} = 4.5 - 1.5 \times 1 = 3$
- $y_{p3} = 4.5 - 1.5 \times 2 = 1.5$
- $y_{p4} = 4.5 - 1.5 \times 2 = 1.5$

where the $y_1$ stands for the first observation's value for $y$.

$$e = Y - Y_p$$

- $e_1 = 0 - 0 = 0$
- $e_2 = 3 - 3 = 0$
- $e_3 = 1 - 1.5 = -0.5$
- $e_4 = 2 - 1.5 = 0.5$

Let's call this set of $e$ values $e_1$.

We then have to determine the error terms in the second equation. Remember that $X_1$ is your dependent variable in this second equation.

$$X_{1p} = a + b_{x1x2}x_2$$
\[ X_{p1} = 3.5 - 0.5(3) = 2 \]
\[ X_{p2} = 3.5 - 0.5(1) = 3 \]
\[ X_{p3} = 3.5 - 0.5(2) = 2.5 \]
\[ X_{p4} = 3.5 - 0.5(2) = 2.5 \]

\[ x_1 = 1 \]
\[ x_2 = 2 \]
\[ x_3 = 3 \]
\[ x_4 = 4 \]

\[ e = X_1 - X_p \]

\[ e_1 = 1 - 2 = -1 \]
\[ e_2 = 2 - 3 = -1 \]
\[ e_3 = 3 - 2.5 = 0.5 \]
\[ e_4 = 4 - 2.5 = 1.5 \]

Let’s call this set of \( e \) values \( e_2 \).

You will now correlate the two error terms:

<table>
<thead>
<tr>
<th>Obs</th>
<th>( e_1 ) (from the first reg)</th>
<th>( e_2 ) (from the 2nd reg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>3</td>
<td>-0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>4</td>
<td>0.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Now, use the bivariate formula for determining \( r \) (you may need to look back in your notes to get this) to get the correlation between \( r_{yx1.x2} \), which is given by \( r_{e1e2} \)

\[
\begin{align*}
    r_{e1e2} &= \frac{4(0.5) - 0 \times 0}{\sqrt{[(4 \times 0.5 - 0)][4 \times 4.5 - 0]}} \\
    r_{e1e2} &= \frac{2}{\sqrt{36}} = 0.333
\end{align*}
\]

(You will not have to compute this \( r \) on the test.)

To get the partial b coefficient for \( b_{yx1.x2} \), you would do the following:

\[ e_1 = a + b \times e_2. \]
You would thus determine the \( b \) coefficient by using the bivariate \( b \) formula, where \( e_1 \) is \( Y \) and \( e_2 \) is \( X \), or the independent variable.

**Question D:**

\[
\begin{align*}
    b_{y|x_1} &= 0.400 \quad a=0.50 \\
    b_{x_2|x_1} &= -0.200 \quad a=2.50
\end{align*}
\]

In each of these, we allow \( x_1 \) to explain what it can in \( Y \) and in \( X_2 \). We will then regress the disturbance for the regression with \( Y \) on the regression with \( X_2 \).

\[
Y_i = 0.50 + 0.40x_i + e_i
\]

We now plug in the values for \( x_i \) and \( Y \) to determine the \( e_i \) values.
Our e1 values are given below:
1. Yp=.50+.40*1 = .9  Yi = 0  Yi-Yp = 0-.9 = -.9
2. Yp=.50+.40*2=1.3  Yi=3  Yi-Yp=3-1.3=1.7
3. Yp=.50+.40*3=1.7  Yi=1  Yi-Yp=1-1.7=-.7
4. Yp=.50+.40*4=2.1  Yi=2  Yi-Yp=2-2.1=-.1

We add up these disturbance terms for a value of 0.

For the second equation, we have:
b_{x2x1}=-0.200  \quad a=2.50
or
X_{2p}=2.50+(-.20)*x_1

The e2 values are given below:
1. X_{2p}=2.5+(-.2)*1=2.3  X_2=3  X_2-X_{2p}=3-2.3= .7
2. X_{2p}=2.5+(-.2)*2=2.1  X_2=1  X_2-X_{2p}=1-2.1= -1.1
3. X_{2p}=2.5+(-.2)*3=1.9  X_2=2  X_2-X_{2p}=2-1.9= .1
4. X_{2p}=2.5+(-.2)*4=1.7  X_2=2  X_2-X_{2p}=2-1.7= .3

We would then regress e1 on e2:
e1=a+b*e2.

