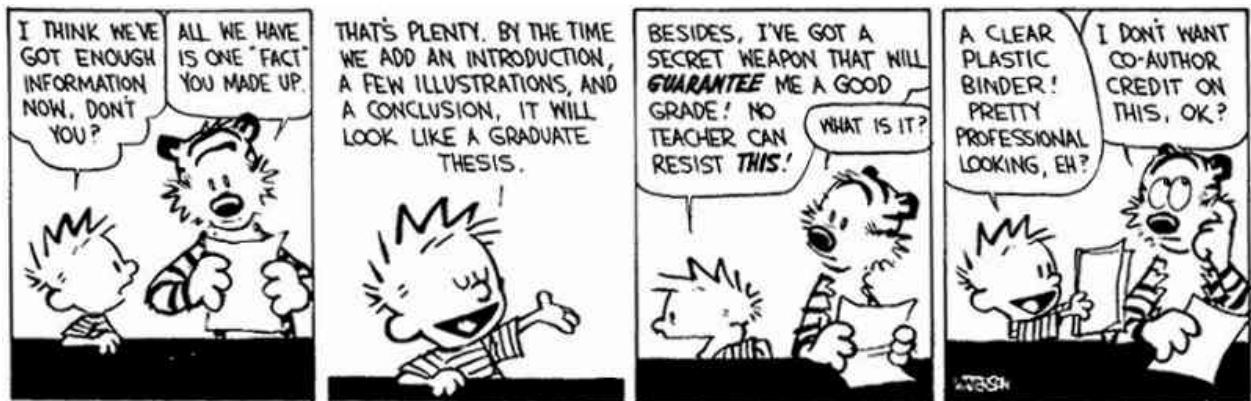


IS/AAE374 Discussion Section  
P-Set 2, T-Test and Scatter Plots

**Problem Set 1** – The answer key is available on the web site.

**1. Presentation and analysis**



Binders, blank front pages (lower your carbon footprint!), colors ...may be too much. BUT, having a well organized problem set is appreciated.

Don't be lazy with the analysis! We are trying to get your insight and identify whether you understand useful concepts or not, so typing two-word answers does not seem enough.

**2. Some problems that came up constantly**

- 1c. Different ratios of max production in North and South?  
Comparative Advantage (Initial endowment & Factor intensity)
- 1d. Marginal productivity  
e.g.,  $(25)^{0.2}(41)^{0.8} - (25)^{0.2}(40)^{0.8} = 0.73$ .
- 2c. Scarcity and real economic prices, and increasing relative prices  
Diminishing returns to scale vs. Diminishing returns to one factor (Comparative advantage has nothing to do with!!)
- 3. Factor price equalization theorem

**Problem Set 2. What you need to know:**

**1. Representing distributions (Handouts 2 & 3)**

- Histograms
- Measures of Central Tendency (mean, median)
- Measures of Spread (variance, standard deviation)

**2. Mean comparison between two groups (population)**

t-test: Null hypothesis  $H_0 : \mu_1 = \mu_2$

Alternative hypothesis  $H_a : \mu_1 \neq \mu_2$

T-Tests are used to compare differences between two groups or segments of the population. In this case they will allow us to check if the difference in means between two segments of the population is statistically significant.

- Go to stat2.exe and select the tab labeled enrollment2000.
- Divide the sample of countries in two according to GDP per capita (countries above and below the median). For each, estimate the mean for infant mortality rate.
- In a separate cell, insert the function “ttest”; a dialogue box should show up.
  - o Arrays: In “array 1” select the first group of countries (low income), and do the same for “array 2” by selecting the other group.
  - o Tails: If you predicted that the mean of group 1 would be lower than that of group 2, pick 1 tail. If you predicted that the mean of group 2 would lower than that of group 1, pick one tail. If you didn’t predict which would be higher, use 2 tails. In this case, select 1 tail.
  - o Type 2: Select type 2 t-test (equal variances); trust me on this one.
- The output of this function represents the p-value associated to the t-test. In this case, a low p-value (0.05) means that there is evidence to reject the null hypothesis; i.e. that the difference in the means is statistically significant.
- You can also estimate the t value as follows (don’t get scared by the expression):

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{X_1X_2} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Where

$$S_{X_1X_2} = \sqrt{\frac{(n_1 - 1)S_{X_1}^2 + (n_2 - 1)S_{X_2}^2}{n_1 + n_2 - 2}}$$

- T-value

- $|t| > 1.64$  reject the null hypothesis at a 90% confidence level
- $|t| > 1.96$  reject the null hypothesis at a 95% confidence level

### 3. Growth Rates (Handout #1)

### 4. Scatter Plots

- Two way graph & Trend line
- Display the data from two variables in the Cartesian plane. Each point in the graph will be determined by the value of the two variables: one variable will determine the position on the horizontal axis and the value of the other variable the position on the vertical axis.
- It suggests a relationship between two variables. In fact, if you draw a trendline for the scatter plot you get a first measure for the correlation of the variables.
- Go to stat2.exe and select the tab labeled enrollment2000. Draw a scatter plot between infant mortality and child enrollment and include the trendline. What is the relation between infant mortality and child school enrollment?
- Covariance & Correlation ( $-1 \leq \rho \leq 1$ )

$$\rho = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_i (X_i - \bar{X})^2} \sqrt{\sum_i (Y_i - \bar{Y})^2}}$$