

Macroeconometria - Séries de tempo

FAUSTO JOSÉ ARAÚJO VIEIRA

Aula 4

17 de abril a 22 de maio de 2018

RESUMO - MODELO ARMA(AULA ANTERIOR)

Revisão: processo de uma série ARMA(p,q)

1. Ajustar sazonalmente as variáveis econômicas
2. Modelos: Média móvel, variável binária, SARIMA, X-12 e TRAMO SEATS.
3. Estimar, verificar os resíduos e projetar:
 - ▶ Modelos sem sazonalidade - aula 2
 - ▶ Modelos com sazonalidade - aula 3
4. Modelos mais completos, adicionando outras variáveis explicativas.
5. Artigo do Ang et al (2006) - continuaremos nas próximas aulas.
 - ▶ Concluimos a análise dos modelos ARIMA.

Sazonalidade e suavização

- ▶ O processo estocástico é produto de quatro fatores:

$$y_t = C_t \times S_t \times T_t \times U_t$$

C_t é um componente de ciclo de longo prazo;

S_t é um componente sazonal;

T_t é um componente de tendência;

U_t é um componente irregular.

- ▶ Objetivo: estimar S_t e, em seguida, expurgar esse termo da série y_t , para fins de previsão.

Resumo - métodos de ajuste sazonal

Método	Positivo	Negativo
Média móvel tradicional	Fácil de executar, pode ser feito no excel	Pressupõe fatores sazonais constantes
Variável binária	Estimado por MQO e adicionar a equações com variáveis explicativas	Fatores constantes e a hipótese de não correlação das variáveis binárias com alguma variável omitida
Suavização - EWMA	Fácil execução	Determinar o valor de suavização e os valores iniciais
ARMA sazonal	Fatores sazonais variáveis	Não há um modelo único
X-12	Fatores sazonais variáveis e opção de escolha automática	Testes são <i>ad-hoc</i>
Tramo	Fatores sazonais variáveis e análise estatística	pressupõe normalidade dos resíduos

Modelos com defasagens degeneradas

- ▶ Incluir no modelo:

- ▶ Defasagens - garantindo que o resíduo seja um RB;
- ▶ *Dummies* sazonais, caso a série não seja ajustada sazonalmente;
- ▶ Variáveis explicativas que tenham poder preditivo da variável dependente.

- ▶ Modelo geral:

$$y_t = \alpha + \delta_1 y_{t-j} + \delta_2 y_{t-k} + \beta X_t + \gamma D_t + e_t \text{ para } j \text{ e } k > 0$$

- ▶ Importante: resíduo é um RB!

Artigo - Ang et al (2006)

► Modelos

- ARIMA: $\pi_{t+1} = \mu + \phi\pi_t + \psi\varepsilon_t + \varepsilon_{t+1}$
- Curva de Phillips: $\pi_{t+4,4} = \alpha + \beta(L)'X_t + \varepsilon_{t+4,4}$
- Modelo linear da estrutura a termo: $X_{t+1} = \mu + \Phi X_t + \varepsilon_{t+1}$
- Usando projeções de mercado: $\pi_{t+4,4} = \alpha_1 + \beta_1 f_t^S + \varepsilon_{t+4,4}$

► Metodologia de projeção:

$$E_t(\pi_{t+4,4}) = E_t\left(\sum_{i=1}^4 \pi_{t+i}\right)$$

$$\pi_{t+4,4} = \pi_{t+1} + \pi_{t+2} + \pi_{t+3} + \pi_{t+4}$$

RESUMO DA AULA

The background features abstract, overlapping geometric shapes in various shades of green, ranging from light lime to dark forest green. These shapes are primarily located on the right side of the slide, creating a modern, layered effect. The text 'RESUMO DA AULA' is positioned on the left side of the slide in a clean, sans-serif font.

Resumo

1. Tendência estacionária e tendência estocástica
2. Teste Dickey-Fuller
3. Teste Phillip-Perron
4. Teste KPSS
5. Raiz unitária - quebra estrutural
6. Aplicação
7. Filtro Hodrick Prescott
8. Decomposição de Beveridge-Nelson

MODELOS DE AJUSTE SAZONAL



Motivação

- ▶ A série temporal não estacionária não pode ser estimada trivialmente.
- ▶ Problema: é impossível estimar todos os momentos da série e fazer inferências estatísticas.

A variância não condicional de um $AR(1)$ é:

$$\text{var}(y_t) = \frac{1}{1 - \phi^2}.$$

Se $\phi = 1$, o que caracteriza uma série não estacionária de raiz unitária, então a variância explode.

- ▶ Solução: diferenciar a série tantas vezes quantas sejam necessárias para estacionarizá-la.

Tendência estacionária

Suponha o seguinte modelo:

$$y_t = y_0 + \delta t + \psi(L) \varepsilon_t.$$

Tal modelo é chamado de **tendência estacionária**, porque flutua em torno de uma tendência determinística.

A série também poderia ser estacionarizada pela primeira diferença, isto é:

$$\Delta y_t \equiv (1 - L) y_t = y_t - y_{t-1} = \delta + (1 - L) \psi(L) \varepsilon_t.$$

Essa diferenciação estacionariza a série, entretanto torna o erro não invertível.

Logo, se uma série é tendência estacionária, é melhor estimá-la usando a variável explicativa t .

Tendência estocástica

Considere outra possibilidade:

$$\Delta y_t = \delta + \varepsilon_t \implies y_t = y_{t-1} + \delta + \varepsilon_t.$$

Compondo recursivamente y_t , obtém-se:

$$y_t = y_0 + \delta t + \sum_{i=1}^t \varepsilon_i.$$

A variável aleatória y_t é dada pela composição de todos os choques havidos, $\sum_{i=1}^t \varepsilon_i$.

Define-se tal série como sendo **tendência estocástica** ou **diferença estacionária**.

Os choques produzem mudanças permanentes na série y_t , ainda que aleatórias.

Séries, cuja tendência é estocástica, são **séries integradas** e denotadas por $I(d)$, em que d é a ordem de integração. Séries integradas com erros estacionários são chamadas de séries *ARIMA* (p, d, q) . Diferenciando d vezes a série, obtém-se uma série estacionária.

Passeio aleatório

No caso $I(1)$ com $\delta = 0$, define-se o **passeio aleatório** ou **tendência estocástica pura** pela equação:

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

A previsão condicional H passos à frente é dada pela observação atual, isto é:

$$E_t(y_{t+H}) = y_t + \sum_{h=1}^H E_t(\varepsilon_{t+h}) = y_t.$$

A covariância é dependente do tempo:

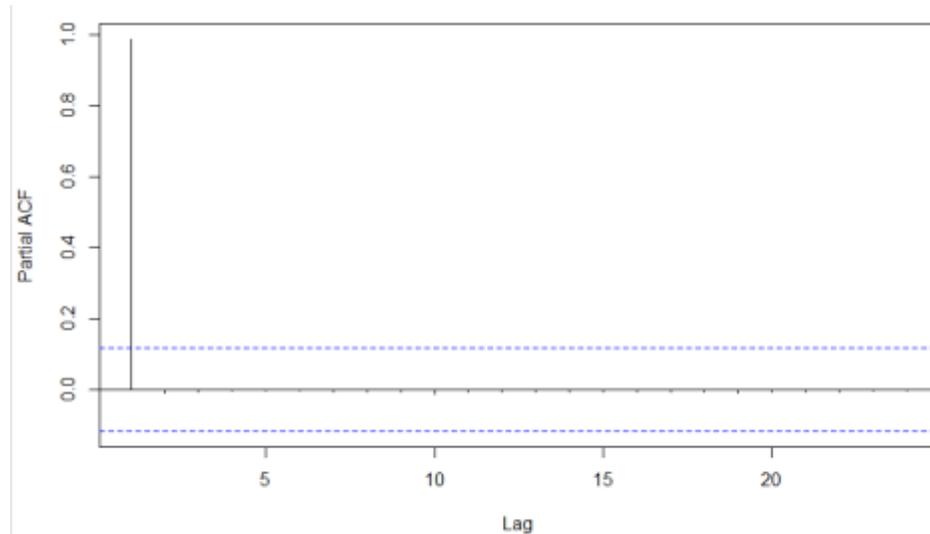
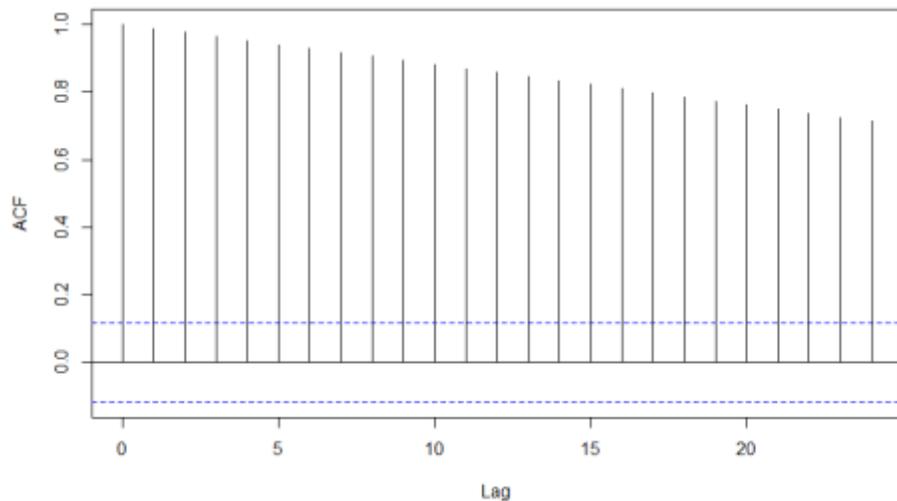
$$\begin{aligned} \text{Var}(y_t) &= \text{Var}\left(\sum_{i=1}^t \varepsilon_i\right) = t\sigma^2; \\ \text{Cov}(y_t, y_{t-j}) &= E\left(\sum_{i=1}^t \varepsilon_i\right)\left(\sum_{s=1}^{t-j} \varepsilon_s\right) = (t-j)\sigma^2. \end{aligned}$$

