Statistics
Tutorial 05

Philipp Scherer & Jens Wiederspohn

27.05.2020

Disclaimer!

- **The content of the slides partly relies on material of Philipp Prinz, a former Statistics tutor. Like us, he's just a student. Therefore we provide no guarantee for the content of the slides or other data/information of the tutorial.**
- **Please note that the slides will not cover the entire lecture content. To pass the exam, it is still absolutely necessary to deal with the Wooldridge in detail!**

OLS asymptotics

- So far, we covered only finite sample OLS properties
  - Gauss-Markov assumptions 1-5 hold for any sample size n
  - Without normality assumption, for any sample size:
    - Distribution of a t statistic is not exactly t
    - F statistic does not have an exact F distribution
  - What happens to OLS estimators/test statistics **as n grows**?
- OLS has satisfactory large sample properties
  - For large sample sizes t and F statistics have approximately t and F distributions, even without normality assumption

Consistency

- Thought experiment: what happens if n approaches infinity?
  - Consistent, if $\hat{\beta}$ approaches β as n increases

$$
\begin{aligned}
\operatorname*{plim}_{n\to\infty} \hat{\beta}_1 &= \beta_1 + \frac{\overbrace{Cov(x_1, u)}^{0 \text{ (MLR.4')}}}{Var(x_1)} \\
&= \beta_1
\end{aligned}
$$

- Consistency is minimal requirement for an estimator
  - If more data $\neq$ better estimates $\rightarrow$ poor estimation procedure
  - Consistency is built on same assumptions as unbiasedness, but OLS is also consistent in the simple regression and general case if we assume only zero correlation instead of zero conditional mean.
- Sidenote: plim = 'probability limit'
  - $\lim_{n\to\infty} Pr(|\hat{\beta}_1 - \beta_1| \geq \epsilon) = 0,$ for all $\epsilon > 0$
  - Probability that $\hat{\beta}_1 - \beta_1 \neq 0$ goes to 0 as n approaches ∞
    - For large n, it is very likely that $\hat{\beta}_1 = \beta_1$

Inconsistency

- If $Cov(x_j, u) \neq 0 \rightarrow$ OLS is biased and inconsistent
  - Bias persists as the sample size grows
  - Affects <u>all</u> OLS estimators if any x_j is correlated with u (Including intercept β_0)
- Asymptotic bias if we omit x_2 from model with $k = 2$:

$$\begin{aligned} \operatorname*{plim}_{n\to\infty} \hat{\beta}_1 - \beta_1 &= \frac{Cov(x_1, u)}{Var(x_1)} \\ &= \beta_2 \frac{Cov(x_1, x_2)}{Var(x_1)} \\ &= \beta_2 \delta_1 \end{aligned}$$

  - With more data, OLS estimtator approaches $\beta_1 + \beta_2\delta_1$
- Practically, inconsistency = bias
  - Inconsistency is expressed in terms of population parameters
  - Bias is based on sample parameters

Normality

- Consistency is not enough for statistical inference
  - Requires information about sampling distribution of estimator
  - Inference based on t- and F-statistic requires normality
  - Statements about unbiasedness/variance do not require MLR.6
- Normality of estimators depends on normality of u
  - Non-normal u_i: statistics will not have t-/F-distributions
- Usually u is unobserved $\rightarrow$ check y instead
  - Linear functions of normally distr. RVs follow normal distr.
  - $y = \beta_0 + \beta_1 x_1 + u \rightarrow$ linear function
- What if the y in our data are not distributed normally?
  - No need for despair $\rightarrow$ CLT to the rescue!
  - If n is large, y is approximately normally distributed
    $\rightarrow$ OLS estimtors satisfy **asymptotic normality**
    - We need not assume normality of u_i as long as $Var(u_i)$ is finite
    - Allows use of t-statistic, even if MLR.6 does not hold

Asymptotic normality of OLS

From an asymptotic point of view:

- It does not matter that we need to replace $sd(\hat{\beta}_j)$ with $se(\hat{\beta}_j)$
- Standardized $\hat{\beta}$ has asymptotic standard normal distribution
- $\frac{(\hat{\beta}_j - \beta_j)}{sd(\hat{\beta}_j)}$ and $\frac{(\hat{\beta}_j - \beta_j)}{se(\hat{\beta}_j)} \sim \mathcal{N}(0,1)$, if:
    1. $\sqrt{n}(\hat{\beta}_j - \beta_j)$ has asymptotic variance $\frac{\sigma^2}{a_j^2}$, with $a_j^2 = \text{plim} \frac{\sum \hat{r}_{ij}}{n}$
        - Gauss-Markov assumptions must hold
        - Caution: Under heteroskedasticity, test statistics are invalid! $\rightarrow$ No matter the sample size, asympotitcs do not help
    2. $\hat{\sigma}^2$ is a consistent estimator of σ^2
- Implication: We can use critical values of $\mathcal{N}(0,1)$, if n is large
    - Even if we estimate σ^2 with $\hat{\sigma}^2$ or if y violates MLR.6
- Keep in mind that larger n does not change u in any way. It is a population parameter that is not affected by our sample!

Exercise 1a and b

To obtain Equation 2 from Equation 1, what steps do you need?

$$\sum_i (x_i - \bar{x})(y_i - \bar{y}) = \sum_i (x_i - \bar{x}) y_i - \sum_i (x_i - \bar{x}) \bar{y}$$
$$= \sum_i (x_i - \bar{x}) y_i - \bar{y} \sum_i (x_i - \bar{x})$$

$\rightarrow$ Multiply out the $(y_i - \bar{y})$
$\rightarrow$ Since $\bar{y}$ is a constant, we can put it before the sum sign.

