Basic Skills Learned in This Course
Econ B2000
Kevin R Foster, CCNY
Fall 2013

By the end of the class you should have the following skills (the exams will test these):

**General Outcome:** Students will be able to apply mathematically rigorous analysis to topics such as analyzing data, hypothesis testing, and regression analysis.

Before beginning the course, students are expected to know how to:

1. create and interpret basic statistics on large datasets, such as
   - mean
   - median
   - measures of spread such as standard error

2. create and interpret data tabulations including
   - crosstabs of counts and fractions
   - marginal and conditional probabilities
   - conditional means

In the course students will learn topics in these basic areas:

3. conducting hypothesis tests for equality of means and regression t-tests including
   - calculating areas under t and normal distributions
   - calculating t-values
   - getting critical values
   - creating confidence intervals
   - determining p-values
   - explaining significance test results including Type I/Type II error

4. determining regression coefficients using statistical software such as SPSS
   - explaining the coefficient estimates as slope values
   - testing statistical significance of these estimates
   - with datasets with thousands of observations

5. estimating extensions of the basic linear regression including:
   - heteroskedasticity-consistent standard errors
   - nonlinear effects
   - binary dependent variable models, where students must be able to calculate changing marginal effects
   - panel data
   - instrumental variables to address endogeneity
   - autocorrelation

Examples
Topic Area 2
Using ATUS data from 2003-2009, we look at the crosstabs of race and ethnicity; this gives the number of each group:

<table>
<thead>
<tr>
<th></th>
<th>Native American Indian / Inuit / Hawaiian</th>
<th>Asian</th>
<th>African-American</th>
<th>White</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Hispanic</td>
<td>1440</td>
<td>2834</td>
<td>12385</td>
<td>69721</td>
<td>86380</td>
</tr>
<tr>
<td>Hispanic</td>
<td>325</td>
<td>77</td>
<td>337</td>
<td>11659</td>
<td>12398</td>
</tr>
<tr>
<td>Total</td>
<td>1765</td>
<td>2911</td>
<td>12722</td>
<td>81380</td>
<td>98778</td>
</tr>
</tbody>
</table>

The fractions of each demographic category are:

<table>
<thead>
<tr>
<th></th>
<th>Native American Indian / Inuit / Hawaiian</th>
<th>Asian</th>
<th>African-American</th>
<th>White</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Hispanic</td>
<td>0.014578145</td>
<td>0.0286906</td>
<td>0.1253822</td>
<td>0.7058353</td>
<td>0.8744862</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.003290206</td>
<td>0.0007795</td>
<td>0.0034117</td>
<td>0.1180324</td>
<td>0.1255138</td>
</tr>
<tr>
<td>Total</td>
<td>0.017868351</td>
<td>0.0294701</td>
<td>0.1287939</td>
<td>0.8238677</td>
<td>0.9162518</td>
</tr>
</tbody>
</table>

Conditional by row:

<table>
<thead>
<tr>
<th></th>
<th>Native American Indian / Inuit / Hawaiian</th>
<th>Asian</th>
<th>African-American</th>
<th>White</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Hispanic</td>
<td>0.016670526</td>
<td>0.0328085</td>
<td>0.1433781</td>
<td>0.8071429</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.026213905</td>
<td>0.0062107</td>
<td>0.0271818</td>
<td>0.9403936</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.017868351</td>
<td>0.0294701</td>
<td>0.1287939</td>
<td>0.8238677</td>
<td>0.9162518</td>
</tr>
</tbody>
</table>

So 14% of non-Hispanics are African-American while just 2.7% of Hispanics are African-American.

Conditional by column:

<table>
<thead>
<tr>
<th></th>
<th>Native American Indian / Inuit / Hawaiian</th>
<th>Asian</th>
<th>African-American</th>
<th>White</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Hispanic</td>
<td>0.815864023</td>
<td>0.9735486</td>
<td>0.9735105</td>
<td>0.8567338</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.184135977</td>
<td>0.0264514</td>
<td>0.0264895</td>
<td>0.1432662</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.900000000</td>
<td>1.0000000</td>
<td>1.0000000</td>
<td>0.9100000</td>
<td></td>
</tr>
</tbody>
</table>

Alternately, 97% of African-Americans are not Hispanic while just 86% of whites are not Hispanic. Native Americans are the most Hispanic ethnic group.

Topic Areas 1 & 4
Using 2010 CPS data, restrict to only fulltime workers with a non-zero wage. Regression will have earnings (annual wage and salary) as the dependent variable.

