Possible Solutions for Homework #4

Econ B2000, MA Econometrics

Kevin R Foster, CCNY

Previous hw questions weren't really the sort with answers so this is the first solution set that's relevant.

1. What are the names of the people in your study group?

2. Go back to the PUMS data and do a few interesting linear regressions. Clearly explain the null hypothesis and statistical tests for them.
   The next questions are from last year's exam.

3. A recent report asserted that people who worked more hours also tended to be fatter (among those in certain occupations). (The paper doesn't give precise numbers so I'll make them up – don't bother with Google.) The paper did much more econometric analysis of course. Nevertheless, suppose that, of the 7219 women working non-strenuous occupations, 23% are working more than 40 hours/week. Of those women in non-strenuous occupations working more than 40 hours/week, 27.3% were obese; of those women in non-strenuous occupations working less than 40 hours/week, 24.6% were obese. There were also 734 women in strenuous occupations with 22% working more than 40 hours/week. Of the women in strenuous occupations working more than 40 hours/week, 28.1% were obese while 37.4% were obese among those working fewer hours. Does it seem likely that overtime makes certain groups more likely to be obese? J Abramowitz, "Working Hours, Body Mass Index, and Health Status: A Time Use Analysis"

   Among females in non-strenuous occupations, there were 24.6% vs 27.3% obese; diff is 2.7%; std err is 1.2% so tstat is -2.18, p-value is .029.

   Among females in strenuous occupations, there were 37.4% vs 28.1% obese; diff is 9.3%, std err is 4.2%, tstat is 2.22, p-value is .027.

4. I used the CEX data to look at the fraction of spending going to health insurance. I get the following table, grouped by education of the reference person:

<table>
<thead>
<tr>
<th>%Insurance</th>
<th>No HS</th>
<th>HS diploma</th>
<th>Some college, no degree</th>
<th>Assoc degree</th>
<th>Bach degree</th>
<th>Adv degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 10%</td>
<td>467</td>
<td>1385</td>
<td>1191</td>
<td>615</td>
<td>1181</td>
<td>521</td>
</tr>
<tr>
<td>11% - 20%</td>
<td>82</td>
<td>231</td>
<td>157</td>
<td>71</td>
<td>122</td>
<td>58</td>
</tr>
<tr>
<td>21% - 30%</td>
<td>21</td>
<td>65</td>
<td>27</td>
<td>10</td>
<td>32</td>
<td>7</td>
</tr>
<tr>
<td>more than 30%</td>
<td>8</td>
<td>18</td>
<td>14</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

a. Conditional on the reference person having a college degree (Associate's, Bachelor's or Advanced), what fraction devote more than 20% of spending to health insurance? 2.09%

b. Conditional on the reference person having less than a college degree, what fraction spend more than 20% on health insurance? 4.17%

c. Is this difference statistically significant? t-stat is almost 5, yes
d. What is the overall share (in this sample) of people with any college degree? What share of people spending more than 20% is made up of people with any college degree? 72% vs 36%
e. Are those break points (+/- 20%; any degree) reasonable? Can you suggest better? Explain.
f. What problems might there be, with the classification and analysis here? Can you do better with the CEX data?
g. Consider the following simple distribution questions. Your answer can be as detailed as necessary. You might sketch each case.
   a. For a Normal Distribution with mean 2 and standard deviation 5.6, what is area to the right of 10.4? \(0.0668\)
   b. For a Normal Distribution with mean -1 and standard deviation 4.9, what is area to the right of -7.86? \(0.9192\)
   c. For a Normal Distribution with mean 10 and standard deviation 5.2, what is area to the left of 0.12? \(0.0287\)
   d. For a Normal Distribution with mean 7 and standard deviation 4, what is area to the left of 9.8? \(0.758\)
   e. For a Normal Distribution with mean 11 and standard deviation 8.4, what is area in both tails farther from the mean than 19.4? \(0.3173\)
   f. For a Normal Distribution with mean -8 and standard deviation 0.4, what is area in both tails farther from the mean than -7.52? \(0.2301\)
   g. For a Normal Distribution with mean 11 and standard deviation 6.3, what is area in both tails farther from the mean than 0.29? \(0.0891\)
   h. For a Normal Distribution with mean 10 and standard deviation 1.5, what is area in both tails farther from the mean than 12.7? \(0.0719\)
   i. For a Normal Distribution with mean -1 and standard deviation 1.2, what is area in both tails farther from the mean than 1.52? \(0.0357\)
   j. For a Normal Distribution with mean 9 and standard deviation 4.9, what values leave probability 0.08 in both tails? \(0.422 17.578\)
   k. For a Normal Distribution with mean 3 and standard deviation 9.6, what values leave probability 0.311 in both tails? \(-6.726 12.726\)
   l. A regression coefficient is estimated to be equal to 8.25 with standard error 7.5; there are 35 degrees of freedom. What is the p-value (from the t-statistic) against the null hypothesis of zero? \(0.2788407\)
   m. A regression coefficient is estimated to be equal to 10.08 with standard error 5.6; there are 7 degrees of freedom. What is the p-value (from the t-statistic) against the null hypothesis of zero? \(0.1148845\)
   n. A regression coefficient is estimated to be equal to -12.04 with standard error 8.6; there are 14 degrees of freedom. What is the p-value (from the t-statistic) against the null hypothesis of zero? \(0.1832826\)

5 Consider the following simple distribution questions. Your answer can be as detailed as necessary. You might sketch each case.
   a. For a Normal Distribution with mean 2 and standard deviation 5.6, what is area to the right of 10.4? \(0.0668\)
   b. For a Normal Distribution with mean -1 and standard deviation 4.9, what is area to the right of -7.86? \(0.9192\)
   c. For a Normal Distribution with mean 10 and standard deviation 5.2, what is area to the left of 0.12? \(0.0287\)
d. For a Normal Distribution with mean 7 and standard deviation 4, what is area to the left of 9.8? 0.758

e. For a Normal Distribution with mean 11 and standard deviation 8.4, what is area in both tails farther from the mean than 19.4? 0.3173

f. For a Normal Distribution with mean -8 and standard deviation 0.4, what is area in both tails farther from the mean than -7.52? 0.2301

g. For a Normal Distribution with mean 11 and standard deviation 6.3, what is area in both tails farther from the mean than 0.29? 0.0891

h. For a Normal Distribution with mean 10 and standard deviation 1.5, what is area in both tails farther from the mean than 12.7? 0.0719

i. For a Normal Distribution with mean -1 and standard deviation 1.2, what is area in both tails farther from the mean than 1.52? 0.0357

j. For a Normal Distribution with mean 9 and standard deviation 4.9, what values leave probability 0.08 in both tails? 0.422 17.578

k. For a Normal Distribution with mean 3 and standard deviation 9.6, what values leave probability 0.311 in both tails? 6.726 12.726

l. A regression coefficient is estimated to be equal to 8.25 with standard error 7.5; there are 35 degrees of freedom. What is the p-value (from the t-statistic) against the null hypothesis of zero? 0.2788407

m. A regression coefficient is estimated to be equal to 10.08 with standard error 5.6; there are 7 degrees of freedom. What is the p-value (from the t-statistic) against the null hypothesis of zero? 0.1148845

n. A regression coefficient is estimated to be equal to -12.04 with standard error 8.6; there are 14 degrees of freedom. What is the p-value (from the t-statistic) against the null hypothesis of zero? 0.1832826