We can obtain Equation 6 from Equation 5 since

- $\sum_i (x_i - \bar{x}) = 0$
- *Do not believe this? Try to calculate* $\sum_i (x_i - \bar{x}) = 0$ *for* $x_i = [3, 16, 7, 0]$ *or any other pair of numbers! :-)*

Exercise 1c and d

What does *plim* stand for?

- probability limit
- See Slide 4

What does $\text{plim}(\hat{\beta}_1) = \beta_1$ **substantively mean?**

- If we have an infinitely large sample, $\hat{\beta}$ converges to β.

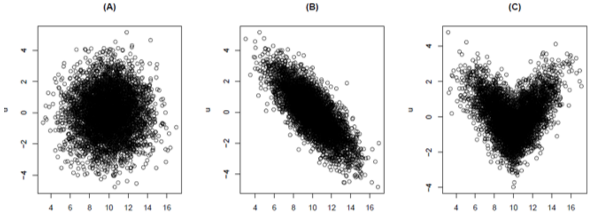
$\rightarrow$ Consistency!

Exercise 1e and f

What law do you need to obtain the statement in Eq. 8 from 7?

- Law of large numbers
- "As a sample size grows, its mean gets closer to the average of the whole population"

In the following figure, which panel clearly violates $Cov(x, u) = 0$?



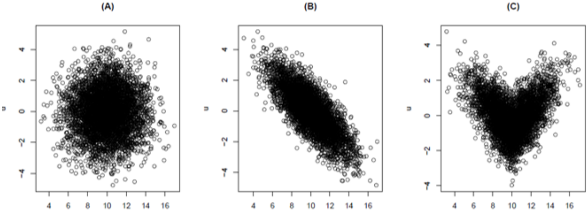
- A and C do not show a clear trend of increasing/decreasing $Cov(x, u)$ (altough C has a quite interesting Cov distribution).
- B on the other hand clearly shows a decreasing trend!

Exercise 1g and h

Which assumption of the linear regression leads to $Cov(x, u) = 0$?

- Zero conditional mean ($E(u|x) = 0$)
- Given all independent variables, u is 0 on average

In the above figure, which panel fulfils the assumption in the last task at most?



- A $\rightarrow$ The expected value of u seems to be independent of x and approximately zero

Exercise 1i and j

Can we derive the assumption in the last task from $Cov(x, u) = 0$?

- No
- $Cov(x, u) = 0$ does not automatically imply $E(u|x) = 0$

What property of the OLS estimator does Equation 8 demonstrate?

- Consistency
- $\hat{\beta}$ approaches β as n increases

Exercise 1k - 1m

Is the property of the last task a finite sample property?

- No
- Keep $\operatorname*{plim}_{n\to\infty} \hat{\beta}_1 = \beta_1$ in mind here!

What property of the OLS estimator does Equation 9 demonstrate?

- Unbiasedness

Is the property of the last task a finite sample property?

- Yes
- Keep in mind that
  - Inconsistency is normally expressed in terms of (approximate) population parameters
  - Bias is based on (finite) sample parameters

Exercise 2a

You have a classical linear model. What distribution does $\hat{\beta}$ have if you know the exact variance of errors. If the distribution has degrees of freedom, you also have to state them.

- A normal distribution (not necessarily the standard normal distribution)
  - $\hat{\beta}$ is just a linear combination of the errors in the sample. An important fact about independent normal random variables is that a linear combination of such random variables is normally distributed.

Exercise 2b

You have a classical linear model. What distribution does $\hat{\beta}$ have if you estimate the variance of errors by using the estimated residuals. If the distribution has degrees of freedom, you also have to state them.

- t distribution with n-k-1 degrees of freedom
  - The t distribution in comes from the fact that the constant σ in $sd(\hat{\beta})$ has been replaced with the random variable $\hat{\sigma}$.

Exercise 2c

Your regression analysis satisfies the Gauss-Markov-assumptions, but it is not a classical linear model. Which condition/s is/are sufficient for $\hat{\beta}$ to have a normal distribution?

- C: Your sample is large enough
  - Even though the y_i are not from a normal distribution, we can use the central limit theorem from to conclude that the OLS estimators satisfy asymptotic normality, which means they are approximately normally distributed in large enough sample sizes.

Exercise 2d

What theorem assures that the condition/s above is/are sufficient for $\hat{\beta}$ to have a normal distribution?

- Asymptotic normality
    - This theorem is useful because the normality assumption MLR.6 has been dropped; the only restriction on the distribution of the error is that it has finite variance, something we will always assume. We have also assumed zero conditional mean (MLR.4) and homoskedasticity of u (MLR.5).

Exercise 2e and f

You have a classical linear model. Does the OLS estimators have the smallest asymptotic variances? Answer with yes or no.

- Yes
  - Under the Gauss-Markov assumptions, the OLS estimators have the smallest asymptotic variances and therefore also when the normality assumption holds.

Your regression analysis satisfies the Gauss-Markov-assumptions, but it is not a classical linear model. Does the OLS estimators have the smallest asymptotic variances? Answer with yes or no.

- Yes
  - Under the Gauss-Markov assumptions, the OLS estimators have the smallest asymptotic variances.

Exercise 2g

Your regression analysis does not satisfy the homoskedasticity assumption. Does the OLS estimators have the smallest asymptotic variances? Answer with yes or no.

- No
  - If the homoskedasticity assumption fails, then there are estimators that have a smaller asymptotic variance than OLS.