Econometrics 871
Time Series

TOPIC 4: TUTORIAL

Replicating the Dickey-Fuller distribution

Review of asymptotic results and hypothesis testing

- Consider the simplest linear regression on a sample of n observations:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \qquad with \quad \varepsilon_i \sim (0, \sigma^2)$$

- OLS estimatpr

$$\hat{\beta}_{1,OLS} = \frac{cov(y,x)}{var(x)}$$

- If the standard OLS assumptions hold then OLS is consistent:

$$\lim_{n\to\infty} \hat{\beta}_{1,OLS} = \beta_1$$

- In a finite sample, $\hat{\beta}_{1,OLS}$ is a random variable
  - If ε_i is i.i.d. normal, then $\hat{\beta}_{1,OLS}$ is also normal with $var(\hat{\beta}_{1,OLS}) = \frac{\sigma^2}{n\, var(x)}$
  - Even if ε_i is i.i.d. but not normal, the central limit theorem proves:

$$\lim_{n\to\infty} \sqrt{n}(\hat{\beta}_{1,OLS} - \beta_1) \sim N\left(0, \frac{\sigma^2}{n\, var(x)}\right)$$

Review of asymptotic results and hypothesis testing

- the central limit theorem proves:

$$\lim_{n\to\infty}\sqrt{n}\left(\hat{\beta}_{1,OLS}-\beta_1\right)\sim N\left(0,\frac{\sigma^2}{n\,var(x)}\right)$$

- Thus, we might use the small sample **approximation** for hypothesis tests:

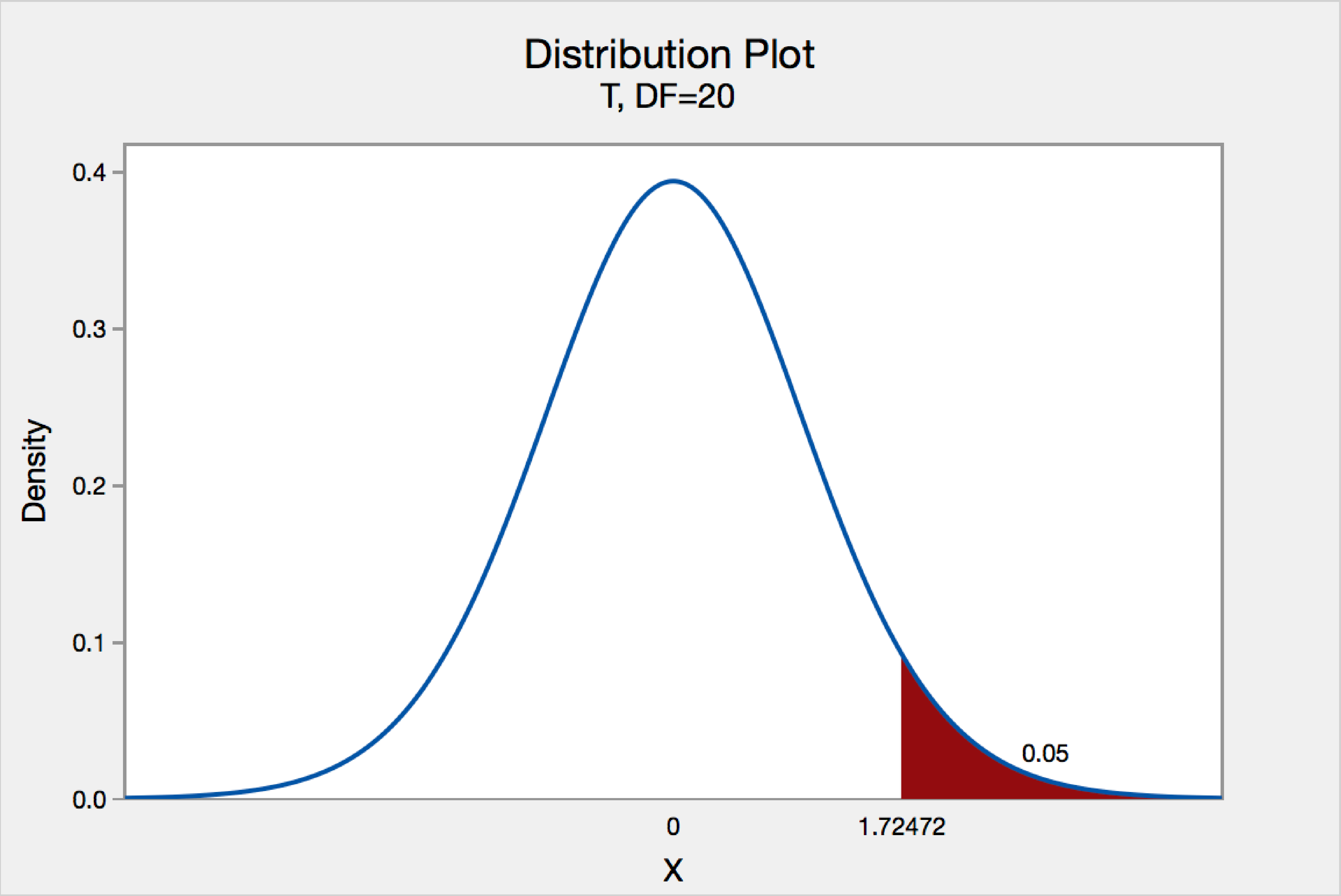
$$\frac{\left(\hat{\beta}_{1,OLS}-\beta_1\right)}{\sqrt{var\left(\hat{\beta}_{1,OLS}\right)}}\sim N(0,1)$$

