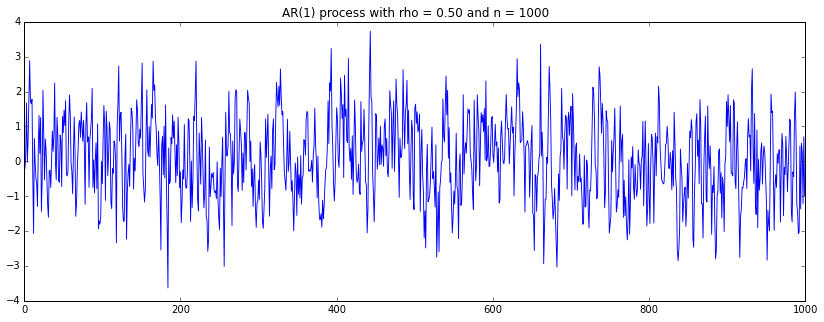
## Problem 1

In this problem I generate three AR(1) models with and estimate using linear regression. In each case the 95% confidence interval included the true DGP value of . The plots and regression outputs are presented below.

#### 1.a.



**OLS Regression Results with rho = 0.5**

==============================================================================

Dep. Variable: y R-squared: 0.261

Model: OLS Adj. R-squared: 0.260

Method: Least Squares F-statistic: 352.5

Date: Mon, 03 Nov 2014 Prob (F-statistic): 1.43e-67

Time: 19:47:55 Log-Likelihood: -1424.5

No. Observations: 999 AIC: 2853.

Df Residuals: 997 BIC: 2863.

Df Model: 1

==============================================================================

coef std err t P>|t| [95.0% Conf. Int.]

------------------------------------------------------------------------------

const 0.0336 0.032 1.053 0.293 -0.029 0.096

x1 0.5109 0.027 18.774 0.000 0.457 0.564

==============================================================================

Omnibus: 0.342 Durbin-Watson: 1.979

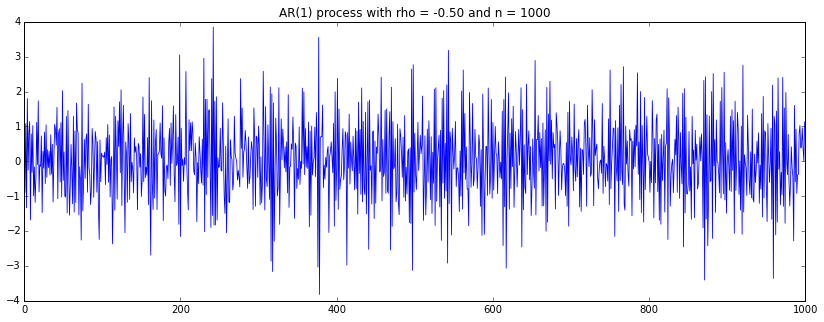
Prob(Omnibus): 0.843 Jarque-Bera (JB): 0.408

Skew: -0.038 Prob(JB): 0.815

Kurtosis: 2.937 Cond. No. 1.19

==============================================================================

#### 1.b.



**OLS Regression Results with rho = -0.5**

==============================================================================

Dep. Variable: y R-squared: 0.284

Model: OLS Adj. R-squared: 0.283

Method: Least Squares F-statistic: 394.9

Date: Mon, 03 Nov 2014 Prob (F-statistic): 2.77e-74

Time: 19:53:19 Log-Likelihood: -1431.2

No. Observations: 999 AIC: 2866.

Df Residuals: 997 BIC: 2876.

Df Model: 1

==============================================================================

coef std err t P>|t| [95.0% Conf. Int.]

------------------------------------------------------------------------------

const 0.0229 0.032 0.713 0.476 -0.040 0.086

x1 -0.5326 0.027 -19.871 0.000 -0.585 -0.480

==============================================================================

Omnibus: 1.516 Durbin-Watson: 1.954

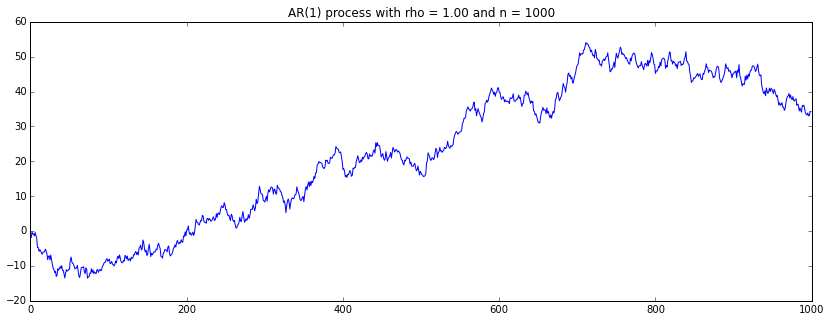
Prob(Omnibus): 0.469 Jarque-Bera (JB): 1.592

Skew: 0.079 Prob(JB): 0.451

Kurtosis: 2.886 Cond. No. 1.20

==============================================================================

#### 1.c. (Random Walk)



**OLS Regression Results with rho = 1**

==============================================================================

Dep. Variable: y R-squared: 0.989

Model: OLS Adj. R-squared: 0.989

Method: Least Squares F-statistic: 9.106e+04

Date: Mon, 03 Nov 2014 Prob (F-statistic): 0.00

Time: 19:53:19 Log-Likelihood: -1401.6

No. Observations: 999 AIC: 2807.