Passeio aleatório

Divide-se a covariância pelo produto do desvio padrão em t e $t-j$:

$$\rho_j = \frac{(t-j)\sigma^2}{\sqrt{t}\sigma\sqrt{(t-j)}\sigma} = \sqrt{\frac{t-j}{t}} = \sqrt{1 - \frac{j}{t}}$$

Num processo não estacionário, a autocorrelação demora a cair, pois j/t se reduz lentamente.



Tendência estocástica com drift

Ao adicionar um *drift* ao modelo, encontra-se o **passeio aleatório com drift**:

$$\begin{aligned}y_t &= y_{t-1} + \delta + \varepsilon_t = \\ &= y_0 + \delta t + \sum_{i=1}^t \varepsilon_i.\end{aligned}$$

Nesse caso, o comportamento de y_t depende de um componente determinístico e de outro estocástico. A previsão H passos à frente é:

$$E_t(y_{t+H}) = y_t + \delta H + \sum_{h=1}^H E_t(\varepsilon_{t+h}) = y_t + \delta H.$$

Tendência estocástica com drift e ruído

É possível generalizar o modelo de passeio aleatório adicionando um ruído a ele. É o **passeio aleatório com ruído**:

$$y_t = y_0 + \sum_{i=1}^t \varepsilon_i + \eta_t,$$

em que

$\{\eta_t\}$ é um ruído branco;

$\varepsilon_t \perp \eta_{t-j}$.

Pode-se, com isso, encontrar que:

$$\Delta y_t = \varepsilon_t + \Delta \eta_t.$$

Tendência estocástica com drift e ruído

Importância: ser $I(1)$ com uma correlação menor do que naquele passeio aleatório puro, em razão da presença de σ_η^2 :

$$\text{Var}(y_t) = \text{Var}\left(\sum_{i=1}^t \varepsilon_i + \eta_t\right) = t\sigma^2 + \sigma_\eta^2;$$

$$\text{Cov}(y_t, y_{t-j}) = E\left(\sum_{i=1}^t \varepsilon_i + \eta_t\right)\left(\sum_{s=1}^{t-j} \varepsilon_s + \eta_{t-j}\right) = (t-j)\sigma^2;$$

$$\rho_j = \frac{(t-j)\sigma^2}{\sqrt{(t\sigma^2 + \sigma_\eta^2) [(t-j)\sigma^2 + \sigma_\eta^2]}} < \sqrt{1 - \frac{j}{t}}.$$

Tendência geral mais componente irregular

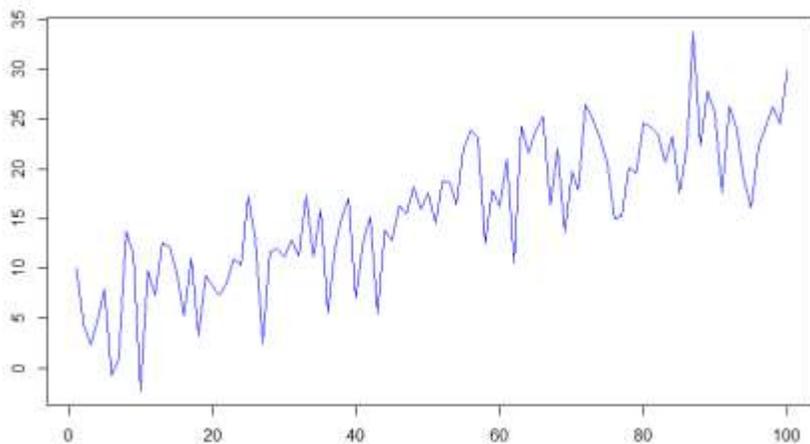
O modelo mais geral possível inclui tendência determinística e estocástica e resíduos que seguem um processo $ARMA(p, q)$. O modelo é chamado de **tendência geral mais componente irregular**:

$$y_t = \delta t + y_0 + \sum_{i=1}^t \varepsilon_i + \psi(L) \eta_t. \quad (1)$$