- However, we do not know σ^2 - it must be estimated with

$$\mathrm{s}^2=\frac{1}{\mathrm{n}}\sum\varepsilon_i^2$$

- Then we use the standard t-distribution

$$\frac{\left(\hat{\beta}_{1,OLS}-\beta_1\right)}{\sqrt{\frac{s^2}{n\,var(x)}}}\sim t(n-k)$$



Distribution Plot
T, DF=20
Density
0.4
0.3
0.2
0.1
0.0
0.05
0
1.72472
X

Fundamental Setting

Given an unknown AR(1) process:

$$y_t = a_1 y_{t-1} + \varepsilon_t$$

- If $|a_1| < 1$
  - The process is stationary
  - An OLS regression of y_t on y_{t-1} yields a consistent (but biased) estimate of a_1
  - Let the sample be of size T
  - Biased means: $E(\hat{a}_1) \neq a_1$
  - Consistent means: $\lim_{T\to\infty} \hat{a}_1 = a_1$

- If $a_1 = 1$
  - The process is non-stationary
  - An OLS regression of y_t on y_{t-1} yields an inconsistent estimate of a_1: $\lim_{T\to\infty} \hat{a}_1 \neq a_1$
  - In this setting: $\lim_{T\to\infty} \hat{a}_1 < a_1$

Test equation:

- Subtracting y_{t-1} from both sides yields the test equation:

$$\Delta y_t = (a_1 - 1)y_{t-1} + \varepsilon_t$$

$$= \gamma y_{t-1} + \varepsilon_t$$

  – If $|a_1| < 1 \Leftrightarrow \gamma < 0$,
    - y_t is stationary, thus so is Δy_t
    - a regression of Δy_t on y_{t-1} yields a consistent estimate of γ, with standard distributional results (i.e. $\frac{\hat{\gamma}_{OLS} - \gamma}{s.e.(\hat{\gamma})}$ has an asymptotic t-distribution centred at zero)
    - Consistency: $\lim_{T \to \infty} \hat{\gamma}_{OLS} = \gamma$
    - However, in a small sample $\hat{\gamma}_{OLS}$ will be biased because y_{t-1} is not exogenous with respect to ε_t: I.e. the condition $\mathrm{E}(y_t \varepsilon_s) = 0 \forall t, s$ does not hold

Test equation:

- Subtracting y_{t-1} from both sides yields the test equation:

$$\Delta y_t = (a_1 - 1)y_{t-1} + \varepsilon_t$$

$$= \gamma y_{t-1} + \varepsilon_t$$

- If $a_1 = 1 \Leftrightarrow \gamma = 0$
  - the I(1) term, y_{t-1}, falls out of the regression at the null of a unit root, so the regression is valid, but $\hat{\gamma}_{OLS}$ has a non-standard distribution
  - We will show that the *mode* of the distribution of $\hat{\gamma}_{OLS}$ is equal to γ, but the mean and median are not, so $\lim\limits_{T\to\infty} \hat{\gamma}_{OLS} \neq \gamma$
  - Moreover, the distribution is non-standard ($\frac{\hat{\gamma}_{OLS}-\gamma}{s.e.(\hat{\gamma})}$ does not have a t-distribution)
  - Thus the critical values of a hypothesis test are different from those of a t-distribution at the null hypothesis of a unit root (i.e. $\mathrm{H}_0: \gamma = 0$)

Exercise for the day:

- Construct a Monte Carlo simulation that reconstructs the Dickey Fuller distribution and critical values for the t-test of a null of a unit root
- We will do a general simulation, for any value of γ (unit root and no unit root)
- We will show that:
  - If $\gamma < 0$, $\hat{\gamma}_{OLS}$ is on average correct/consistent and the distribution of the test of H_0: $\hat{\gamma}_{OLS} = \gamma$ has an approximate t-distribution *only if* the sample of observations T is large enough
    - This raises a subtle point not often discussed: for near-unit root processes, small sample test statistics can be misleading
  - If $\gamma = 0$, $\hat{\gamma}_{OLS}$ is on average incorrect/inconsistent and the distribution of the test of H_0: $\hat{\gamma}_{OLS} = 0$ does not have a t-distribution *no matter* how large the sample of observations is

Monte Carlo Simulation

- For a process defined by a given AR coefficient a_1:

$$y_t = a_1 y_{t-1} + \varepsilon_t$$

  - Generate N different time-paths of length T
  - For each time-path $i \in N$,
    - do the OLS regression of the test equation:

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t$$

    - Store $\hat{\gamma}_{OLS}$ and $\frac{\hat{\gamma}_{OLS}-\gamma}{s.e.(\hat{\gamma})}$
  - Approximate the density function of $\hat{\gamma}_{OLS}$ and $\frac{\hat{\gamma}_{OLS}-\gamma}{s.e.(\hat{\gamma})}$
  - Compare the density function of $\frac{\hat{\gamma}_{OLS}-\gamma}{s.e.(\hat{\gamma})}$ to that of a standard t-distribution
  - Compute the empirical critical t-statistic and compare to the theoretical t-statistic of an α significance level

- Study the impact of varying a_1 and T