We would use the bivariate regression formula to determine the value of b. This b value would be the same as the b value for b_{yx2.x1}

11. A. All of the coefficient estimates are significant except those for jailhd and NC.
   1. HDSMOKES. Those who have a head of household who smokes have incomes that are $16,891 less relative to those who do not have a head of household who smokes.
   2. Females have income levels that are $12,989 less than males.
   3. Those who live in the west have income levels that are $16,272 more than those who live in the north central and north eastern parts of the U.S.
   4. For each additional child, income rises by $3,270.
   5. For each additional year of age, income decreases by $342.

B. The primary reason for this is because the standard error for this coefficient estimate is high. The reason the coefficient estimate is high is because there are so few households with a head who is in jail (.22% of all households).

C. Yp= 81309+(-16891)+(-12989) = $51429.

D. Age of the head has a larger effect than kids – the standardized b coefficient is larger (in absolute value) than the standardized coefficient estimate for kids.

12. The difference in income between the two groups is 6 -- females have an income
level that is 6 higher than males. The mean income level for females is 15 and the mean income level for males is 9. The t value for females is 6, meaning that this is significant at the 1% level. The mean income levels for the two groups are significantly different from one another.

13. The difference in income between the two groups is 5 -- males have an income level that is 5 lower than females. The mean level of income for males is 3.2 and for females it is 8.2. The t value for the equation is 7.2; therefore, there the mean income level is significantly different for the two groups.

14. The slope for each group will be (0.5) age. The intercepts for the different groups are:

<table>
<thead>
<tr>
<th>Group</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big City</td>
<td>7.3</td>
</tr>
<tr>
<td>Urban</td>
<td>5.2</td>
</tr>
<tr>
<td>Suburban</td>
<td>2.2</td>
</tr>
<tr>
<td>Rural</td>
<td>1.0</td>
</tr>
</tbody>
</table>

The difference between each of the residential groups and rural residents is the B coefficient.

The t values for the coefficients are:

<table>
<thead>
<tr>
<th>Group</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big city</td>
<td>2.52</td>
</tr>
<tr>
<td>Urban</td>
<td>3.82</td>
</tr>
<tr>
<td>Suburban</td>
<td>1.2</td>
</tr>
<tr>
<td>Age</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Thus, all but suburban will be significant at the 5% level. Anything above 1.96 will be significant at the .05 level for a two-tailed, .05 test.

15. Spearman’s r. Use the formula:

\[
r_s = 1 - \frac{6 \sum_{i=1}^{n} D_i^2}{n(n^2 - 1)}
\]

<table>
<thead>
<tr>
<th>Student</th>
<th>Difference</th>
<th>D^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student 1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Student 2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Student 3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Student 4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Student 5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Student 6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sum</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>
Thus, $D^2 = 8$, $n = 6$.

\[ r_s = 1 - \frac{\sum_{i=1}^{n} d_i^2}{6(n^2 - 1)} = \]
\[ r_s = 1 - \frac{48}{210} = 1 - .22857 = .771 \]

Now use a z score to test for significance:

\[ z = \frac{r_s}{\sqrt{\frac{n-1}{1}}} = \frac{.771}{\sqrt{6-1}} = \frac{.771}{2.23} = .346 \]

Because $z$ is less than 1.96, you will fail to reject at the .05 level for a two-tailed test. For a one-tailed test, you could reject the null hypothesis for a .05 test.

**Using Kendall’s tau, use the formula:**

\[ \tau = \frac{S}{\sqrt{\frac{N(N-1)}{2}}} \]

<table>
<thead>
<tr>
<th>Student</th>
<th>Ranking in Reading Ability</th>
<th>Ranking in Socioeconomic Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student 1</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Student 2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Student 3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Student 4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Student 5</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Student 6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Put these in order according to one of the variables, here the rank of reading. The students will be put in a different order, according to their order for their reading skills.