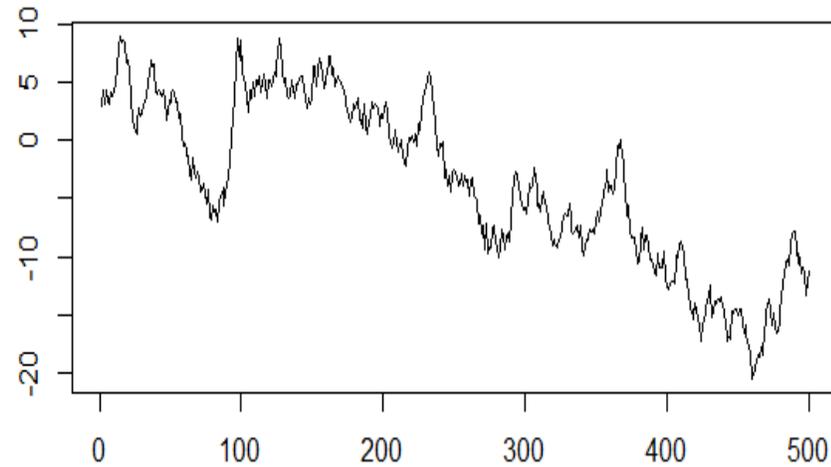
Tendência

- Exercício feito na aula 2
- u_t é um ruído branco

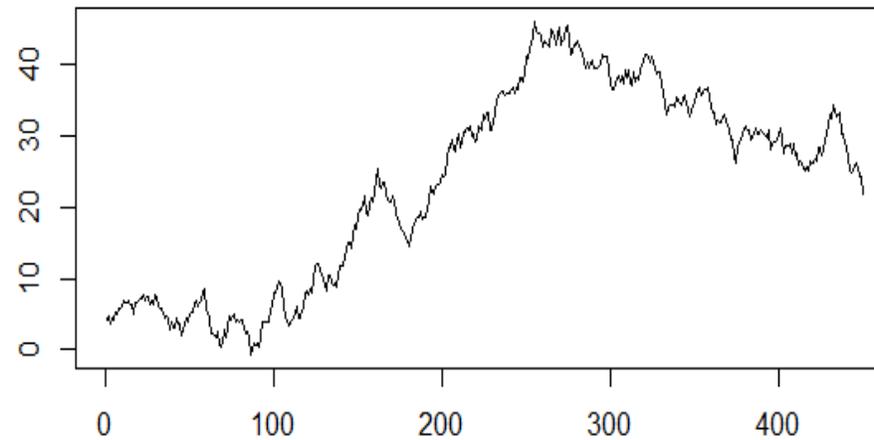
$$x_t = \mu + \delta t + u_t$$



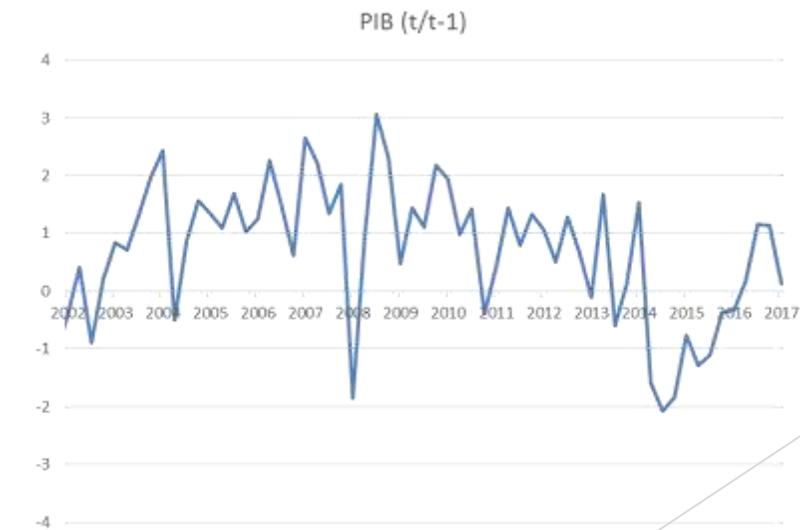
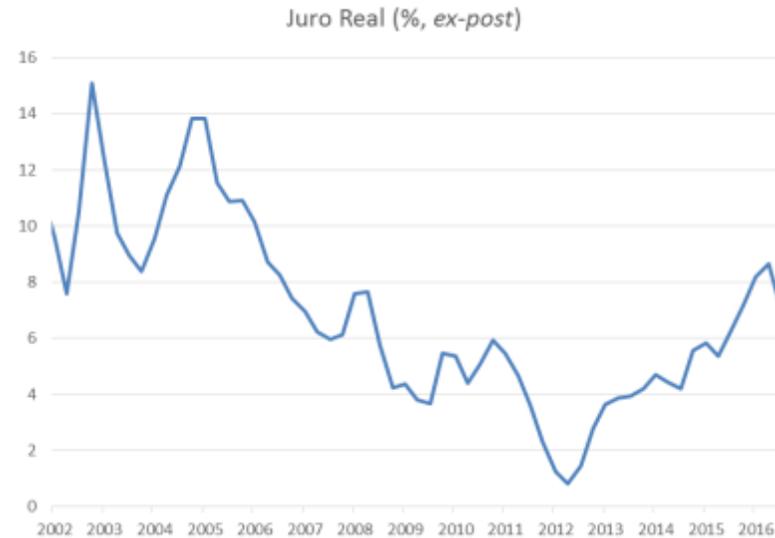
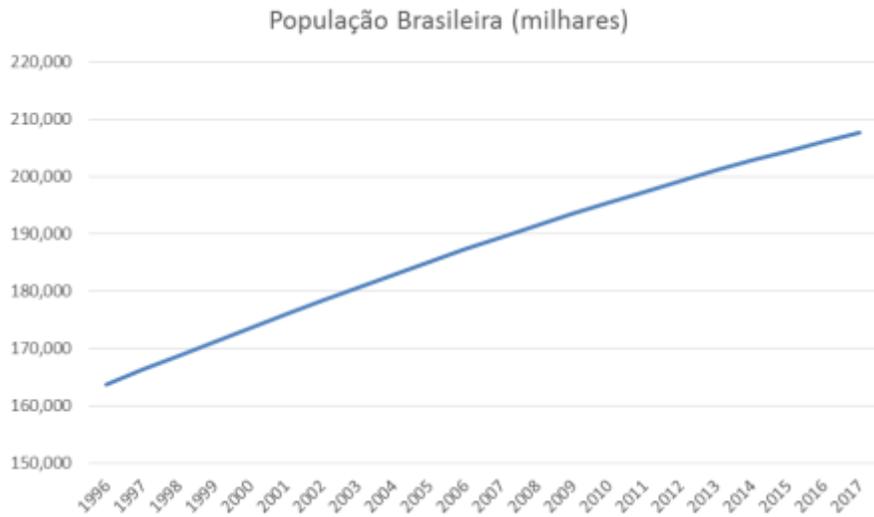
$$x_t = x_{t-1} + u_t$$



$$x_t = \mu + x_{t-1} + u_t$$



Como classificar as séries econômicas?



Removendo a tendência

No modelo com tendência estocástica, basta diferenciá-lo, inclusive se houver tendência determinística:

$$\Delta y_t = \delta + \varepsilon_t + \psi(L) \Delta \eta_t.$$

Se y_t for integrado de ordem d , toma-se a d -ésima diferença. Porém, como estimar uma série cuja tendência é determinística?

- ▶ Estime por mínimos quadrados ordinários:

$$y_t = \delta_0 + \delta_1 t + \delta_2 t^2 + \cdots + \delta_n t^n + e_t,$$

em que $e_t = \psi(L) \varepsilon_t$, e obtenha os resíduos estimados: \hat{e}_t .

- ▶ Estime o modelo $ARMA(p, q)$ a partir dos resíduos estimados.

Para determinar n , use testes t , F ou AIC/BIC . Em geral, estima-se o modelo com um n máximo, n_{\max} . Se o teste t sobre $\delta_{n_{\max}}$ não é rejeitado, retira-se t^n e estima-se o modelo até t^{n-1} , repetindo o teste. Proceda-se assim até rejeitar que $\delta_{n-i} = 0$.

Removendo a tendência

- ▶ É proibido diferenciar uma série que é tendência estacionária, porque isso adiciona ruído à série original.
- ▶ É proibido estimar uma série que é tendência estocástica usando tendência determinística, porque isso não elimina a tendência estocástica.

Regressão Espúria

- ▶ A necessidade de teoria econômica para definir variável explicada e explicativa torna-se muito importante na presença de raiz unitária.
- ▶ Podem-se encontrar relações estatísticas entre duas ou mais variáveis econômicas sem qualquer relação de causalidade entre uma e outra por puro acaso.
- ▶ Por exemplo, a regressão de uma variável $I(1)$ com outra $I(1)$ obtida independentemente gera alto R^2 e significativo t -estatístico. Contudo, o resultado é sem significado econômico.

Regressão Espúria

Considere a seguinte experiência. Gere duas séries I (1) independentemente uma da outra e regrida uma contra a outra. Qual resultado você obtém? Em 75% das vezes, parecer-lhe-á que elas são correlacionadas.

Importante lembrar que espera-se que 95% das vezes, espera-se que as séries não sejam correlacionadas.

Regressão Espúria

Criar um script no R

Digitar as seguintes linhas:

```
y<- e <- rnorm(100,0,1)
```

```
y[0]=0
```

```
y[1]=0
```

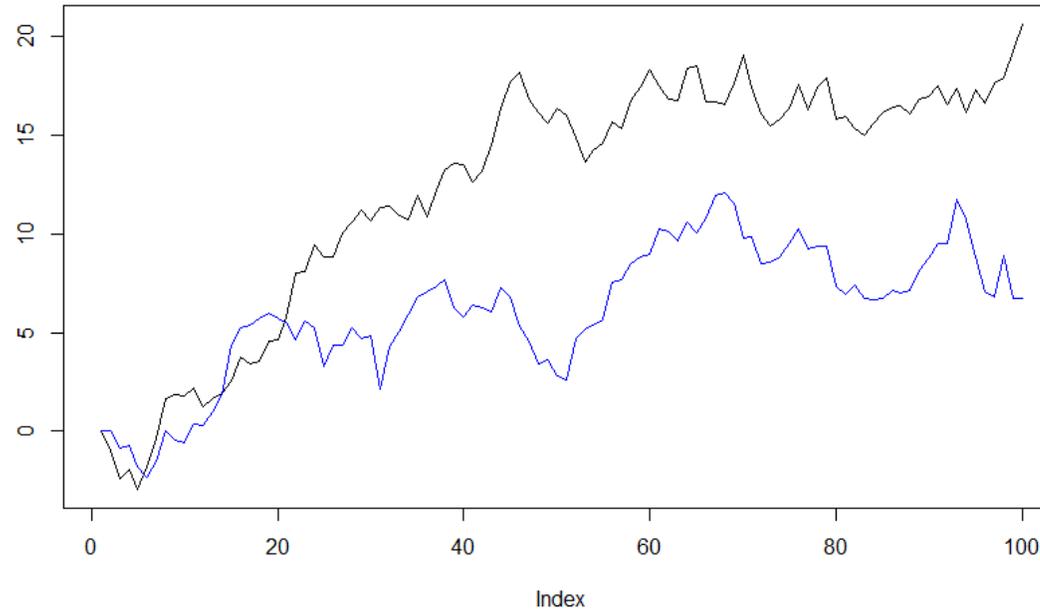
```
for (t in 2:100) y[t] <- e[t]+1*y[t-1]
```

```
plot(y,type= "l")
```

```
z<-y
```

- ▶ Rodar novamente, substituir a última linha por **x<-y**

Regressão Espúria



```
eq=lm(z~x)
summary(eq)
```

```
Call:
lm(formula = z ~ x)

Residuals:
    Min       1Q   Median       3Q      Max
-8.204 -1.659  0.070  1.787  9.049

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.0475     0.7403   4.117 8.02e-05 ***
x            1.5184     0.1065  14.261 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.594 on 98 degrees of freedom
Multiple R-squared:  0.6748,    Adjusted R-squared:  0.6715
F-statistic: 203.4 on 1 and 98 DF,  p-value: < 2.2e-16
```

CONCLUSÃO: CUIDADO COM A REGRESSÃO QUE SE ESTIMA!!!!

Dickey-Fuller

Considere o seguinte modelo:

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

Tendência: estimar esse modelo e usar um teste convencional de t sobre ϕ , tendo como hipótese nula $H_0 : \phi = 1$.

Alternativamente, poder-se-ia alterar o teste subtraindo y_{t-1} de ambos os lados:

$$\Delta y_t = (\phi - 1) y_{t-1} + \varepsilon_t = \alpha y_{t-1} + \varepsilon_t, \quad (2)$$

em que se define $\alpha \equiv \phi - 1$.

Assim, $H_0 : \phi = 1$ é equivalente a $H_0 : \alpha = 0$.

Problema: sob a nula, a distribuição do teste não é convencional, ou seja, não é igual à distribuição t estatística, pois y_t não é estacionário.

