Postscript Problem Set 1

Lars A. Lochstoer
UCLA Anderson School of Management

Winter 2018

Postscript to Problem Set 1

1. Stationarity and nonstationarity: Practical implications
2. Intuition for effect of White-correction on standard errors
   - In homework, it mattered a lot!
3. Asymptotic standard errors
   - When does the central limit theorem 'kick in'?

Stationarity: review

For a stationary series

- The unconditional mean and variance exists (are finite numbers)
- The autocovariances also exist (e.g., *cov* $(x_t, x_{t+1}) = \gamma_1$)
  - These moments are *not* a function of time

Practical implications:

1. A set of sample moments, that we base our model calibration (estimation) on, are "close" to population moments and therefore representative also of future data
2. OLS regressions (and most other methods we will see) yield biased estimates of regression coefficients and standard errors if data is nonstationary
   - Easy solution: transform nonstationary series so that it is stationary

Log differencing or normalizing

Take a price series, $P_t, \quad t = 1, ..., T$

- Typically nonstationary
  - Calculate sample autocorrelation $(corr\,(P_t, P_{t+1}))$
    - If higher than about $\frac{T-10}{T}$, you may be in trouble...

Simple solution: take logs, take the first difference

$$\begin{aligned} p_t &= \ln P_t \\ \Delta p_t &= p_t - p_{t-1} \end{aligned}$$

- Like a log return
  - Exact if no payouts (dividends)

Alternatively, work with a *ratio* that is not too highly autocorrelated

- E.g., market value divided by book value (a kind of *normalization*)

OLS and Stationarity

Consider univariate OLS:

$$y = x\beta + \varepsilon$$

Unconditional mean:

$$E[y] = E[x]\beta$$

Unconditional variance:

$$var(y) = \beta^2 var(x) + var(\varepsilon)$$

- Need both y and x variables to be stationary

Some regression do's and don'ts

Do not:

1. Regress a nonstationary variable on another nonstationary variable
   - Tools exist for working with nonstationary time series, but OLS ain't it
2. Regress a stationary variable on a nonstationary variable
   - Think about your null and alternative hypotheses in this case
3. Regress a nonstationary variable on a stationary variable
   - Think about your null and alternative hypotheses in this case
4. Regress one highly persistent variable on another highly persistent variable
   - In "small" samples, this is similar to 1. above

Do:

1. Regress a stationary dependent variable on stationary independent variables
2. Use OLS even if data is non-normal, but adjust standard errors appropriately and make sure sample is sufficiently long
   - Perhaps use Monte-Carlo analysis to check small sample behavior of test-statistic

Intuition for effect of White-correction on standard errors

Problem Set 1 – Asymptotic Standard Errors

With simple returns, the assumption of normally distributed error terms become even worse than for log returns

- Recall: limited liability
- In addition, there are extreme events and time-varying volatility...

White standard errors and a reasonably large sample allows for correct inference about standard errors given these real-world complexities

- E.g., 10 times more observations than parameters to be estimated and at least 30 degrees of freedom
- This number depends on exact distributional issues, so ultimately application-dependent

OLS and White Standard Error Implementation

Linear regression (in vector/matrix form as discussed in Lecture 2):

$$Y = X\beta + \varepsilon$$

- OLS large sample covariance matrix of estimated betas:

$$var\left(\hat{\beta}\right)_{OLS} = \left(X'X\right)^{-1}\left(X'X\right)s_{\varepsilon}^{2}\left(X'X\right)^{-1} = \left(X'X\right)^{-1}s_{\varepsilon}^{2}$$

  where s_{ε}^{2} is the sample variance of the residuals.

- OLS large sample covariance matrix of estimated betas:

$$var\left(\hat{\beta}\right)_{White} = \left(X'X\right)^{-1}\Sigma_{t=1}^{T}x_t x_t'\varepsilon_t^2\left(X'X\right)^{-1}$$

- Under what circumstances are White standard errors higher?
  - In PS1, OLS s.e. was 0.01 vs. White s.e. at 0.018

Covariance definition

Recall:

$$\begin{aligned} E\left[(x - E[x])(y - E[y])\right] &= cov(x, y) \\ &= E[xy] - E[x]E[y] \end{aligned}$$

Recall OLS assumption:

$$E\left[xx'\varepsilon^2\right] = E\left[xx'\right]E\left[\varepsilon^2\right]$$

White relaxes this:

$$E\left[xx'\varepsilon^2\right] = E\left[xx'\right]E\left[\varepsilon^2\right] + cov\left(xx', \varepsilon^2\right)$$

Thus: Higher White standard errors (relative to standard OLS) means the squared residuals are positively correlated with squared explanatory variables

- In PS1: When market return large (in absolute value), residuals also large
- Intuition, large values of x are very informative for β. *Unless*, of course, this is when the noise term is large as well...

Why you should care about correct standard errors

- Coefficient standard errors important for two things:
  1. Statistical inference
  2. Prediction errors

- (1) includes testing null hypotheses regarding coefficient restrictions. E.g., can we get by with a simpler more restrictive model?
  - More parsimonious models typically do better out of sample

- (2) is important for the degree of confidence you have in your model's prediction
  - Perhaps you should do nothing if prediction error is too large?
  - Perhaps hedging when the prediction error is large *increases* risk out of sample?
  - How much should you shrink your model prediction to unconditional prediction?
    - Example, if large standard error on a firm's market beta, perhaps better estimate in terms of future mean-squared-error is to shrink the estimate close to 1 (the market average)

Asymptotics: When does the CLT kick in?

Consider a slightly more extreme version of the model from PS1

- mu = 0.012; sig = 0.05; p = 0.15; mu_j = -0.05; sig_j = 0.17;
- $E[r] = 0.45\%$, $\sigma(r) = 8.46\%$, $skewness(r) = -0.92$, $kurtosis(r) = 9.93$
- So, like monthly log stock market return data (ref lecture1 notes)

Monte-Carlo excercise:

1. Simulate N return samples of length T
2. Calculate the sample mean for each of the N samples
3. Calculate the t-statistic for the sample mean in each sample as $\hat{\mu}\sqrt{T}/\sigma(r)$
   - Asymptotics preduct this quantity has a Normal distribution with unit variance
4. Give the first four moments of this statistic across the N samples
5. Do this for $T = 12, 60, 600$

Monte-Carlo simulation results

Denote the sample mean t-stat for each samples t_n.

- Sample mean is positive, so t-stat should on average be positive against null of zero
- N = 50,000

T = 12 (1 year)

- $E_N\left[t_n\right] = 0.37$, $\sigma_N\left(t_n\right) = 1.19$, $skewness_N\left(t_n\right) = 0.37$, $kurtosis_N\left(t_n\right) = 3.44$

T = 60 (5 years)

- $E_N\left[t_n\right] = 0.51$, $\sigma_N\left(t_n\right) = 1.07$, $skewness_N\left(t_n\right) = 0.25$, $kurtosis_N\left(t_n\right) = 3.04$

T = 600 (50 years)

- $E_N\left[t_n\right] = 1.33$, $\sigma_N\left(t_n\right) = 1.03$, $skewness_N\left(t_n\right) = 0.10$, $kurtosis_N\left(t_n\right) = 3.00$

CLT kicks in pretty quickly in this case (5 years of data seems fine), but it is a simple case (sample mean and truly iid data)