Stationarity / Nonstationarity

Natalia Khorunzhina

Copenhagen Business School

Stationarity

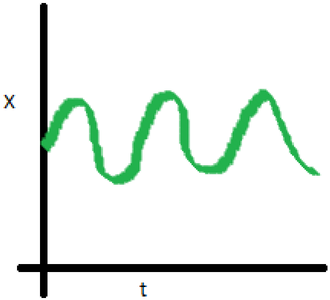
Using regression analysis, we quantified the past historical relationship and generalize it to the future. That is, we assumed that the future is like the past, that is, we assumed **stationarity**.

- In general, a time series is stationary if its probability distribution does not change over time.
- Thus, a time series Y_t is stationary if the joint distribution of $(Y_{s+1}, Y_{s+2}, \ldots, Y_{s+T})$ does not depend on s.
- Otherwise, Y_t is said to be non-stationary.
- X_t and Y_t are said to be jointly stationary if the joint distribution of $(X_{s+1}, Y_{s+1}, X_{s+2}, Y_{s+2}, \ldots, X_{s+T}, Y_{s+T})$ does not depend on s.
- Stationarity requires the future to be linked to the past (in a probabilistic sense)

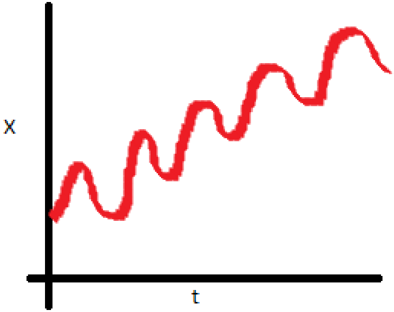
Covariance Stationary

- In this class, we'll require covariance stationarity
- Also called weak stationarity
- A series Y_t is cov-stationary if it has a constant mean, constant variance, and if the covariance only depends on the time difference, e.g. $\gamma_k = cov(y_t, y_{t-k})$ only depends on k and not t