Df Residuals: 997 BIC: 2817.

Df Model: 1

==============================================================================

coef std err t P>|t| [95.0% Conf. Int.]

------------------------------------------------------------------------------

const -0.0618 0.043 -1.426 0.154 -0.147 0.023

x1 0.9941 0.003 301.761 0.000 0.988 1.001

==============================================================================

Omnibus: 1.123 Durbin-Watson: 1.968

Prob(Omnibus): 0.570 Jarque-Bera (JB): 1.202

Skew: 0.071 Prob(JB): 0.548

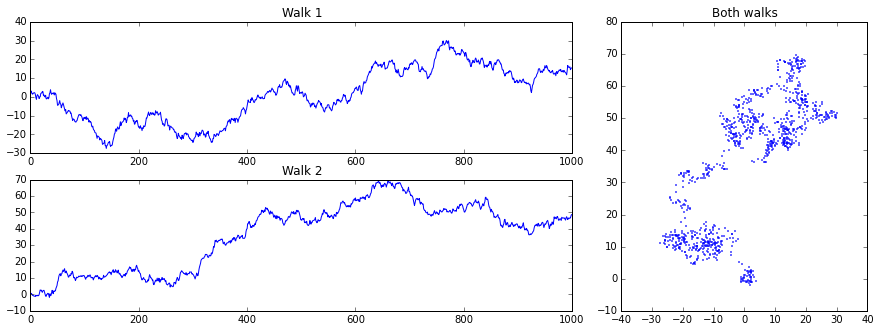
Kurtosis: 2.906 Cond. No. 18.4

==============================================================================

## Problem 2.

The plots and regression results for two independently generated random walks.

#### 2.a.

****

**OLS Regression Results of Walk1 ~ Walk2**

==============================================================================

Dep. Variable: y R-squared: 0.544

Model: OLS Adj. R-squared: 0.543

Method: Least Squares F-statistic: 1190.

Date: Mon, 03 Nov 2014 Prob (F-statistic): 2.52e-172

Time: 20:10:14 Log-Likelihood: -3688.6

No. Observations: 1000 AIC: 7381.

Df Residuals: 998 BIC: 7391.

Df Model: 1

==============================================================================

coef std err t P>|t| [95.0% Conf. Int.]

------------------------------------------------------------------------------

const -18.1120 0.632 -28.642 0.000 -19.353 -16.871

x1 0.5278 0.015 34.499 0.000 0.498 0.558

==============================================================================

Omnibus: 53.852 Durbin-Watson: 0.013

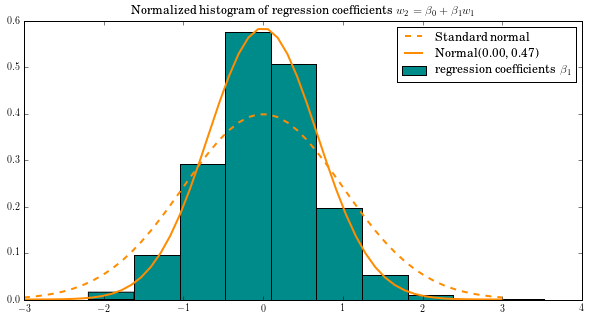
Prob(Omnibus): 0.000 Jarque-Bera (JB): 28.245

Skew: 0.237 Prob(JB): 7.36e-07

Kurtosis: 2.326 Cond. No. 85.4

#### 2.b.

The normalized empirical distribution of regression coefficients is pictured by the histogram below with the two normal curves added for comparison: and added for comparison.



#### 2.b.c.

In assignment 4, question 3c, almost all of the distribution was between -0.1 and 0.1, so the histogram was much tighter. This presents strong evidence that although slopes that are not very close to zero are extremely rare if OLS is fitted to two samples of 1000 points taken from independent normal populations, large slopes are relatively common when OLS is fitted to two independent random walks. More concisely, large independent samples are almost always uncorrelated, but large independent random walks often exhibit spurious correlation.

In the first plot of the two random walks I plotted the random walk on a plane as a scatter plot to make it look more like a collection of possibly correlated points than a path. Here’s another plot showing the plane walk as a path and using a more meaningful subplot arrangement. It would be fun to animate a point traveling from the green start to the red finish. (next page)

Macintosh HD:Users:petervarshavsky:Documents:Git_NYU:applied_data_science:assignment-5:random_walk.pdf