Dickey-Fuller: experimento de Monte Carlo

Visualização:

- ▶ Gere uma seqüência de erros normais com esperança nula e variância σ^2 , $\{\varepsilon_t\}$, com $T + n$ observações;
- ▶ Gere a seqüência $\{y_t\}$ sob a hipótese nula de raiz unitária, dado y_0 ;
- ▶ Estime a equação (2) usando as T últimas realizações e armazene o valor da estatística t ;
- ▶ Retorne a 1 outras S vezes (em geral, $S \geq 10.000$);
- ▶ Faça o gráfico da distribuição da estatística t .

$$y_t = \delta t + y_0 + \sum_{i=1}^t \varepsilon_i + \psi(L) \eta_t. \quad (1)$$

$$\Delta y_t = (\phi - 1) y_{t-1} + \varepsilon_t = \alpha y_{t-1} + \varepsilon_t, \quad (2)$$

Dickey-Fuller: experimento de Monte Carlo

Replicando o teste no R

Transformando as variáveis:

```
x_diff=diff(x)
```

```
z_diff=diff(z)
```

```
x_lag=x[1:99]
```

```
z_lag=z[1:99]
```

Estimando a equação:

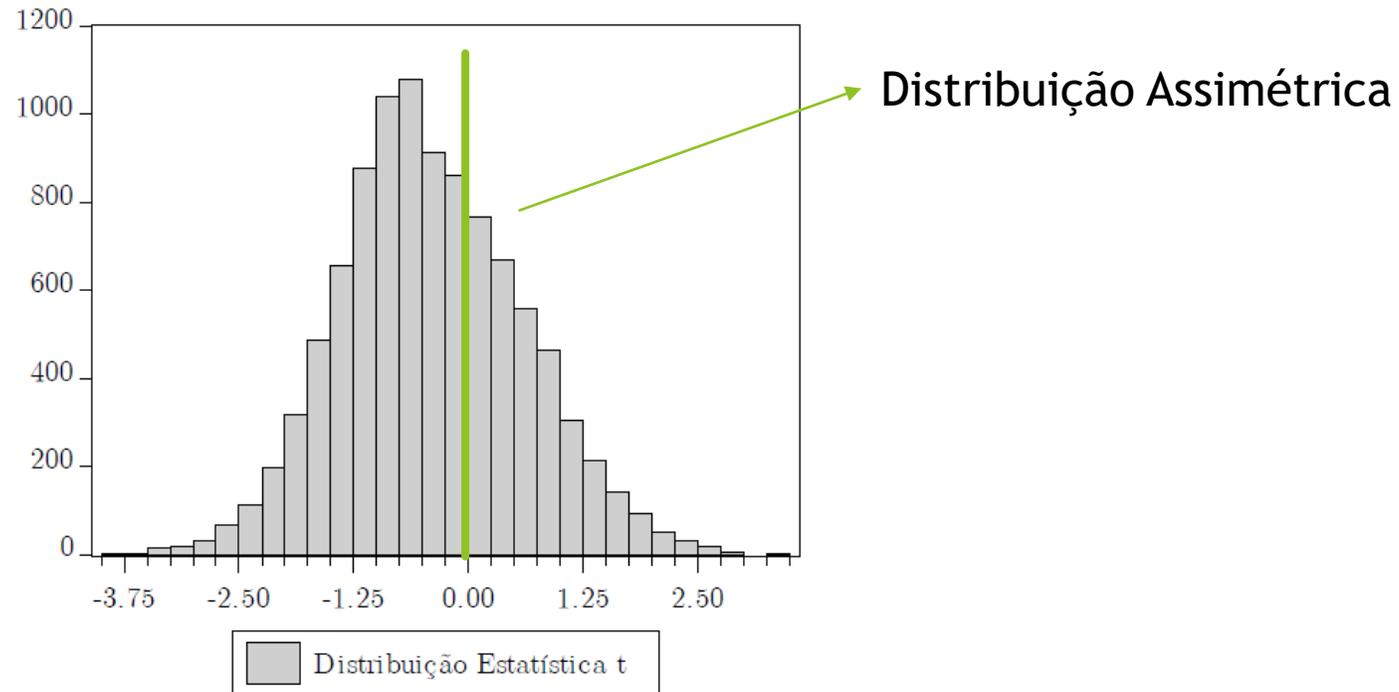
```
lm(x_diff~x_lag-1)
```

```
lm(z_diff~z_lag-1)
```

O TESTE REFAZ N VEZES (normalmente $n > 10.000$)

Dickey-Fuller: experimento de Monte Carlo

A Figura a seguir mostra o histograma empírico dessa estatística em que $T = 100$, $n = 50$, $S = 10.000$ e $y_0 = 0$.



Distribuição da estatística t – *student* sob $H_0 : \phi = 1$.

Dickey-Fuller: experimento de Monte Carlo

Buscamos a $H_1: \Phi < 1$

A média da estatística t não é zero.

Em 10% das vezes, a estatística $t < -1,60$; em 5%, $t < -1,95$; e em 1%, $t < -2,60$.

Ou seja, o uso da estatística t olhando para a tabela convencional implicaria cometer o erro do tipo I com muito mais frequência.

Erro do tipo I: Quando a hipótese nula é verdadeira e você a rejeita.

Por isso há uma adaptação na estatística para estimação do teste.

Dickey-Fuller: experimento de Monte Carlo

O valor da estatística t se altera, conforme se define a equação de regressão e segundo o tamanho da amostra:

$$\Delta y_t = \alpha y_{t-1} + \varepsilon_t \rightarrow \tau;$$

$$\Delta y_t = \mu + \alpha y_{t-1} + \varepsilon_t \rightarrow \tau_\mu;$$

$$\Delta y_t = \mu + \delta t + \alpha y_{t-1} + \varepsilon_t \rightarrow \tau_\tau.$$

Sob $H_0 : \alpha = 0$, as três estatísticas associadas foram obtidas por meio de simulações de Monte Carlo.

Dickey-Fuller: experimento de Monte Carlo

Valores críticos assintóticos para os testes DF (Dickey-Fuller)

estatística	1 %	2.5%	5%	10%
const. e tend.	-3.96	-3.67	-3.41	-3.13
só const.	-3.42	-3.12	-2.86	-2.57
sem const.	-2.58	-2.23	-1.95	-1.62

Fonte: ISEG

Dickey-Fuller: calculando a estatística DF

- ▶ Supondo $T + 1$ observações, $\{y_t\}_{t=0}^T$, faça OLS e subtraia 1 do parâmetro ϕ , para proceder ao teste sob $H_0 : \alpha = 0$:

$$\hat{\alpha} = \frac{\sum_{t=1}^T y_{t-1} y_t}{\sum_{t=1}^T y_{t-1}^2} - 1.$$

- ▶ Calcule a variância amostral:

$$s^2 = \frac{\sum_{t=1}^T (\Delta y_t - \hat{\alpha} y_{t-1})^2}{T}.$$

- ▶ Calcule o desvio padrão do coeficiente $\hat{\alpha}$, $s(\hat{\alpha})$:

$$s(\hat{\alpha}) = \frac{s}{\sqrt{\sum_{t=1}^T y_{t-1}^2}}.$$

- ▶ Obtenha o valor calculado da estatística t :

$$\hat{\tau} = \frac{\hat{\alpha}}{s(\hat{\alpha})}.$$

Dickey-Fuller aumentado

- ▶ Problema do teste anterior: o erro é um ruído branco. Será?

Suponha que y_t seja um processo auto-regressivo de ordem p , com raiz unitária:

$$y_t = \mu + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \phi_{p+1} y_{t-p-1} + \varepsilon_t.$$

- ▶ Como testar esse modelo para raiz unitária?
- ▶ Idéia: estimar o modelo com as variáveis auto-regressivas. Forma de corrigir o desvio do valor correto da estatística, ou seja, trata-se de encontrar os desvios de y_t em relação à sua "média", para deslocar a distribuição de α em direção a zero, caso a hipótese nula seja verdadeira.

Dickey-Fuller aumentado

Adicione e subtraia $\phi_{p+1}y_{t-p}$ à equação anterior:

$$\begin{aligned}y_t &= \mu + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \phi_{p+1} y_{t-p-1} + \\ &\quad + \phi_{p+1} y_{t-p} - \phi_{p+1} y_{t-p} + \varepsilon_t = \\ &= \mu + \phi_1 y_{t-1} + \cdots + (\phi_p + \phi_{p+1}) y_{t-p} - \phi_{p+1} \Delta y_{t-p} + \varepsilon_t.\end{aligned}$$

Utilizando o mesmo procedimento, desta vez com $(\phi_p + \phi_{p+1}) y_{t-p+1}$:

$$\begin{aligned}y_t &= \mu + \phi_1 y_{t-1} + \cdots + (\phi_p + \phi_{p+1}) y_{t-p+1} - (\phi_p + \phi_{p+1}) y_{t-p+1} + \\ &\quad + (\phi_p + \phi_{p+1}) y_{t-p} - \phi_{p+1} \Delta y_{t-p} + \varepsilon_t = \mu + \phi_1 y_{t-1} + \cdots + \\ &\quad + (\phi_{p-1} + \phi_p + \phi_{p+1}) y_{t-p+1} \\ &\quad - (\phi_p + \phi_{p+1}) \Delta y_{t-p+1} - \phi_{p+1} \Delta y_{t-p} + \varepsilon_t.\end{aligned}$$

Dickey-Fuller aumentado

Repetindo isso p vezes, obtém-se ao final:

$$\Delta y_t = \mu + \alpha y_{t-1} + \sum_{i=1}^p \lambda_i \Delta y_{t-i} + \varepsilon_t,$$

em que

$$\alpha = - \left(1 - \sum_{i=1}^{p+1} \phi_i \right);$$

$$\lambda_i = - \sum_{j=i}^p \phi_{j+1}.$$

O teste então pode ser feito, usando os mesmo valores críticos encontrados por Dickey e Fuller.