Mean Stationary - constant mean

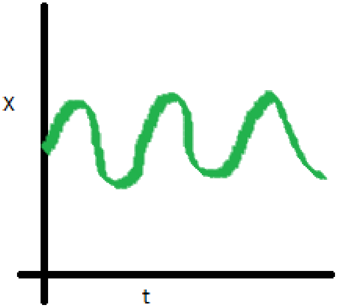


Stationary series

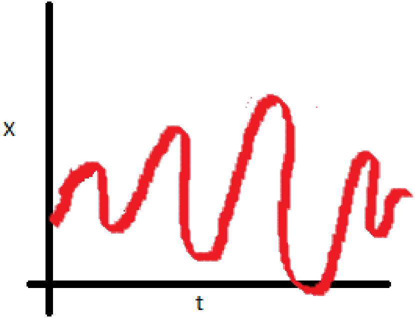


Non-Stationary series

Variance Stationary - constant variance

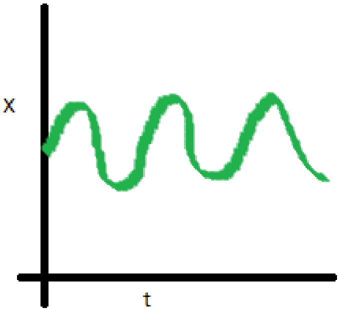


Stationary series

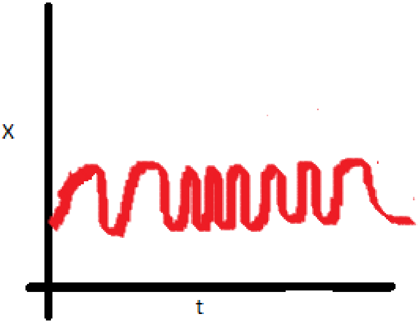


Non-Stationary series

Covariance Stationary - covariance does not depend on t

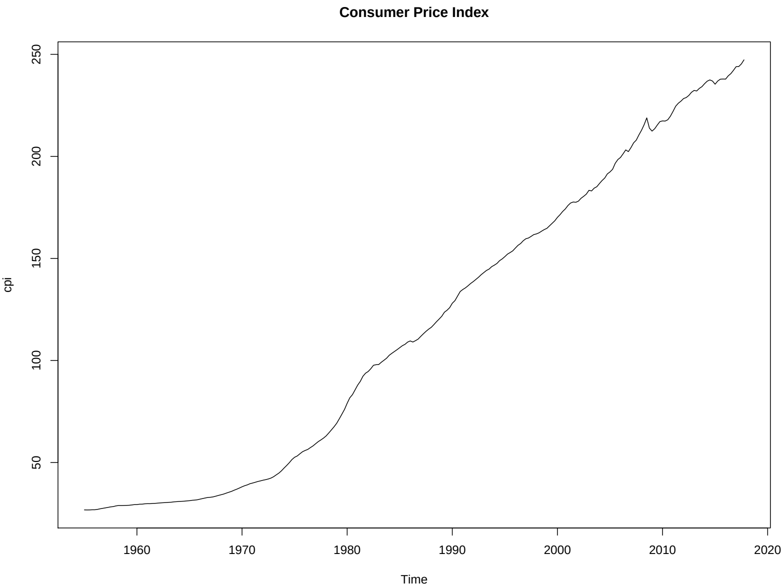


Stationary series

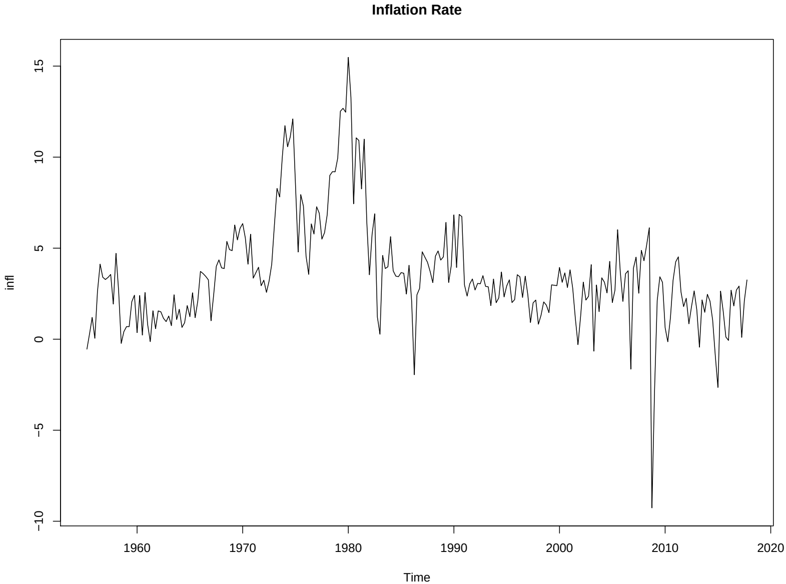


Non-Stationary series

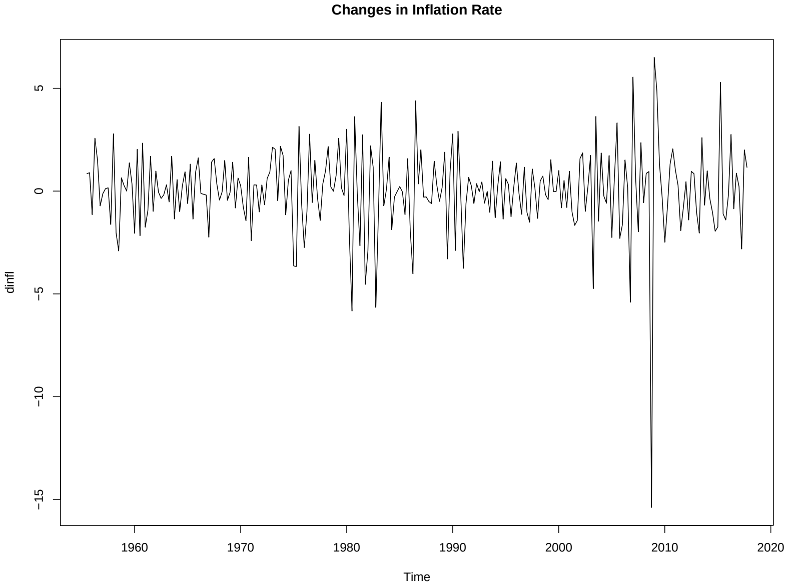
Example of nonstationarity – CPI



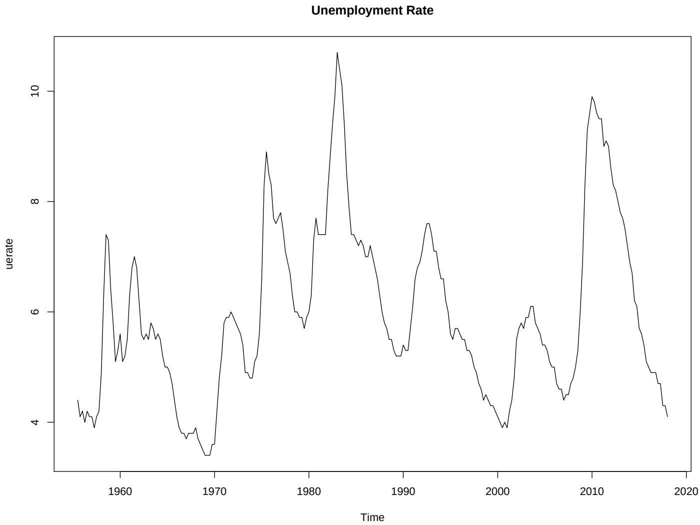
Looking more stationary - Inflation



Example of stationarity - Changes in Inflation



One more example of stationarity - Unemployment Rate



The lag operator

- Learning lag operators will help us prove stationarity for different models
- Let L be the lag operator and y_t be a time series
- $Ly_t = y_{t-1}$
- $L^2 y_t = L(y_{t-1}) = y_{t-2}$
- In general, $L^k y_t = y_{t-k}$
- The lag operator and multiplication are commutative: $L(ay_t) = aLy_t = ay_{t-1}$ where a is a constant
- The lag operator is distributive over addition: $L(y_t + x_t) = Ly_t + Lx_t = y_{t-1} + x_{t-1}$

