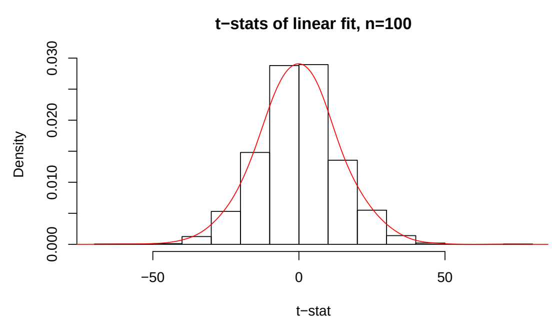
Testing for Unit Roots

Overview

1. Ideas.
2. Estimators.
3. Essential asymptotic properties.

Unit roots: two questions

First question Is that a random walk or a trend? Simulation results show that the distribution of the usual t-statistic is very fat tailed: *i.e.*, we'll often reject $H_0 : \beta_1 = 0$ for a linear model when there's a unit root. This histogram shows the distribution of t-statistics for testing the slope in a linear model $\mathbb{E}\, X_t = \beta_0 + \beta_1\, t$ in which X_t is in fact a Gaussian random walk.



The shape is Gaussian, but the scale is way off. Almost 90% of the t-stats are larger than 2 in absolute value. This situation is a special case of what's known as *spurious regression*.

Second question Is $|\phi| < 1$ in an AR(1) model?

Stick with the simple case of a AR(1) model with mean zero,

$$X_t = \phi X_{t-1} + w_t, \quad w_t \sim WN(0, \sigma^2) \tag{1}$$

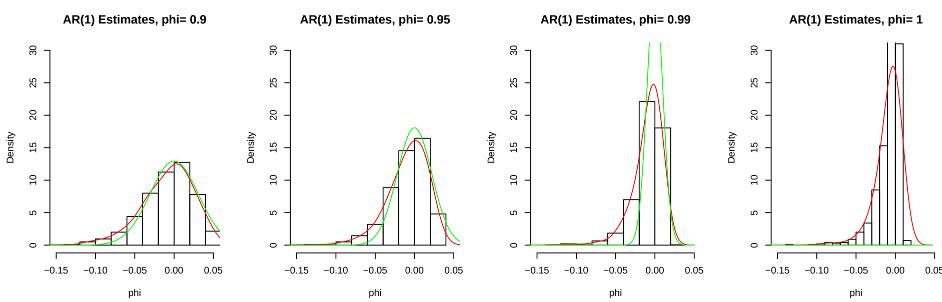
Variations on these results apply when there's a constant term or a time trend.

Problem arises in this case as ϕ approaches 1. The distribution of the OLS estimator of ϕ (let $X_t = 0$ if $t < 1$)

$$\hat{\phi} = \frac{\sum_1^n X_t X_{t-1}}{\sum_1^n X_{t-1}^2} \tag{2}$$

is "non-standard" when $\phi = 1$. Rather than approach the asymptotic normal distribution, the distribution of the sampling estiamates resembles this other distribution for ϕ near 1. (Note: YW estimator or the estimator obtained from the periodogram would not be suitable for this problem since $\hat{\phi} < 1$ by construction.)

Simulation results These histograms compare the distribution of $\hat{\phi}$ for $\phi = 0.95, 0.99, 1.0$ with $n = 100$ and gaussian white noise.



Dickey Fuller tests

Literature goes back to work of Dickey and Fuller (1979 *JASA*, 1981 *Econometrica*) and classical methods that show that

$$\sqrt{n}(\hat{\phi} - \phi) \sim N(0, 1 - \phi^2), \quad |\phi| < 1$$

whereas $n(\hat{\phi} - 1)$ has an asymptotic distribution if $|\phi| = 1$ (*i.e.*, the standard error of $\hat{\phi}$ is proportional to $1/n$ rather than $1/\sqrt{n}$.