Dickey-Fuller aumentado

- ▶ E se o modelo for ainda mais complexo, com termos de médias móveis, o que fazer?
- ▶ Mesmo procedimento: se $ARIMA(m, 1, n)$, transformar um $MA(q)$ num $AR(\infty)$.
- ▶ Como estimar um modelo de infinitas defasagens?
- ▶ Provou-se que um modelo $ARIMA(m, 1, n)$ pode ser bem aproximado por um $ARIMA(p, 1, 0)$, em que $p \leq T^{\frac{1}{3}}$.
- ▶ Experimentos de Monte Carlo mostraram que o valor da estatística t permanece inalterado.

Demais testes de Dickey e Fuller

O teste de Dickey e Fuller pode ser feito conjuntamente para dois ou três coeficientes. Considere as seguintes especificações:

$$\Delta y_t = \alpha y_{t-1} + \sum_{i=1}^p \lambda_i \Delta y_{t-i} + \varepsilon_t;$$

$$\Delta y_t = \mu + \alpha y_{t-1} + \sum_{i=1}^p \lambda_i \Delta y_{t-i} + \varepsilon_t;$$

$$\Delta y_t = \mu + \delta t + \alpha y_{t-1} + \sum_{i=1}^p \lambda_i \Delta y_{t-i} + \varepsilon_t.$$

Dickey e Fuller (1981) calcularam estatísticas F para testes conjuntos, chamando-as de Φ_i , $i = 1, 2, 3$, com distribuições não convencionais.

Demais testes de Dickey e Fuller

As hipóteses a testar são:

$$H_0 : \alpha = \mu = 0 \rightarrow \Phi_1;$$

$$H_0 : \alpha = \delta = \mu = 0 \rightarrow \Phi_2;$$

$$H_0 : \alpha = \delta = 0 \rightarrow \Phi_3.$$

Dickey-Fuller aumentado - aplicação

Atualizando o arquivo

```
setwd("C:/Users/01366987608/Desktop/aula enap/series de tempo/aula2")
```

```
list.files()
```

```
X<-read.csv("aula4_exercicio.csv",sep=";", dec=".", head=TRUE)
```

Transformando as variáveis em séries temporais

```
ipca<-ts(X$ipca, start=c(2000,1), frequency=12)
```

```
populacao<-ts(X$populacao, start=c(2000,1), frequency=12)
```

```
varejo<-ts(X$varejo, start=c(2000,1), frequency=12)
```

```
selic<-ts(X$selic, start=c(2000,1), frequency=12)
```

```
selic_real<-ts(X$selic_real, start=c(2000,1), frequency=12)
```

```
brl<-ts(X$brl, start=c(2000,1), frequency=12)
```

Dickey-Fuller aumentado - aplicação

Instalando o pacote tseries - escolha automática

```
install.packages("tseries")
```

```
require(tseries)
```

```
detach("package:tseries",unload = TRUE) - retirando o pacote que tem função semelhante
```

Analisando o resultado para quatro séries

```
adf.test(ipca)
```

```
adf.test(populacao)
```

```
adf.test(selic)
```

```
adf.test(selic_real)
```

Escolha automática da defasagem e se há constante ou tendência

Dickey-Fuller aumentado - aplicação

Instalando o pacote fUnitRoots - escolha manual

```
install.packages("fUnitRoots")
```

```
require(fUnitRoots)
```

Fazer o gráfico das variáveis - decidir se o modelo é melhor com constante ou tendência

Analisando o resultado para quatro séries

```
adfTest(ipca,lags=1,type = "c") ou adfTest(ipca,lags=1,type = "nc")
```

```
adfTest(populacao,lags=2,type = "c") ou adfTest(populacao,lags=2,type = "ct")
```

```
adfTest(selic,lags=2,type = "c")
```

```
    adfTest(diff(selic),lags=2,type = "c")
```

```
adfTest(selic_real,lags=2,type = "c")
```

Phillips-Perron

- ▶ Usar um modelo auto-regressivo gera perda de graus de liberdade. Talvez fosse melhor um teste especificado independentemente das ordens p e q do modelo.
- ▶ Phillips e Perron (1988) usam essa idéia e propõem uma correção não paramétrica ao teste de Dickey e Fuller, gerando uma estatística consistente mesmo que haja variáveis defasadas dependentes e correlação serial nos erros.
- ▶ As equações estimadas e os testes designados são idênticos aos de Dickey e Fuller.
- ▶ A interpretação também é análoga.

Phillips-Perron

Possibilidades com as respectivas estatísticas associadas:

$$\begin{aligned}\Delta y_t &= \alpha y_{t-1} + u_t \rightarrow z_t, \\ \Delta y_t &= \mu + \alpha y_{t-1} + u_t \rightarrow z_{t,\mu}, \\ \Delta y_t &= \mu + \delta t + \alpha y_{t-1} + u_t \rightarrow z_{t,\tau},\end{aligned}$$

em que u_t é um processo estacionário.

- ▶ Phillips e Perron (1988) também definem testes diretamente sobre os coeficientes do modelo, em vez de usar a estatística t , como anteriormente. Eles chamaram tais testes de z_α .
- ▶ A correção, $z_{t,\mu}$, empregada por Phillips e Perron para τ_μ é seqüencialmente estimada da seguinte forma, dado y_0 :
- ▶ Estime as seguintes médias: $\bar{y} = \frac{\sum_{t=1}^T y_t}{T}$, $\bar{y}_{-1} = \frac{\sum_{t=1}^T y_{t-1}}{T}$;
- ▶ Estime o parâmetro de maior interesse: $\hat{\alpha} = \frac{\sum_{t=1}^T (y_{t-1} - \bar{y}_{-1})(y_t - \bar{y})}{\sum_{t=1}^T (y_{t-1} - \bar{y}_{-1})^2} - 1$;

Phillips-Perron

Dickey-Fuller

$$\hat{\alpha} = \frac{\sum_{t=1}^T y_{t-1} y_t}{\sum_{t=1}^T y_{t-1}^2} - 1$$

Phillips-Perron

$$\hat{\alpha} = \frac{\sum_{t=1}^T (y_{t-1} - \bar{y}_{-1})(y_t - \bar{y})}{\sum_{t=1}^T (y_{t-1} - \bar{y}_{-1})^2} - 1$$

Phillips-Perron - aplicação

Instalando o pacote tseries - escolha automática

Analisando o resultado para quatro séries

```
pp.test(ipca)
```

```
pp.test(populacao)
```

```
pp.test(selic)
```

```
pp.test(selic_real)
```

KPSS

Um dos problemas do teste de raiz unitária desenvolvido por DF e PP é seu baixo poder.

Ou seja, o teste de raiz unitária não consegue rejeitar a hipótese nula para alguns indicadores econômicos

KPSS - Kwiatkowski, Phillips, Schmidt e Shin (1992)

Hipóteses: $H_0 : y_t \sim I(0)$ contra $H_1 : y_t \sim I(1)$

$$\text{O teste: } y_t = x_t + u_t,$$

$$x_t = x_{t-1} + v_t$$

Ideia - testar a variância do resíduo de x_t . Se a variância de x_t é nula, então $x_t = x_{t-1} = c$

$$H_0 : \sigma_v^2 = 0 \times H_1 : \sigma_v^2 > 0$$

KPSS

Pode-se acrescentar uma tendência determinística ao modelo da seguinte forma:

$$\begin{aligned}x_t &= x_{t-1} + \delta + v_t \\ \Delta y_t &= \delta + v_t + \Delta u_t.\end{aligned}$$

KPSS - Aplicação

Instalando o pacote - tem que desativar o tseries :

```
install.packages("aTSA")
```

```
require(aTSA)
```

```
detach("package:tseries",unload = TRUE) #os dois pacotes tem comandos iguais
```

Analisando o resultado para quatro séries

```
kpss.test(ipca)
```

```
kpss.test(selic)
```

```
kpss.test(selic_real)
```

Quebra Estrutural

- ▶ Na presença de quebra estrutural, os testes são viesados na direção da não rejeição da hipótese de raiz unitária.
- ▶ Devem-se consultar os valores tabelados para teste com quebra estrutural por Perron, em que se assume uma **única** e **conhecida** quebra estrutural usando toda a amostra disponível.
- ▶ Considerando um passeio aleatório com *drift*, há três tipos de quebras estruturais possíveis. Uma mudança de nível da série, uma mudança de inclinação e ambas as mudanças.