The lag operator and AR models

- The AR(1) model: $Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t$
- $(1 - \beta_1 L) Y_t = \beta_0 + u_t$
- The AR(2) model: $Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + u_t$
- $(1 - \beta_1 L - \beta_2 L^2) Y_t = \beta_0 + u_t$
- The AR(p) model: $Y_t = \beta_0 + \beta_1 Y_{t-1} + \cdots + \beta_p Y_{t-p} + u_t$
- $(1 - \beta_1 L - \cdots - \beta_p L^p) Y_t = \beta_0 + u_t$
- Thus, the AR(p) model is Y_t multiplied by a polynomial of order p.

Stationarity of the AR(p) model

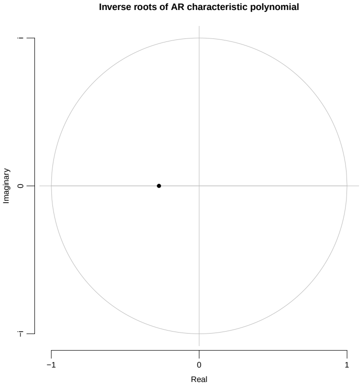
- The AR(p) is stationary if the roots of the "characteristic" polynomial, $(1 - \beta_1 z - \cdots - \beta_p z^p) = 0$, (replaced L with z) lie outside the unit circle

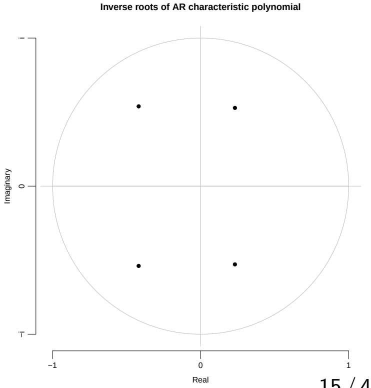
Example stationary AR(1)

- $Y_t = 0.2 + 0.7Y_{t-1} + u_t$
- $(1 - 0.7L)Y_t = 0.2 + u_t$
- This AR(1) is stationary if the roots of $(1 - 0.7z) = 0$ lie outside the unit circle
- Solve for z: $z = 10/7$. So, the roots of the model lie outside the unit circle. The model is stationary
- An easy way to tell if an AR(1) is stationary is to look at the coefficient on Y_{t-1}. If this coefficient is less than one in absolute value, the model is stationary.

Example: Characteristic roots for AR(1) and AR(4) inflation model

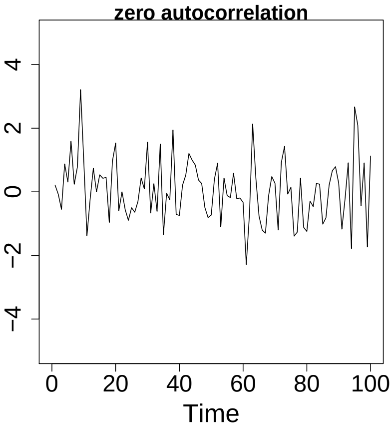
It is actually convenient to get <u>inverse</u> roots, these will be inside the unit circle if AR(p) model is stationary.

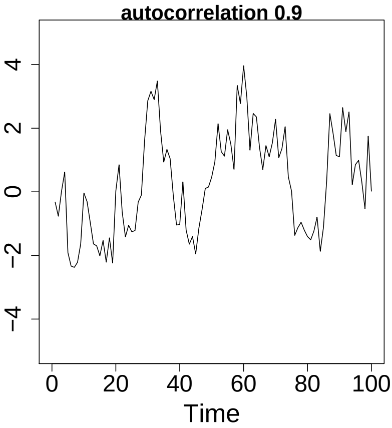




Simulate the stationary AR(1) for 100 observations

- Let u_t be distributed $N(0, 1)$. Let Y_0 (the starting value) be equal to zero



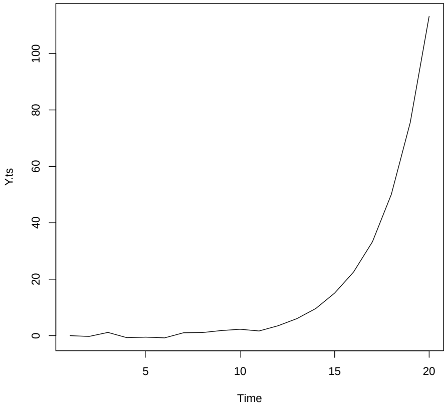


Example of nonstationary AR(1)

- $Y_t = 0.2 + 1.5Y_{t-1} + u_t$
- $(1 - 1.5L)Y_t = 0.2 + u_t$
- Check the roots of $(1 - 1.5z) = 0$.
- $z = 2/3$. The model is nonstationary.
- An easy way to tell if an AR(1) is nonstationary is to look at the coefficient on Y_{t-1}. If this coefficient is greater than one in absolute value, the model is nonstationary.