<table>
<thead>
<tr>
<th>Student</th>
<th>Ranking in Reading Ability</th>
<th>Ranking in Socioeconomic Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student 3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Student 2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Student 5</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Student 4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Student 1</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Student 6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>
Here, A stands for the ranking in reading, B is for the students ranking in socioeconomic status. The small case letters stand for each of the students.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>

We then determine if the rankings are in the correct order ((1,2) is in the correct order and would get a value of +1, while (2,1) is not in the proper order and would get a ranking of -1).

Pair (a,b)=-1  
Pair (a,c)=+1  
Pair (a,d)=+1  
Pair (a,e)=+1  
Pair (a,f)=+1  
Pair (b,c)=+1  
Pair (b,d)=+1  
Pair (b,e)=+1  
Pair (b,f)=+1  
Pair (c,d)=-1  
Pair (c,e)=-1  
Pair (c,f)=+1  
Pair (d,e)=+1  
Pair (d,f)=+1  
Pair (e,f)=+1  

\[ S = +3 + 4 + (-1) + 2 + 1 = +9 \]

We then use the formula to determine the value of the correlation coefficient.

\[ t_a = \frac{9}{\sqrt{\frac{1}{2}(6-1)}} \]

\[ t_a = \frac{9}{\sqrt{\frac{1}{2}6(5)}} \]

\[ t_a = \frac{9}{15} = .6 \]

To test for significance, use the formula:
\[ z = \frac{S}{\sqrt{\frac{1}{18} \cdot n \cdot (n-1)(2n+5)}} \]

\[ z = \frac{9}{\sqrt{\frac{1}{18} \cdot 6 \cdot (6-1)(2 \cdot 6+5)}} \]

\[ z = \frac{9}{\sqrt{\frac{1}{18} \cdot 6 \cdot (5)(17)}} \]

\[ z = \frac{9}{\sqrt{28.33}} = 1.69 \]

We can reject at the .05 level for a one-tailed test but not for a two-tailed test.

#16. First, you can determine the total sums of squares by adding the explained and unexplained totals = 100+50=150. Therefore, the R\(^2\) value is 100/150, or .667. From here, use the F test to determine significance:

\[ F_{4,995} = z = \frac{.667 / 4}{.333 / 995} = \frac{.1667}{.0003346} = 498.0975 \]

The critical value is xxx for a .05 test and yyyy for a .01 test. You will reject at both levels of significance.

#17.

A. You already have the partial b coefficient estimates, so all you need is the intercept, which is given by

\[ a_{Y \cdot x_1, x_2} = \bar{Y} - b_{x_1, x_2} \bar{X}_1 - b_{x_2, x_1} \bar{X}_2 \]

In this case, a=1.5-.667*2.5 – (-.667)*2.0 = 1.1665

B. Use the formula:

\[ s_{b_{y \cdot x_1, x_2}} = \sqrt{\frac{s_{y, x_1, x_2}^2}{\sum x_1^2 - \left( \frac{\sum x_1}{n} \right)^2 \left( 1 - r_{x_1 x_2}^2 \right)}} \]

The error sums of squares is 1. n-k-1=4-2-1=1. Thus, the Mean Square Error, or the numerator in this case, is 1/1=1.

The denominator:
\( r^2 = (-.316)^2 = .10 \) (or \(.099856\)). \((1-.10)=.90\)

\[ 30 - 100/4 = 30-25 = 5 \]

\[ s_{b_{x1x2}} = \sqrt{\frac{1}{[30-25](.90)}} = \]

\[ s_{b_{x1x2}} = \sqrt{\frac{1}{5(.90)}} = \sqrt{\frac{1}{4.5}} = .471 \]

To determine the standard error for the coefficient estimate for x2 (which was not asked for in this question):

The denominator:

\[-18-64/4 = 18-16=2\]

\[ s_{b_{x2}} = \sqrt{\frac{1}{[18-16](.90)}} = .745 \]

C. To determine how much of the variance is being explained, use the total sums of squares and the explained sums of squares. We know that the unexplained sums of squares is 1. The total sums of squares is derived by:

\[ \sum(Y_i - \bar{Y})^2 \]

\[ \bar{Y} = 1.5 \]