Quebra Estrutural - mudança de nível

- ▶ Há uma quebra estrutural em $T_b < T$. Logo, há mudança de nível em $T_b + 1$:

$$H_0^A : y_t = \mu + y_{t-1} + d_1 DP_t + \varepsilon_t,$$
$$DP_t = \begin{cases} 1 & \text{se } t = T_b + 1, \\ 0, & \text{caso contrário,} \end{cases}$$

em que DP_t representa a *dummy* de nível transitória (ou pulso).

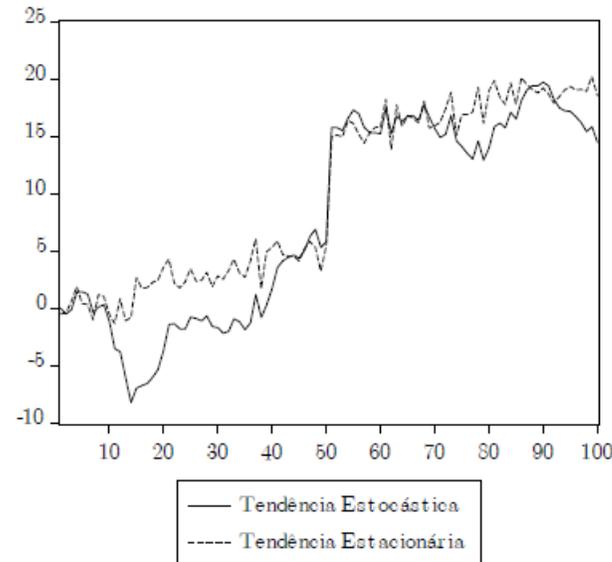
- ▶ O efeito de um choque no nível é permanente num processo com raiz unitária.
- ▶ Alternativa: o processo tendência estacionária tem mudança permanente de nível:

$$H_1^A : y_t = \mu + \delta t + d_2 DL_t + \varepsilon_t,$$
$$DL_t = \begin{cases} 1 & \text{se } t > T_b, \\ 0, & \text{caso contrário,} \end{cases}$$

em que DL_t é a *dummy* de nível, cuja mudança é permanente.

Quebra Estrutural - mudança de nível

- ▶ Por que, na hipótese alternativa, a mudança de intercepto é permanente, enquanto na hipótese nula, a mudança de intercepto é temporária?
- ▶ O choque temporário no intercepto da nula tem efeitos permanentes.



Séries simuladas com $\mu = 0,2$; $\delta = 0,1$; $d_1 = d_2 = 10$

Quebra Estrutural - Mudança Permanente de nível

No caso em que há mudança permanente de nível no modelo com raiz unitária, então a hipótese nula torna-se:

$$H_0^B : y_t = \mu + y_{t-1} + d_2 DL_t + \varepsilon_t.$$

O modelo alternativo equivalente a esse caso é uma mudança permanente da inclinação, isto é:

$$H_1^B : y_t = \mu + \delta t + d_3 DS_t + \varepsilon_t,$$
$$DS_t = \begin{cases} t - T_b & \text{se } t > T_b \\ 0, & \text{caso contrário,} \end{cases}$$

em que DS_t é uma tendência determinística, efetiva a partir de $T_b + 1$, quando o choque já se deu.

Quebra Estrutural - Mudança Permanente e Transitória de Nível

No caso mais geral, há mudança transitória e permanente de nível:

$$H_0^C : y_t = \mu + y_{t-1} + d_1 DP_t + d_2 DL_t + \varepsilon_t.$$

A hipótese alternativa é uma composição dos dois efeitos num modelo determinístico com quebra estrutural:

$$H_1^C : y_t = \mu + \delta t + d_2 DL_t + d_3 DS_t + \varepsilon_t.$$

Há dois procedimentos a seguir, conforme o tipo de ajustamento, instantâneo ou gradual.

Quebra Estrutural - modelagem

Obtenha os resíduos, \hat{y}_t^h , $h = A, B, C$ estimando a hipótese alternativa:

$$\hat{y}_t^A = y_t - \left(\hat{\mu} + \hat{\delta}t + \hat{d}_2 DL_t \right);$$

$$\hat{y}_t^B = y_t - \left(\hat{\mu} + \hat{\delta}t + \hat{d}_3 DS_t \right); \text{ ou}$$

$$\hat{y}_t^C = y_t - \left(\hat{\mu} + \hat{\delta}t + \hat{d}_2 DL_t + \hat{d}_3 DS_t \right).$$

Estime a seguinte regressão (distribuição de α depende da proporção de observações havidas antes do choque, $\lambda = \frac{T_b}{T}$):

$$\Delta \hat{y}_t^h = \alpha \hat{y}_{t-1}^h + \sum_{i=1}^p \lambda_i \Delta \hat{y}_{t-i}^h + \varepsilon_t.$$

Sob a nula, os resíduos serão $I(1)$. Logo, espera-se que $\alpha = 0$.

Compare o valor de τ^h com aqueles gerados por Perron. Exemplo, se $\lambda = 0,5$, $\tau^h = -3,76 < \tau = -3,41$ de D-F, a 5%.

Quebra Estrutural

- ▶ Para Zivot e Andrews (1992) Propõem um modelo em que o ponto de quebra é escolhido de forma que a quebra estrutural obtenha o maior peso possível para aceitar o modelo tendência estacionário.
 - ▶ Além disso, λ é escolhido de forma a minimizar a estatística "t-student".
- ▶ Perron (1997) propôs uma metodologia para encontrar endogenamente λ , à semelhança de Zivot e Andrews (1992).
 - ▶ λ é selecionado de forma a minimizar as estatísticas t associadas aos demais coeficientes d_2 ou d_3 .

Quebra Estrutural - aplicação

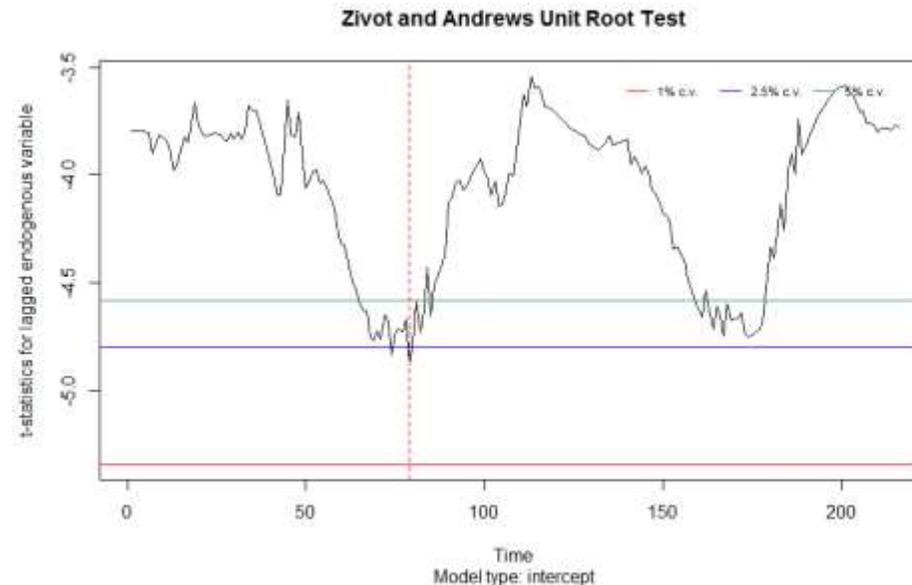
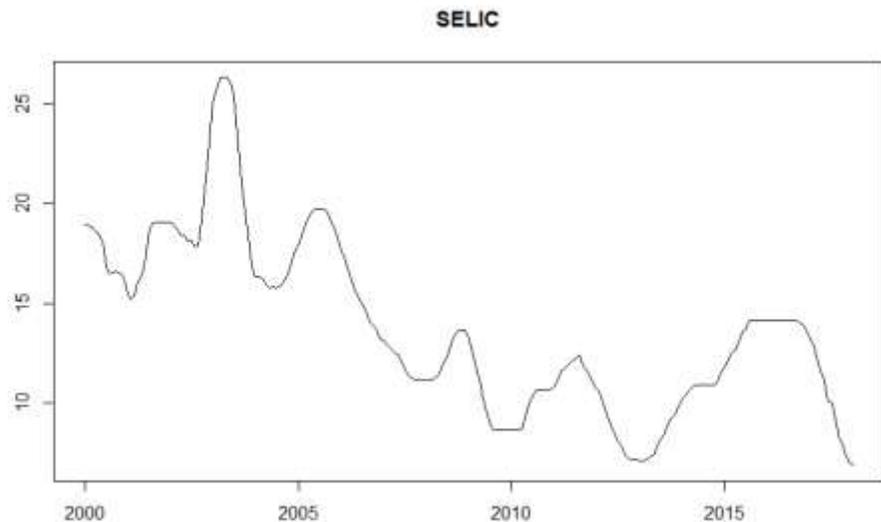
Instalando o pacote no R:

```
install.packages("urca")
```

```
require(urca)
```