Simulate the nonstationary AR(1) for 20 observations

- Let u_t be distributed $N(0, 1)$. Let Y_0 (the starting value) be equal to zero



The stationary vs nonstationary AR(1)

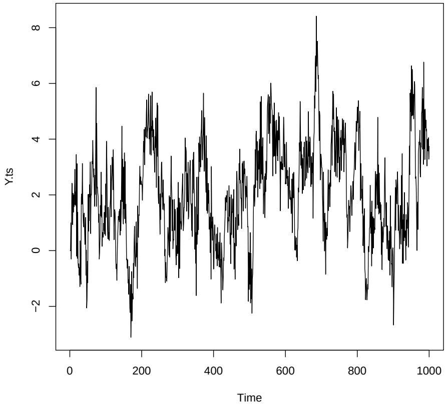
- Look at the examples above.
- Even after 100 observations, the stationary AR(1) revolves within a fixed interval
- The nonstationary AR(1) model, however, explodes even after just 20 observations

Example of Stationary AR(2)

- $Y_t = 0.2 + 0.4Y_{t-1} + 0.5Y_{t-2} + u_t$
- $(1 - 0.4L - 0.5L^2)Y_t = 0.2 + u_t$
- Check the roots of $(1 - 0.4z - 0.5z^2) = 0$
- $z = -1.87, 1.07$
- Thus, the model is stationary

Simulate the stationary AR(2)

- Let u_t be distributed $N(0, 1)$. Let Y_0 and Y_{-1} (the starting values) be equal to zero

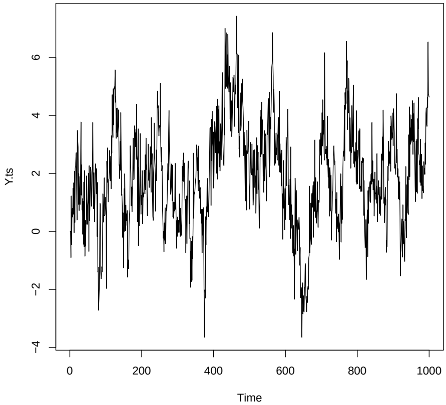


Example of nonstationary AR(2)

- $Y_t = 0.2 + 0.6Y_{t-1} + 0.7Y_{t-2} + u_t$
- $(1 - 0.6L - 0.7L^2)Y_t = 0.2 + u_t$
- Check the roots of $(1 - 0.6z - 0.7z^2) = 0$
- $z = -1.7, 0.84$
- One of the roots is inside the unit circle. The model is non-stationary

Simulate the nonstationary AR(2)

- Let u_t be distributed $N(0, 1)$. Let Y_0 and Y_{-1} (the starting values) be equal to zero



Stationary vs Nonstationary AR(2)

- Notice from the graphs above that it is sometimes difficult to tell if a model is stationary.
- Later we'll learn statistical tests for nonstationarity
- If an AR(p) model has a root that equals 1, the series is said to have a unit root

Formal Test for Unit Root

Start with the simple AR(1) process:

$$Y_t = \beta Y_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim WN(0, \sigma^2)$.

We can regress Y_t on Y_{t-1} and then use the standard t-test for testing $\beta = 1$.

But remember that t-statistic is for the null hypothesis that $\beta = 0$, while we want to test that $\beta = 1$.
H_0: $\beta = 1 \Rightarrow$ Unit root or stochastic trend
H_1: $\beta < 1 \Rightarrow$ No unit root, a stationary time series

Formal Test for Unit Root

We can trick the system by rewriting the regression as

$$\begin{aligned} Y_t &= \beta Y_{t-1} + \varepsilon_t \\ Y_t - Y_{t-1} &= \beta Y_{t-1} - Y_{t-1} + \varepsilon_t \\ Y_t - Y_{t-1} &= (\beta - 1) Y_{t-1} + \varepsilon_t \\ \Delta Y_t &= \delta Y_{t-1} + \varepsilon_t \end{aligned}$$

where $\delta = \beta - 1$, and testing the null hypothesis that $\delta = 0$.
H_0: $(\beta_1 - 1) = 0 \Rightarrow$ unit root
H_0: $\delta = 0 \Rightarrow$ unit root

The test statistic follows the Dickey-Fuller Distribution (not the normal t-distribution).