The first set of basic explanatory variables is hypothesized to be factors such as age, sex, education, race/ethnicity, marital status, veteran status, and if a union member.

Average values of regression variables, for this subset, are:
Wage/Salary (annual) $ 49,773.79
Age 41.88
Female 44.5%
White 79.7%
African-American 11.8%
Asian-American 5.8%
Native American/ Indian/ Alaskan/ Inuit/ Hawaiian 2.8%
Hispanic 16.1%
   Mexican 9.8%
   Puerto Rican 1.4%
   Cuban 0.6%
Immigrant 17.5%
1 or more Parents were immigrants 23.8%
Education: no high school 8.6%
Education: High School Diploma 28.9%
Education: Some College (incl no degree or Assoc degree) 27.9%
Education: Some College but no degree 17.5%
Education: Associate in vocational 5.0%
Education: Associate in academic 5.4%
Education: 4-yr degree 22.5%
Education: Advanced Degree 12.1%
Married 62.0%
Divorced or Widowed or Separated 14.8%
Unmarried 23.2%
Union member 2.2%
Veteran (any) 7.4%

The regression estimates are made with three basic specifications: Spec 1 has just the listed variables; Spec 2 included dummies for industry, occupation, and state of residence; Spec 3 has dummy interactions for female*age, African-American*age, female*African-American*age, Hispanic*age, female*Hispanic*age, and female*education. An asterisk indicates statistical significance.

<table>
<thead>
<tr>
<th></th>
<th>Spec 1</th>
<th>Spec 2</th>
<th>Spec 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-$28,685.56</td>
<td>$13,744.52</td>
<td>-$10,978.43</td>
</tr>
<tr>
<td></td>
<td>1954.106</td>
<td>3025.180</td>
<td>3685.959</td>
</tr>
<tr>
<td>Age</td>
<td>$2,517.92</td>
<td>$2,012.04</td>
<td>$3,052.09</td>
</tr>
<tr>
<td></td>
<td>93.814</td>
<td>88.514</td>
<td>133.158</td>
</tr>
<tr>
<td>Age-squared</td>
<td>-$23.60</td>
<td>-$18.55</td>
<td>-$29.40</td>
</tr>
<tr>
<td></td>
<td>1.055</td>
<td>.994</td>
<td>1.504</td>
</tr>
<tr>
<td>Female</td>
<td>-$17,380.74</td>
<td>-$14,587.20</td>
<td>$26,912.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
African American  \(-6,136.77\)  \(-5,315.62\)  \(17,924.27\)  \\
Asian  \(-783.89\)  \(-3,160.09\)  \(-3,196.33\)  \\
Native American Indian or Alaskan or Hawaiian  \(-4,615.72\)  \(-3,077.92\)  \(-3,030.05\)  \\
Hispanic  \(-5,776.56\)  \(-4,433.05\)  \(32,492.36\)  \\
Immigrant  \(-7,377.88\)  \(-4,669.63\)  \(-4,080.20\)  \\
1 or more parents were immigrants  \(4,513.48\)  \(1,231.87\)  \(892.78\)  \\
Education: High School Diploma  \(7,658.27\)  \(3,819.68\)  \(4,208.53\)  \\
Education: Some College but no degree  \(15,430.94\)  \(7,791.73\)  \(9,434.14\)  \\
Education: Associate in vocational  \(15,719.42\)  \(8,376.06\)  \(9,873.19\)  \\
Education: Associate in academic  \(19,907.99\)  \(9,660.31\)  \(11,310.63\)  \\
Education: 4-yr degree  \(35,565.50\)  \(20,756.84\)  \(24,651.87\)  \\
Education: Advanced Degree  \(63,729.94\)  \(40,911.95\)  \(46,708.57\)  \\
Married  \(8,100.77\)  \(7,074.38\)  \(6,912.90\)  \\
Divorced or Widowed or Separated  \(1,646.98\)  \(1,893.12\)  \(1,881.97\)  \\
Union member  \(-3,992.75\)  \(2,282.96\)  \(2,372.64\)  \\
Veteran (any)  \(-1,186.63\)  \(-884.41\)  \(-905.22\)  \\
R-squared  \(0.213\)  \(0.315\)  \(0.319\)  \\

Sample age-wage profiles are shown below, for a white male with just a high-school diploma, unmarried, neither immigrant, veteran nor union member. The estimated peak earning year is 53 in Specification 1, 54 in Specification 2, and 52 in Specification 3.