Analisando se há quebra para a taxa Selic:

```
urzaTest(selic,model="intercept",lag=2)
```



Quebra Estrutural - aplicação

Analisando se há quebra para a população:

```
urzaTest(populacao,model="trend",lag=3)
```

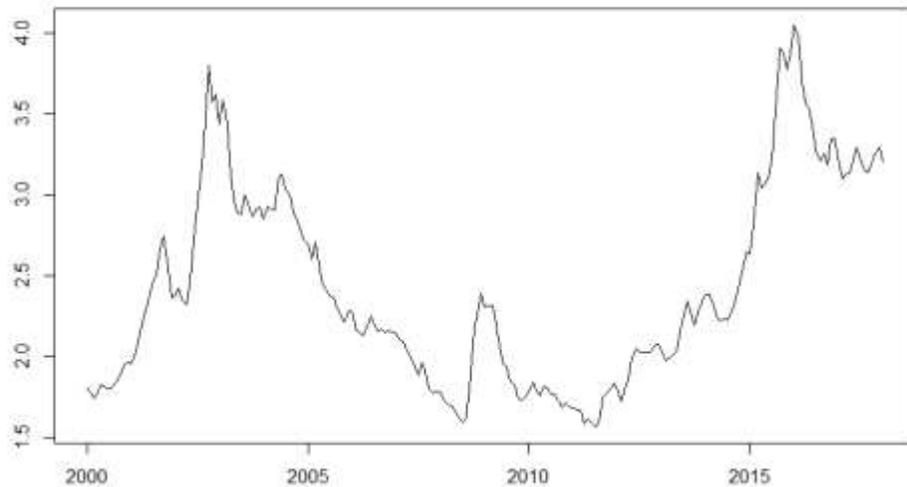
```
urzaTest(populacao,model="both",lag=3)
```

```
urzaTest(diff(populacao),model="intercept",lag=3)
```

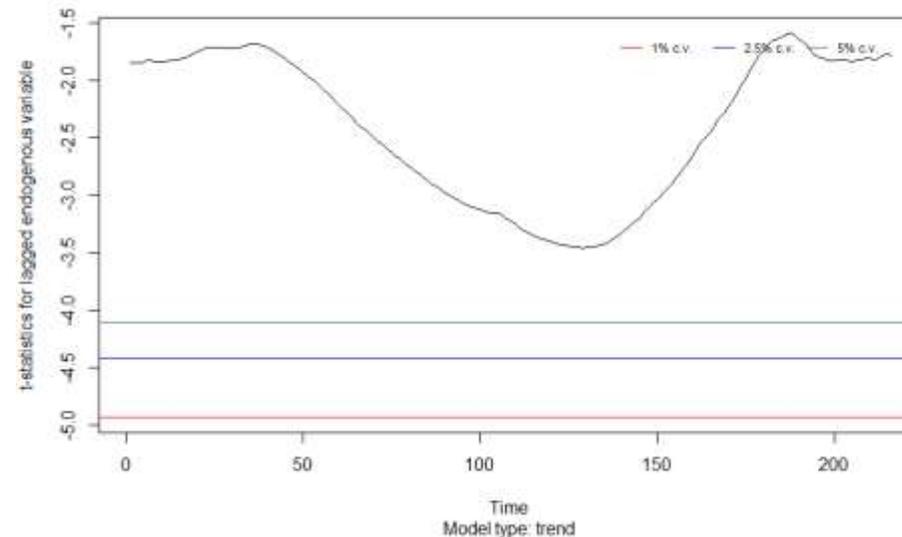
Analisando se há quebra para a taxa de câmbio:

```
urzaTest(brl,model="trend",lag=3)
```

Taxa de câmbio



Zivot and Andrews Unit Root Test



Artigo - ANPEC 2014

- ▶ “Validade da paridade de poder de compra para países selecionados da América Latina: teste de raiz unitária linear e não linear”
- ▶ O estudo examinou a validade da PPC em um contexto de teste de raiz unitária linear e não linear para a taxa de câmbio real efetiva dos países Argentina, Brasil, Chile, Colômbia, México, Peru e Venezuela.
- ▶ Os resultados evidenciaram que as séries com características lineares são Argentina, Brasil, Chile, Colômbia e Peru e não lineares México e Venezuela
- ▶ Os testes de raiz unitária lineares indicaram a taxa de câmbio real efetiva é estacionária para Chile e Peru e o teste não linear para o México.

Artigo - ANPEC 2014 (modelo)

- ▶ Sachs e Larrain - Macroeconomia (2000)

- ▶ Paridade do poder de compra:

$$P = E P^*$$

- ▶ A lei do preço único afirma que o preço interno P deve ser igual a $E P^*$
- ▶ Grama do ouro no Brasil em real é igual ao internacional multiplicado pelo câmbio:

$$146,0 = 3,43 \times 42,6$$

- ▶ Caso o preço seja diferente, haverá arbitragem
- ▶ A doutrina da paridade do poder de compra amplia a lei do preço único
 - ▶ Hipótese menos rígida é que a variação cambial replica o diferencial da inflação no médio prazo.
- ▶ Motivos para a falha da paridade do poder de compra:
 1. Barreiras comerciais;
 2. Todos os produtos são comercializados internacionalmente;
 3. Índice de preços doméstico tem a mesma cesta de bens do internacional

Artigo - ANPEC 2014

- ▶ Fonte: BIS (Bank for International Settlements)
- ▶ Período: janeiro de 1994 até o abril de 2014, totalizando 244 observações
- ▶ Países: Argentina, Brasil, Chile, Colômbia, México, Peru e Venezuela

Artigo - ANPEC 2014

Atualizando o arquivo

```
Z<-read.csv("aula4_artigoanpec.csv",sep=";", dec=".", head=TRUE)
```

Transformando as variáveis em séries temporais

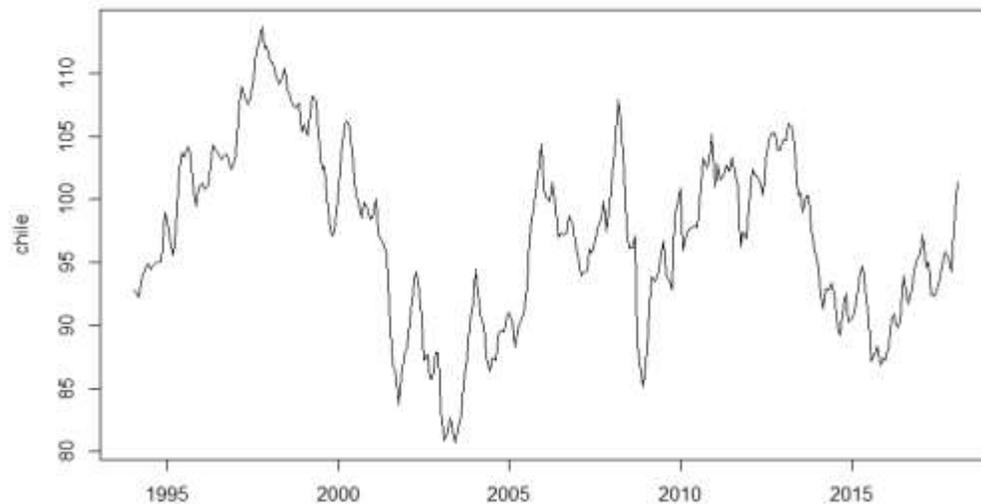
```
brasil<-ts(Z$Brazil, start=c(1994,1), frequency=12)
```

```
chile<-ts(Z$Chile, start=c(1994,1), frequency=12)
```

Artigo - ANPEC 2014 (Chile)

```
plot(chile)
```

```
adfTest(chile,lags=2,type = "c")
```



Title:
Augmented Dickey-Fuller Test

Test Results:
PARAMETER:
Lag Order: 2
STATISTIC:
Dickey-Fuller: -2.8725
P VALUE:
0.05096