Testing for a Unit Root – ADF test

- Augmented Dickey-Fuller Test for a Unit root
- No trend:
  $\Delta Y_t = \beta_0 + \delta Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \cdots + \gamma_p \Delta Y_{t-p} + u_t$
- trend:
  $\Delta Y_t = \beta_0 + \alpha t + \delta Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \cdots + \gamma_p \Delta Y_{t-p} + u_t$
- Null hypothesis: $\delta = 0$, Y_t has a stochastic trend
- Can choose p using the AIC
- The test statistic follows the Dickey-Fuller Distribution (not the normal t-distribution)

Critical Values for the ADF statistic

Table: Dickey-Fuller Critical Values

Deterministic Regressors	10%	5%	1%
Intercept Only	-2.57	-2.86	-3.43
Intercept and time trend	-3.12	-3.41	-3.96

The ADF test in practice

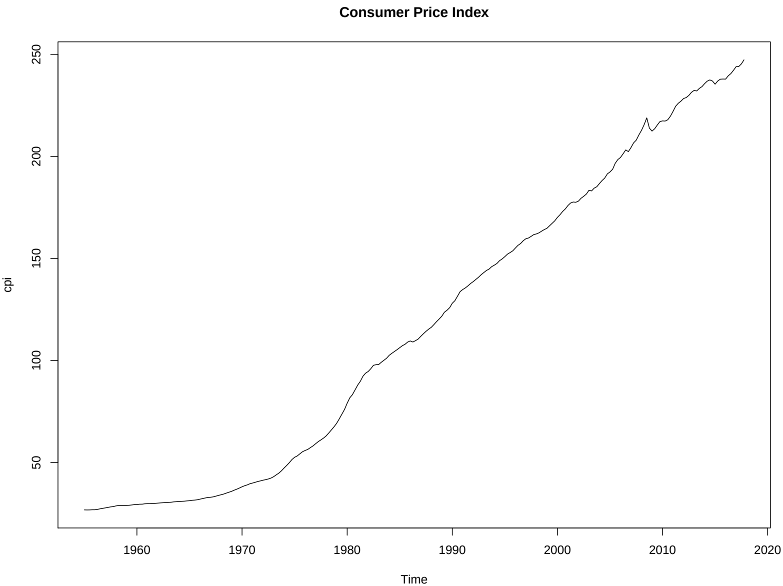
- I first test for a unit root in the presence of a trend.
- If the trend is insignificant, I drop it and re-test
- I always keep an intercept in my tests even if it's not significant so that the residuals sum to zero
- Recall that the null hypothesis is that the series has a unit root.
- We can remove a unit root by differencing
- Unit root tests have "low power." Intuitively, this means that the test will not always reject the null hypothesis even when it's false (we'll see how this can be a problem in a minute)

ADF Test (with trend)

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################
Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.1853873  0.0681252   2.721  0.00692 **
z.lag.1     -0.0320276  0.0116882  -2.740  0.00655 **
tt          -0.0001432  0.0000533  -2.686  0.00768 **
z.diff.lag1  0.3885315  0.0594140   6.539 3.03e-10 ***
z.diff.lag2 -0.1072465  0.0632457  -1.696  0.09108 .
z.diff.lag3  0.1066671  0.0590731   1.806  0.07207 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02596 on 273 degrees of freedom
Multiple R-squared: 0.1555, Adjusted R-squared:  0.14
F-statistic: 10.05 on 5 and 273 DF,  p-value: 7.566e-09
Value of test-statistic is: -2.7402 3.1295 3.8205

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -3.98 -3.42 -3.13
phi2  6.15  4.71  4.05
phi3  8.34  6.30  5.36
```

We test the null hypothesis that there is unit root in z. We do NOT reject that there is unit root if |test statistic|<|critical value|. Here, we do not reject that there is unit root process because test statistic in abs.values is smaller than abs. critical values at all levels of significance

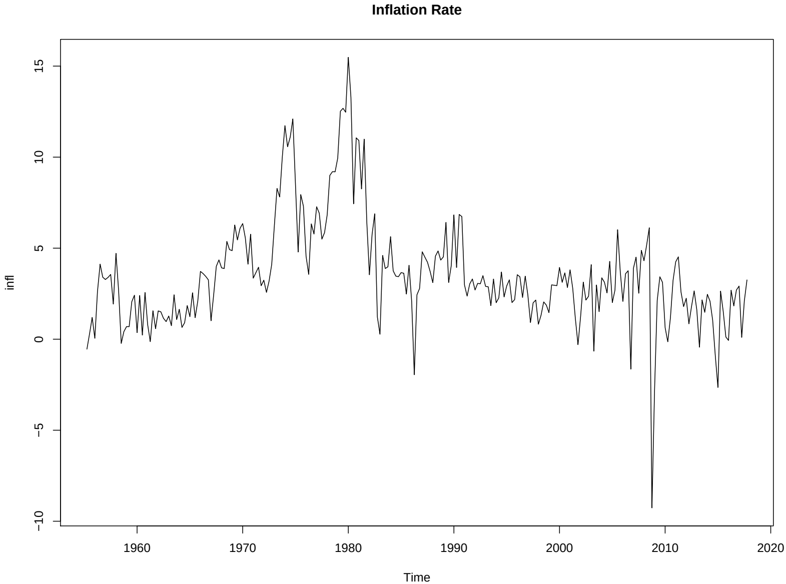
Example of nonstationarity – CPI



CPI Example: ADF Test (with trend)

