Wooldridge ECM Learning Notes 00: Pre Knowledge

Avienya JANE

September 2022

These are some of my personal understandings while studying the Wooldridge material. It is not suitable for first-time study. The content of this text corresponds to Appendix A, B and C. Due to the heavy course load and time constraints, I apologize in advance for any possible intellectual and linguistic errors in the text.

1 Proportions, Percentages and the Natural Logarithm

We denote $\%\Delta x = 100(\Delta x/x_0)$ as the percentage change, which is read as "the percentage change in x".

Notice that $log(1+x) \approx x \, if \, x \asymp 0$. Let x_1 and x_0 be positive numbers. Then we get $log(x_1) - log(x_0) \approx (x_1 - x_0)/x_0 = \Delta x/x_0$. Thus, $100\Delta log(x) \approx \%\Delta x$ for small changes in x. This is important if we consider the conception of **elasticity** in economics which is defined as, the elasticity of y with respect to x is the percentage change in y when x increases by 1%. It's obvious that elasticity depends on the value of x if we write the equation in a linear form. But if we use the log version, the elasticity is approximately equal to $\Delta log(y)/log(x)$, meaning that β_1 is exactly the elasticity of y with respect to x. The introduction of elasticity is essential when explaining coefficients in regression models.

We call equation in the form of

$$log(y) = \beta_0 + \beta_1 log(x) \tag{1}$$

as a **constant elasticity model** and it plays an important role in empirical ECM. Similarly we define the **semi-elasticity** which is the percentage change in y when x increase by one *unit*.

2 Conditional Distributions

2.1 Conditional Expectation

The essence of conditional expectations is that $E(Y|X)$ is just some function on x telling how Y varies with x.Example is the sum statistics of Mincer Equation. We can generate different average hourly wage by various level years of education. If suppose that $E(WAGE|EDUC) = 1.05 + .45EDUC$, the average wage for people with 8 years of education is $1.05 + 0.45 \cdot 8 = 4.65$\$. The coefficient implies that each year of education increases the expected hourly wage by 0.45\$,

Then we state some basic properties of conditional expectations.

Property 2.1.1 $E[c(X)|X = c(X)]$. We emphasize the intuition, if we know X, we also know X^2.

Property 2.1.2 If X and$Y$ are independent, then we have $E(Y|X) = E(Y)$. A special case for this property is, when $E(Y) = 0$, then $E(U|X) = 0$, which serves as an essential point in OLS model.

Property 2.1.3 (Law of Iterated Expectations) $E[E(Y|X)] = E(Y)$
Proof.

$$R.H.S. = E(Y) = \int_{-\infty}^{\infty} y f_y(y) dy$$

$$\begin{aligned} L.H.S. &= E_X(\int_{-\infty}^{\infty} y \frac{f(x,y)}{f_x(x)} dy) \\ &= \int_{-\infty}^{\infty} f_x(x) \int_{-\infty}^{\infty} y \frac{f(x,y)}{f_x(x)} dydx \\ &= \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f(x,y) dydx \\ &= \int_{-\infty}^{\infty} y f_y(y) dy \end{aligned}$$

Property 2.1.3' (General Version of Law of Iterated Expectations) $E(Y|X) = E[E(Y|X,Z)|X]$. Meaning that if we want to find $E(Y|X)$, we can firstly find $E(Y|X,Z)$ for any other random variable Z, and then find the expected value conditional on X of $E(Y|X,Z)$.

Property 2.1.4 If $E(Y|X) = E(Y)$, then $Cov(X, Y) = 0$. Proof.

$$\begin{aligned}
Cov(X, Y) &= E[(X - E(X))(Y - E(Y))] \\
&= E_X E_Y[(X - E(X)(Y - E(Y)|X] \\
&= E_X[(X - E(X))E_Y[(Y - E(Y)|X]] \\
&= E_X[(X - E(X))(E(Y|X) - E(Y))] \\
&= E_X[(X - E(X)) \cdot 0] \\
&= 0
\end{aligned}$$

Reconsider the special case in property 2 and we get an important conclusion for OLS. if U and X are random variables and $E(U|X) = 0$, then $E(U) = 0$ with U and X are uncorrelated.

2.2 Conditional Variance

The calculation is $Var(Y|X = x) = E(Y^2|x) - (E(Y|x))^2$.
Property 2.2.1 If X and Y are independent, then $Var(Y|X) = 0$.
Due to time limit, I show other properties of conditional variance and conditional covariance in figure below.

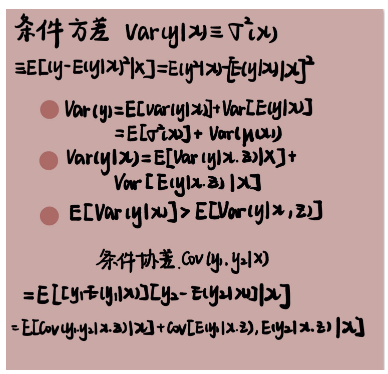

Figure 1: Property of CV and CCV

3 Unbiasedness, Efficiency and Consistency

Consistency is the minimal requirement of an estimator since if not, it won't help us to learn about θ no matter how large the sample size. Unbiased estimators are not necessarily consistent. Those whose variance shink to 0 as the sample size grows are consistent. Unbiasedness is something about the expectation and Consistency is that of variance. We can use three facts about probability limits to get other different consistent estimators.