14.32 Recitation 5

Paul Schrimpf

March 6, 2009

- Problem set 2 solutions
- Difference in log $\approx$ percent change
- Regression overview

1 Difference in logs $\approx$ percent change

We have stated in class many times that the difference in logs is an approximate percent change. Let's see exactly why that is. First, we'll just think about comparing the average of logs between two groups. After that, we'll think about regression with the dependent variable in logs. Let $\beta = E[\log y_1] - E[\log y_2]$. The basic idea is that:

$$\beta \approx e^{\beta} - 1 \approx \frac{E[y_1]}{E[y_0]} - 1$$

The first $\approx$ comes from a Taylor expansion of e^{β} around $\beta = 0$. That is,

$$e^{\beta} = 1 + \beta + \frac{\beta^2}{2} + o(\beta^2) \tag{1}$$

It will be more accurate the smaller β is. The second $\approx$ comes from saying that $\exp\left(E[\log y_1] - E[\log y_0]\right) \approx \frac{E[y_1]}{E[y_0]}$. This approximation is exact when there is no variance in y_1 and y_0 and when $\frac{E[y_1]}{E[y_0]} = \frac{e^{E[\log y_1]-E[y_1]}}{e^{E[\log y_0]-E[y_0]}}$. The later would be the case when y_0 and y_1 have the same type of distribution but one is rescaled. For example, if $y_0 \sim Exp(\lambda)$ and $y_1 \sim Exp(a\lambda)$

In summary, $\beta = E[\log y_1] - E[\log y_0] \approx \frac{E[y_1]-E[y_0]}{E[y_0]}$ and this approximation is better (i) the smaller β is and (ii) the closer $e^{E[\log y_k]}$ is to $E[y_k]$.

The following table shows the quality of the approximation $\beta \approx e^{\beta} - 1$ as a function of β:

β	0	0.01	0.05	0.1	0.2	0.3	0.4	0.5
$e^{\beta} - 1$	0	0.01	0.051	0.11	0.22	0.35	0.49	0.65
Error	0	5e-05	0.0013	0.0052	0.021	0.05	0.092	0.15

In general, the log approximation understates the percent change.

Regression with Logs When you estimate a regression with dependent variable in logs, say

$$\log y = \alpha + x\beta + \epsilon$$

Then β is the approximate percent change in $E[y|x]$ per unit change in x. The reason is exactly that given above.

2 Regression overview

These are some important facts about regression. Some we have already seen, others we will cover soon. I'll go through some of the details in recitation.

- Population regression: $\min_{\alpha,\beta} E\left[(y-\alpha-x\beta)^2\right]$ gives $\alpha = E[Y] - \beta E[X]$ and $\beta = \frac{COV(X,Y)}{V(X)}$
- Best linear approximation to CEF: population regression also solve $\min_{\alpha,\beta} E\left[(E[y|x]-\alpha-x\beta)^2\right]$
- Sample regression: $\min_{\alpha,\beta} \sum (y_i-\alpha-x_i\beta)^2$ gives $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$, $\hat{\beta} = \frac{\sum(y_i-\bar{y})(x_i-\bar{x})}{\sum(x_i-\bar{x})^2}$
- Variance of OLS: $V(\hat{\beta}) = \frac{\sigma_\epsilon^2}{\sigma_x^2}$
- Gauss-Markov theorem: under the classical regression assumptions, OLS is the best linear unbiased estimator. That is among all estimators that are linear in y, $\tilde{\beta} = \sum z_i y_i$ where z_i is potentially some function of x and unbiased, $E[\tilde{\beta}] = \beta$, OLS has the smallest variance.
- Firsch-Waugh Theorem / partialing out: the following are equivalent ways to estimate $\hat{\beta}_1$
  - Multiple regression: $\min_{\alpha,\beta_1,\beta_2} \sum (y_i - \alpha - x_{i1}\beta_1 - x_{i2}\beta_2)^2$
  - Partial out x_{i2}: regress y on x_2, call the residuals $\hat{e}_y$. Regress x_1 on x_2, call the residuals $\hat{e}_x$. Regress $\hat{e}_y$ on $\hat{e}_x$. The coefficient on $\hat{e}_x$ is $\hat{\beta}_1$.