```
Residuals:
    Min      1Q  Median      3Q     Max 
-6.8058 -0.1980 -0.0217  0.2637  1.5684 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.006405   0.113797   0.056  0.95516    
z.lag.1     -0.014000   0.004485  -3.121  0.00203 ** 
tt           0.015970   0.004757   3.357  0.00092 ***
z.diff.lag1  0.359362   0.064947   5.533 8.47e-08 ***
z.diff.lag2 -0.120005   0.069121  -1.736  0.08386 .  
z.diff.lag3  0.159501   0.069715   2.288  0.02304 *  
z.diff.lag4 -0.075021   0.069835  -1.074  0.28382    
z.diff.lag5  0.165761   0.069956   2.370  0.01863 *  
z.diff.lag6  0.076624   0.070007   1.095  0.27486    
z.diff.lag7 -0.010748   0.069984  -0.154  0.87807    
z.diff.lag8 -0.012494   0.065504  -0.191  0.84890    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7027 on 232 degrees of freedom
Multiple R-squared:  0.3324,	Adjusted R-squared:  0.3036 
F-statistic: 11.55 on 10 and 232 DF,  p-value: 4.768e-16


Value of test-statistic is: -3.1212 7.2212 6.4084 

Critical values for test statistics: 
      1pct  5pct 10pct
tau3 -3.98 -3.42 -3.13
```

Looking more stationary - Inflation



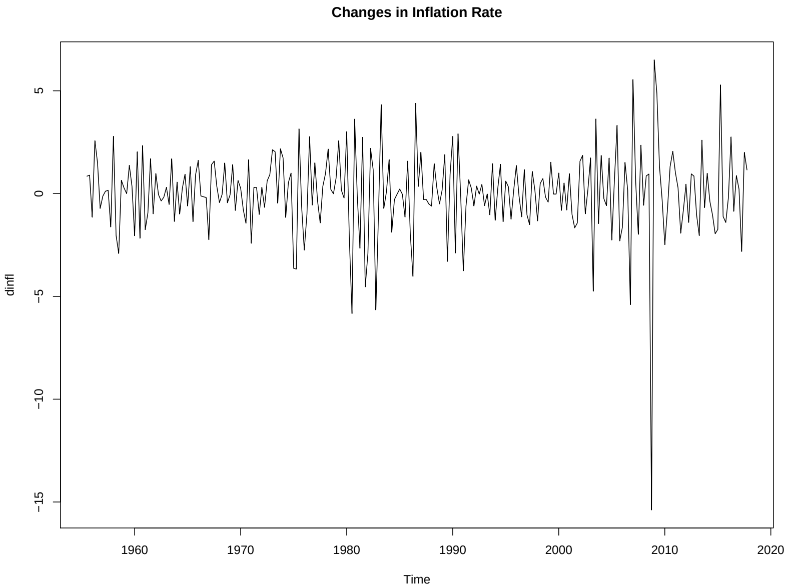
ADF Test on Inflation Rate (dCPI, without trend)

- Since, CPI has a unit root, we'll take the first (log) difference and test for a unit root in inflation rate.

```
Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-14.2157  -0.7876   0.1439   1.0485   4.5566

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.44326    0.20393   2.174  0.03074 *
z.lag.1     -0.12346    0.04574  -2.699  0.00745 **
z.diff.lag1 -0.29203    0.07146  -4.087 6.02e-05 ***
z.diff.lag2 -0.34474    0.07343  -4.695 4.54e-06 ***
z.diff.lag3 -0.02319    0.07054  -0.329  0.74263
z.diff.lag4 -0.12692    0.06511  -1.949  0.05247 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.849 on 234 degrees of freedom
Multiple R-squared:  0.2386,	Adjusted R-squared:  0.2223
F-statistic: 14.67 on 5 and 234 DF,  p-value: 1.652e-12


Value of test-statistic is: -2.6994 3.6437

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.44 -2.87 -2.57
```

Reject at the 10 percent level, but not at the 5 percent level

Example of stationarity - Changes in Inflation



ADF Test on Changes in Inflation Rate (dInfl, without trend)

- Since, Inflation rate may not be stationary, we'll take the first difference again and test for a unit root in changes in inflation rate.

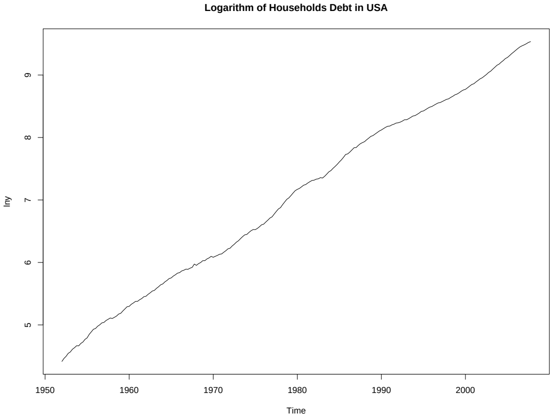
```
Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-14.3152  -0.7506   0.2220   1.0055   4.9350