Model comes in various flavors, but the most interesting combines possible non-stationarity with a time trend. Consider the model

$$X_t = \beta_0 + \beta_1(t - \bar{t}) + \phi X_{t-1} + W_t \tag{3}$$

We'd like to test things such as whether $\beta_1 = 0$ when $\phi = 1$

Dickey-Fuller test Authors contribution was to find the asymptotic distribution for $n(\hat{\phi} - 1)$ when $\phi = 1$ and table it. They also propose that one test $H_0 : \phi = 1$ rather than assume the process is stationary. They show that the likelihood ratio tests of the various parameters are functions of the "usual" t-statistic (or F-tests). Their tests work as follows. Rather than fit the regression specified in (3), subtract X_{t-1} from both sides of the quation and fit

$$(1 - B)X_t = \beta_0 + \beta_1(t - \bar{t}) + (\phi - 1)X_{t-1} + W_t$$

Estimate $\phi - 1$ via OLS and test whether this estimate is different from 0. You test this using the standard t-statistics, but don't find the p-value from the t-distribution. Instead, you have to use their tables of the sampling distributions of the test statistics.)

Augmented Dickey-Fuller test The augmented test allows dependent (but stationary) noise in (3). The model is

$$X_t = \beta_0 + \beta_1(t - \bar{t}) + \phi X_{t-1} + Z_t, \quad Z_t = \sum_{j=1}^{p} \rho_j Z_{t-j} + w_t$$

Under H_0 that the process is a simple random walk rather than a time trend plus stationary noise ($\beta_0 = \beta_1 = 0, \phi = 1$), $Z_t = X_t - X_{t-1} = \Delta X_t$), this expression is equivalent to the following:

$$X_t = X_{t-1} + \sum_{j=1}^{p} \rho_j \Delta X_{t-j} + w_t \tag{4}$$

As in the regular Dickey-Fuller test, you estimate the coefficient of X_{t-1} and $t - \bar{t}$ in the equation (subtract X_{t-1} from both sides of (4) with coefficient ϕ)

$$(1 - B)X_t = \beta_0 + \beta_1(t - \bar{t}) + (\phi - 1)X_{t-1} + \sum_j \rho_j \Delta X_{t-j} + w_t$$

by least squares, then compare the "usual" t-statistics of the estimates to a tabled distribution.

Issues Do these tests have much power? If the tests have low power, you're left with a non-stationary process.

What's the right number of lags p to use in (4)? Should you pick these, for example, by using AIC? (Some answers include the tests of Newey and West that allow a more nonparametric estimate, needing only the variance of the errors.)

Examples

Null models used in testing are

$$\Delta X_t = \beta_0 + \beta_1(t - \bar{t}) + (\phi - 1)X_{t-1} + \sum_j \rho_j \Delta X_{t-j} + w_t \quad (5)$$

$$\Delta X_t = \beta_0 + (\phi - 1)X_{t-1} + \sum_j \rho_j \Delta X_{t-j} + w_t \quad (6)$$

$$\Delta X_t = (\phi - 1)X_{t-1} + \sum_j \rho_j \Delta X_{t-j} + w_t \quad (7)$$

Software package urca (The author B Pfaff has a book on the topic, *Analysis of Integrated and Co-integrated Times Series in R*. The package includes several tests that one does sequentially.

```
ur.df(y, type=c(``none'', ``drift'', ``trend''), lags=1,
      selectlags=c(``Fixed'', ``AIC'', ``BIC''))
```

Set `type` to

- "trend" to test $H_0 : \beta_0 = \beta_1 = \phi - 1 = 0$ in (5), to
- "drift" to test the $H_0 : \beta_0 = \phi - 1 = 0$ in (6), and to
- "none" to test $H_0 : \phi = 1$ in the pure random walk (7).

The tests are typically done in this order. If you reject, for example, the first test, then it appears that you have stationary data.

The option `lags` sets p in these equations; you can set p manually (`selectlags` = "Fixed") or use an automatic criterion to choose these for you.

Examples are included in the file `19.R`.