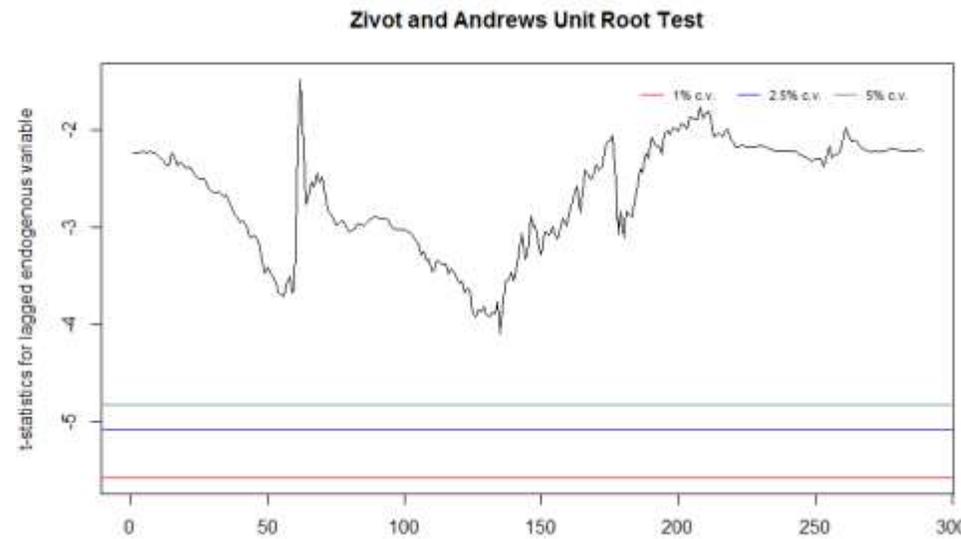
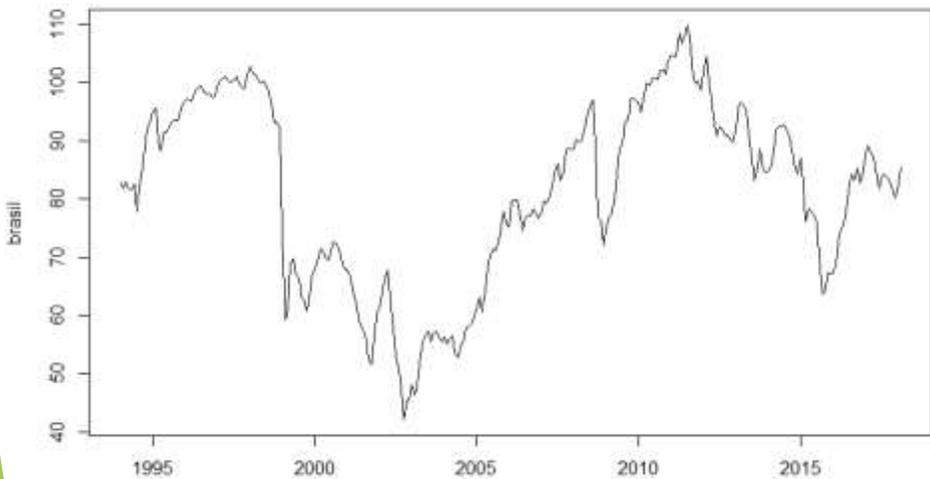
Artigo - ANPEC 2014 (Brasil)

```
plot(brasil)
```

```
adfTest(brasil,lags=1,type = "c")
```

```
adfTest(brasil,lags=1,type = "ct")
```

```
urzaTest(brasil,model="both",lag=1)
```



Artigo - ANPEC 2014 (Brasil)

Exercício:

Fazer o teste de raiz unitária para dois países

Caso não se rejeite H_0 , calcular o teste de raiz unitária com quebra estrutural.

Tendência estocástica x determinística

Como remover a tendência e tornar a série estacionária

► Tendência determinística:

Regressão linear

Filtro HP

► Tendência estocástica:

Diferenciação

Decomposição de Beveridge-Nelson

Filtro de Kalman

Filtro HP

Outros

Filtro Hodrick-Prescott

Trabalho original investiga os ciclos econômicos nos EUA no pós guerra.

O filtro HP decompõe a série y_t em componentes de tendência (crescimento de longo prazo - g_t) e cíclicos (c_t).

$$y_t = g_t + c_t \quad \text{for } t = 1, \dots, T$$

Para estimar as séries não observáveis da tendência e do componente cíclico, o modelo minimiza a seguinte equação com respeito a g_t :

$$\text{Min}_{\{g_t\}_{t=1}^T} \left\{ \sum_{t=1}^T c_t^2 + \lambda \sum_{t=1}^T [(g_t - g_{t-1}) - (g_{t-1} - g_{t-2})]^2 \right\}$$

Onde $c_t = y_t - g_t$ e λ é o parâmetro exógeno para determinar a minimização da função.

O primeiro termo penaliza o desvio da série em relação a tendência, já o segundo termo, penaliza a variação da tendência.

Filtro Hodrick-Prescott

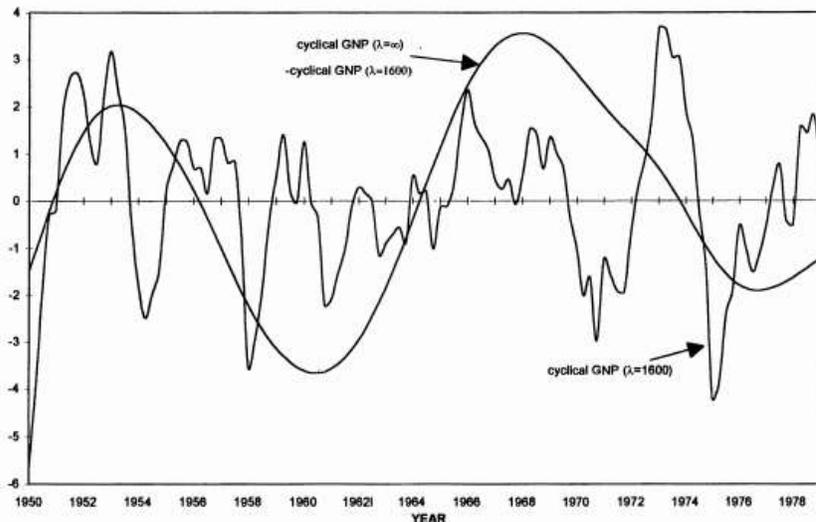
Como determinar o λ ?

Se o crescimento cíclico e a segunda diferença da tendência tiverem média zero e com desvio padrão σ_1^2 e σ_2^2 , então o valor é dado por:

$$\sqrt{\lambda} = \sigma_1^2 / \sigma_2^2$$

- ▶ Pressupõe distribuição normal e que seja iid

Os autores sugerem $\lambda=1600$ para o ciclo dos dados trimestrais EUA



Fonte: HP 1997

Filtro Hodrick-Prescott

Principais desvantagens:

1. Filtro HP pode produzir dinâmicas espúrias
2. É um filtro para ser analisado historicamente, novos dados geram mudanças de tendência e ciclo
3. Problema de final da amostra
4. Formalização estatística dos modelos econômicos, normalmente produz parâmetros de suavização significativamente diferentes do que normalmente são usados $\lambda=1600$

Filtro Hodrick-Prescott - aplicação

Instalando o pacote - tem que :

```
install.packages("mFilter")
```

```
require(mFilter)
```

Estimando o filtro HP para as vendas no varejo

```
hp=hpfilter(varejo,type="lambda",freq = 1600)
```

```
plot(hp$trend)
```

```
plot(hp$cycle)
```

Decomposição de Beveridge-Nelson

- ▶ Instrumento para obter o componente permanente e transitório de uma série.
- ▶ Importante para entender o efeito de curto prazo dos ciclos econômicos e inferir a tendência de longo prazo da série analisada.
- ▶ Por exemplo, suponha que haja uma queda do PIB; então, seria importante saber se essa queda é primordialmente devida a fatores cíclicos ou trata-se de uma tendência de longo prazo.

Decomposição de Beveridge-Nelson

Beveridge e Nelson decompõem a expressão em duas parcelas, admitindo-se por simplicidade que $\varepsilon_0 = 0$:

$$y_t = p_t + c_t;$$

$$p_t = y_0 + \delta t + \psi(1) \sum_{j=1}^t \varepsilon_j;$$

$$c_t = \psi^*(L) \varepsilon_t.$$

Da segunda equação, pode-se subtrair p_{t-1} para obter:

$$p_t - p_{t-1} = \delta + \psi(1) \varepsilon_t,$$

$$p_0 = y_0.$$

em que

$$\psi(1) = \sum_{j=0}^{\infty} \psi_j;$$

$$\psi^*(L) = \sum_{k=0}^{\infty} \psi_k^* L^k$$

Decomposição de Beveridge-Nelson

- ▶ p_t representa o componente permanente, cuja inovação tem variância $\psi(1)^2 \sigma^2$.
- ▶ $\psi(1)$ pode ser maior ou menor do que 1.
- ▶ c_t representa o componente cíclico ou transitório do modelo, que será estacionário se $\sum_{j=0}^{\infty} |\psi_j| < \infty$. Essa hipótese é mais forte que $\sum_{j=0}^{\infty} |\psi_j| < \infty$, e necessária para demonstrar que $\sum_{k=0}^{\infty} |\psi_j^*| < \infty$.

Decomposição de Beveridge-Nelson - Exemplo

Considere um modelo *ARIMA* (0, 1, 2):

$$y_t = y_{t-1} + \delta + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}.$$

É claro que $\psi_1 = \theta_1$, $\psi_2 = \theta_2$ e $\psi_j = 0, j > 2$. Então:

$$\psi(1) = 1 + \theta_1 + \theta_2; \quad \psi_0^* = - \sum_{j=1}^{\infty} \psi_j = -(\theta_1 + \theta_2);$$

$$\psi_1^* = - \sum_{j=2}^{\infty} \psi_j = -\theta_2; \quad \psi_k^* = - \sum_{j=k+1}^{\infty} \psi_j = 0, k > 1.$$

Logo, tem-se que:

$$p_t = p_{t-1} + \delta + (1 + \theta_1 + \theta_2) \varepsilon_t;$$

$$c_t = -(\theta_1 + \theta_2) \varepsilon_t - \theta_2 \varepsilon_{t-1}.$$