<table>
<thead>
<tr>
<th>Yi</th>
<th>Mean for Y</th>
<th>Difference</th>
<th>Difference Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.5</td>
<td>1.5</td>
<td>2.25</td>
</tr>
<tr>
<td>1</td>
<td>1.5</td>
<td>.5</td>
<td>.25</td>
</tr>
<tr>
<td>3</td>
<td>1.5</td>
<td>1.5</td>
<td>2.25</td>
</tr>
<tr>
<td>2</td>
<td>1.5</td>
<td>.5</td>
<td>.25</td>
</tr>
<tr>
<td>Sum</td>
<td></td>
<td></td>
<td>5</td>
</tr>
</tbody>
</table>

Thus, the total sums of squares is 5. The explained sums of squares is 5-1=4.

\[ R^2 = 4/5 = .80 \]. We are explaining 80% of the variation in y with \( x_1 \) and \( x_2 \).

D. Use the t statistic to test for significance of the coefficient estimates.

For \( X_1 \): \( t_1 = 1.414 \)
For \( X_2 \): \( t_2 = 1.894 \).

Neither is larger than the critical value. We therefore fail to reject the null hypothesis.
E. To test the significance of the model, use an F test.

\[ F_{2,3} = \frac{.8/2}{2/1} = 2 \]

This F value is not greater than the critical value for F at the .01 or .05 levels. We therefore fail to reject the null hypothesis.

#18. There is an overall difference in means by race for the level of the county unemployment. We know this because the F statistic is significant at the .00 level. We can then determine which of the mean values by race are different from each other. In the Scheffe test, we see that the unemployment rate is lower for whites than it is for African Americans. The difference between the two groups is significant at the .000 level and the difference is .3575%. That is, the unemployment rate is .3575% lower in areas where Whites live relative to areas where African Americans live. There is also a difference in the unemployment rate between whites and ‘Other’ group. Whites have area unemployment rates that are 2.1632 percentage points lower than those who are of ‘Other’ races. Other differences exist between those who are African American and ‘Other’ races: the area unemployment rate for African Americans is 1.8057 points lower relative to ‘Other’ races. The same is true for Asian, Pacific Islander relative to ‘Other’ races, where the difference is 2.6311 percentage points.

#19. There is an overall difference in means by race for the level of wages for wives. The F test is significant at the .00 level. Differences between groups include the following: White wives have incomes that are $2.8849 higher than African American wives (significant at the .00 level). African American wives have wages that are $5.0816 less than ‘Other’ wives (significant at the .000 level).

#20.
A. To determine significance, need to run a chi-square test.


| Poor as an | Mid Income as an | Rich as an |

C:\WP60_1\LECT1.PHD\Final\Review Questions for the Final Exam1.doc
<table>
<thead>
<tr>
<th></th>
<th>adult</th>
<th>Adult</th>
<th>Adult</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grew Up Poor</td>
<td>50 (30)</td>
<td>30 (30)</td>
<td>10 (30)</td>
</tr>
<tr>
<td>Grew Up Mid Income</td>
<td>40 (43.3)</td>
<td>40 (43.3)</td>
<td>50 (43.3)</td>
</tr>
<tr>
<td>Grew Up Rich</td>
<td>10 (26.7)</td>
<td>30 (26.7)</td>
<td>40 (26.7)</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Chi-square value is:

<table>
<thead>
<tr>
<th>Cell</th>
<th>Observed</th>
<th>Expected</th>
<th>Difference</th>
<th>Diff Squared</th>
<th>Chi-square Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>50</td>
<td>30</td>
<td>20</td>
<td>400</td>
<td>400/30=13.3</td>
</tr>
<tr>
<td>b</td>
<td>30</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0/30=0</td>
</tr>
<tr>
<td>c</td>
<td>10</td>
<td>30</td>
<td>20</td>
<td>400</td>
<td>400/30=13.3</td>
</tr>
<tr>
<td>d</td>
<td>40</td>
<td>43.3</td>
<td>3.3</td>
<td>10.9</td>
<td>10.9/43.3=.25</td>
</tr>
<tr>
<td>e</td>
<td>40</td>
<td>43.3</td>
<td>3.3</td>
<td>10.9</td>
<td>10.9/43.3=.25</td>
</tr>
<tr>
<td>f</td>
<td>50</td>
<td>43.3</td>
<td>6.7</td>
<td>44.9</td>
<td>44.9/43.3=1.04</td>
</tr>
<tr>
<td>g</td>
<td>10</td>
<td>26.7</td>
<td>16.7</td>
<td>278.9</td>
<td>278.9/26.7=10.4</td>
</tr>
<tr>
<td>h</td>
<td>30</td>
<td>26.7</td>
<td>3.3</td>
<td>10.9</td>
<td>10.9/26.7=.41</td>
</tr>
<tr>
<td>i</td>
<td>40</td>
<td>26.7</td>
<td>13.3</td>
<td>176.9</td>
<td>176.9/26.7=6.6</td>
</tr>
</tbody>
</table>

\[ ?^2 = 13.3+0+13.3+.25+.25+1.04+10.4+.41+6.6 = 32.55 \]

DF=(3-1)(3-1)=2*2=4

At 4 DFs, the critical value is 9.49. Therefore, reject the null that there is no relationship between the variables.

B. It’s difficult to indicate the direction of the relationship when there are more than two categories for each of the variables. Here, we can say that there is a positive relationship between growing up poor and being poor as an adult, and a negative relationship between growing up poor and being rich as an adult. We could say just the opposite for those who grew up rich.

C. A residuals test for the adult middle income group for cell b shows the following:

Use the formula:

\[
\frac{f_o - f_e}{\sqrt{f_e(1-rowproportion)(1-columnproportion)}}
\]

For cell b: the numerator is 30-30, therefore we will end up with a value of 0 for the z score.

For cell e: the numerator is 40-43.3=-3.3. The denominator is:

The square root of (43.3)*(1-130/300)*(1-100/300) =

\[ \text{Sqrt}(43.3)*(.57)(.67) = 4.07 \]
Z=-3.3/4.07 =-.81.  The absolute value is less than 1.96, therefore you will fail to reject the null hypothesis.

Generally speaking, the chi-square value for the adult middle income group adds little to the total value of the chi-square for this test.

**Question 21.**

First regression: Allow x2 and x3 to explain what they can in Y.

\[ Y=a+b_2x_2+b_3x_3+e_1 \]

Second regression: Allow x2 and x3 to explain all they can in x1.

\[ X_1=a+b_2x_2+b_3x_3+e_2 \]

Because Y is the dependent variable, you will regress the residual of the Y equation on the residual of the X1 equation.

\[ e_1=a+b*e_2 \]

**Question 22:**

A. You may use log dependent variables because there are outliers in the dependent variable. Log dependent variables will decrease the effects of the outliers. You may therefore get a higher R^2 value when using log dependent variables, and thus have a better fitting model. Sometimes log dependent variables are used to help with the interpretation of indexed or other such dependent variables.

B. All of the coefficient estimates are estimated controlling for or paritalling out for the effects of all other variables included in the model.
   a. Age of child: \( b=0.01713, e^{0.01713} = 1.01718 \). For every one year increase in age of the child, income increases by 1.718%.
   b. College: \( b=1.019, e^{1.019} = 2.77042 \). (2.77042 - 1)*100 = 177.042% higher income relative to those who do not have a college degree.
   c. Persons in the family: \( b=0.05179, e^{0.05179} = 1.05315 \). (1.05315 - 1)*100 = 5.315%. For every additional person in the family, income increases by 5.315%.
   d. Emotional status of the primary caregiver. \( b=-0.501, e^{-0.501} = 0.60592 \). (+0.60592 - 1) = -.3941*100 = -39.41%. Those with emotional problems have incomes that are 39.41% lower relative to those who do not have such problems.

C. \[ Y_p=+9.888+.01713*10+1.019*0+.05179*4+(-.501)*0 \]

\[ Y_p=10.2665 \]

We then need to transform this by using the exponential function:
\[ e^{10.2665} = 28753.1 \]