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.01242    0.12107  -0.103    0.9184
z.lag.1     -2.02960    0.18531 -10.952  < 2e-16 ***
z.diff.lag1  0.65294    0.15485   4.217 3.55e-05 ***
z.diff.lag2  0.23406    0.10981   2.132   0.0341 *
z.diff.lag3  0.16247    0.06455   2.517   0.0125 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.872 on 234 degrees of freedom
Multiple R-squared:  0.6921,    Adjusted R-squared:  0.6868
F-statistic: 131.5 on 4 and 234 DF,  p-value: < 2.2e-16


Value of test-statistic is: -10.9525 59.9787

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.46 -2.88 -2.57
```

Nonstationarity due to trends

- A trend is a persistent long-term movement of a variable over time
- A deterministic trend is a nonrandom function of time
- A stochastic trend is random and varies over time
- A series exhibiting a stochastic trend may have long periods of increases or decreases (e.g. monthly stock prices)
- In economics and finance, we often model time series using stochastic trends.

Deterministic Trends

- A deterministic trend, for example, may be linear in time
- To handle deterministic trend, remove the trend. For a linear trend, run the following regression and use the residuals for modeling purposes: $Y_t = \alpha_0 + \alpha_1 t + e_t$.



Stochastic Trend: Random walk model of a trend

- Random walk: $Y_t = Y_{t-1} + u_t$, u_t is iid
- Called a random walk because the value at time t is the value at time $t-1$ plus some unpredictable change.
- Thus, Y_t consists of random "steps" u_t, that path makes up a "random walk"
- Recall that $E(u_t|Y_{t-1}, Y_{t-2}, \ldots) = 0$. So, $E(Y_t|Y_{t-1}, Y_{t-2}, \ldots) = Y_{t-1}$
- If Y_t is a random walk, the best forecast is Y_{t-1}
- $Y_{T+1|T} = Y_T$ Your best forecast of the value of Y in the future is the value of Y today
- The terms "stochastic trend" and "unit root" are used interchangeably.

Random Walk with drift

- $Y_t = \beta_0 + Y_{t-1} + u_t$
- $E(u_t | Y_{t-1}, Y_{t-2}, \ldots) = 0$
- β_0 is the "drift" term in the random walk
- If β_0 is positive, the Y_t increases on average
- Example: Stock Prices
- $Y_{T+h|T} = \hat{\beta}_0 h + Y_T$ Your best forecast of the value of Y in the future is the value of Y today plus an h-period influence of a drift

A random walk is nonstationary

- The random walk plus drift model is an AR(1) with $\beta_1 = 1$
- The random walk models are not stationary as the distribution of Y_t changes over time
- $var(Y_t) = var(Y_{t-1}) + var(u_t)$. In order for Y_t to be stationary, the variance must be constant over time.
- Thus, $var(Y_t) = var(Y_{t-1})$ which implies that $var(u_t) = 0$. This is false by assumption.
- Also we can write $Y_t = Y_0 + \beta_0 t + (u_1 + u_2 + ... + u_t)$ and $var(Y_t) = t\sigma^2$, depends on t (increases linearly with t).

Estimation of Random Walk

$$Y_t = \beta_0 + Y_{t-1} + \varepsilon_t$$

Under unit root process the series Y_t are not covariance stationary.

However, the differences in Y_t and Y_{t-1}:

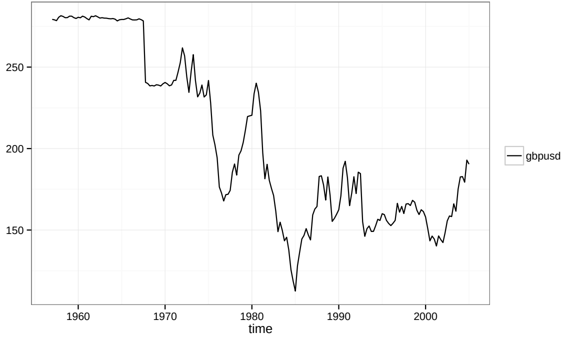
$$Y_t - Y_{t-1} = \Delta Y_t = \beta_0 + \varepsilon_t$$

are covariance stationary, because, by assumption $\varepsilon_t \sim WN(0, \sigma^2)$.

We say that a nonstationary series is integrated if its nonstationarity is appropriately "undone" by differencing.

Nonstationarity due to structural breaks

- Breaks arise from a change in the population regression coefficients at a distinct time or from a gradual evolution of the coefficients
- Example: In macroeconomic time series, the break down of the Bretton Woods system of fixed exchange rates in 1972 produced a break in the Dollar/Pound exchange rate



Nonstationarity due to structural breaks

- Problems Caused By breaks:
    - The OLS regression estimates over the full sample will estimate a relationship that holds "on average" as it combines the estimates over the two different periods.
- Model: $Y_t = \beta_0 + \beta_1 Y_{t-1} + \delta_1 X_{t-1} + u_t$
- Testing for a break date (Chow test):
    - Suppose that we hypothesize that the break date is at time τ.
    - Create a dummy variable st $D_t(\tau) = 1$ if $t > \tau$ and zero otherwise.
    - Augment the model:
    - $Y_t = \beta_0 + \beta_1 Y_{t-1} + \delta_1 X_{t-1} + \gamma_0 D_t(\tau) + \gamma_1 [D_t(\tau) \cdot Y_{t-1}] + \gamma_2 [D_t(\tau) \cdot X_{t-1}] + u_t$
    - Null hypothesis: $\gamma_0 = \gamma_1 = \gamma_2 = 0$ (No break)
    - F-test

Example – Testing for a break in the Inflation Rate

- The Phillips Curve links unemployment to inflation. The idea behind the Phillips curve was that there was a negative relationship between unemployment and inflation
- The Phillips Curve couldn't be used for prediction or policy if there was a structural break in the model.
- Model: $\Delta Inf_t = \beta_0 + \beta_1 \Delta Inf_{t-1} + \ldots + \beta_4 \Delta Inf_{t-4} + \delta_1 Unemp_{t-1} + \ldots + \delta_4 Unemp_{t-4} + u_t$
- Augmented model to test for breaks in the coefficients on unemployment:
- $$\Delta Inf_t = \beta_0 + \beta_1 \Delta Inf_{t-1} + \ldots + \beta_4 \Delta Inf_{t-4} + \delta_1 Unemp_{t-1} + \ldots + \delta_4 Unemp_{t-4} + \gamma_0 D_t(\tau) + \gamma_1 D_t(\tau) Unemp_{t-1} + \cdots + \gamma_4 D_t(\tau) Unemp_{t-4} + u_t$$
- Null Hypothesis: $\gamma_0 = \cdots = \gamma_4 = 0$ (no break, 5 restrictions)

Testing for an unknown break date

- Use Quant-likelihood Ratio (QLR) statistic
  - Let $F(\tau)$ be the Chow test statistic testing the hypothesis of no break at date τ.
  - Calculate the test F-statistic from the Chow Test for a range of different values for τ, $0.15T \leq \tau \leq 0.85T$.
  - The QLR statistic is then the largest of these F-statistics.
  - The QLR statistic doesn't follow one of our usual distributions; thus, we must use the simulated critical values from Andrews (2003)

TABLE 14.6 Critical Values of the QLR Statistic with 15% Trimming			
Number of Restrictions (q)	**10%**	**5%**	**1%**
1	7.12	8.68	12.16
2	5.00	5.86	7.78
3	4.09	4.71	6.02
4	3.59	4.09	5.12
5	3.26	3.66	4.53
6	3.02	3.37	4.12
7	2.84	3.15	3.82
8	2.69	2.98	3.57
9	2.58	2.84	3.38
10	2.48	2.71	3.23
11	2.40	2.62	3.09
12	2.33	2.54	2.97
13	2.27	2.46	2.87
14	2.21	2.40	2.78
15	2.16	2.34	2.71
16	2.12	2.29	2.64
17	2.08	2.25	2.58
18	2.05	2.20	2.53
19	2.01	2.17	2.48
20	1.99	2.13	2.43

These critical values apply when $\tau_0 = 0.15T$ and $\tau_1 = 0.85T$ (rounded to the nearest integer), so the F-statistic is computed for all potential break dates in the central 70% of the sample. The number of restrictions q is the number of restrictions tested by each individual F-statistic. Critical values for other trimming percentages are given in Andrews (2003).

Has the postwar U.S. Phillips Curve been stable?

Recall the ADL(4,4) model of ΔInf and $Unemp$ – the empirical backwards-looking Phillips curve, estimated over (1955 – 2017):

$$\widehat{\Delta Inf}_t = 1.09 - 0.42\Delta Inf_{t-1} - 0.47\Delta Inf_{t-2} - 0.13\Delta Inf_{t-3} - 0.19\Delta Inf_{t-4}$$
$$- 1.94 Unemp_{t-1} + 2.53 Unemp_{t-2} - 1.03 Unemp_{t-3} + 0.26 Unemp_{t-4}$$

Has this model been stable over the full period 1955-2017? Test for constancy of intercept and coefficients on $Unemp_{t-1}, \ldots, Unemp_{t-4}$: QLR = 7.11 (q = 5)

- 1% critical value = 4.53. Reject at 1% level
- Estimate break date: maximal F occurs in 1982:IV
- Conclude that there is a break in the inflation – unemployment relation, with estimated date of 1982:IV

Example – Testing for an unknown break in the Inflation Rate