Séries Temporais: Trabalho 2

Mariana Costa Freitas

O que são os testes de Dickey Fuller e Phillip Perron?

Uma série temporal tem uma raiz unitária se, em sua equação característica, temos uma raiz igual a 1, indicando que a série é não estacionária.

Os testes de Dickey Fuller e Phillip Perron são utilizados para verificar a presença de uma raíz unitária em uma série temporal, ou seja, se uma série é ou não estacionária (garante propriedades como média e variância constantes ao longo do tempo). Isso é importante para a modelagem e previsão de séries temporais, pois muitas técnicas assumem que a série é estacionária.

Teste de Dickey Fuller

O teste de Dickey-Fuller verifica a hipótese nula de que uma série temporal possui uma raiz unitária, ou seja, não é estacionária, contra a hipótese alternativa de que a série é estacionária. Para isso, os seguintes passos são seguidos:

- Estimar a regressão: ajustar um modelo de regressão para a série.
- Calcular a estatística do teste: calcular a estatística $DF = \frac{\hat{\gamma}}{SE(\hat{\gamma})}$ ($\hat{\gamma}$ é é a estimativa do coeficiente associado a y_{t-1} no modelo) com base no coeficiente estimado e seu erro padrão.
- Comparar com valores críticos: comparar a estatística DF com valores críticos tabelados para determinar se rejeitamos ou não a hipótese nula.

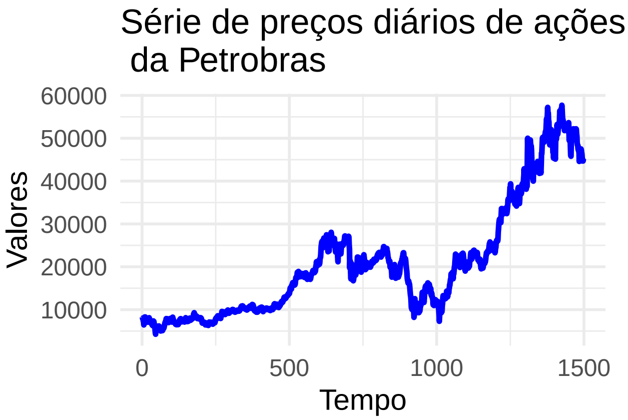
Teste de Phillip Perron

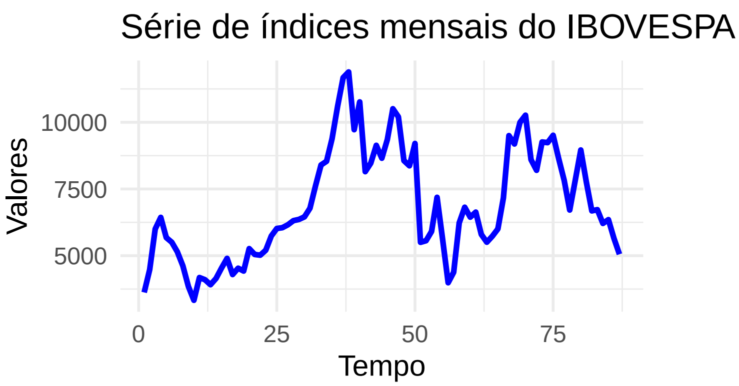
O teste de Phillips Perron é bem parecido com o teste anterior, porém faz ajustes para lidar com autocorrelação e heterocedasticidade.

- Estimar a regressão: ajustar um modelo de regressão para a série.
- Ajustar para autocorrelação e heterocedasticidade: o teste de Phillips-Perron corrige automaticamente a presença de autocorrelação e heterocedasticidade nos resíduos do modelo.
- Calcular a a estatística de teste de Phillips-Perron $Z(t) = \frac{\hat{\gamma}}{SE(\hat{\gamma})}\sqrt{\frac{\hat{\sigma^2}}{\hat{\sigma_u^2}}}$ (σ^2 é a variância dos resíduos do modelo e σ_u^2 é a variância de longo prazo dos resíduos, ajustada para autocorrelação e heterocedasticidade).
- Comparar com valores críticos: comparar a estatística DF com valores críticos tabelados para determinar se rejeitamos ou não a hipótese nula.

Aplicações

Vamos aplicar esses testes a três séries: preços diários de ações da Petrobras, preços diários de ações do Banespa e índices mensais do IBOVESPA, plotados abaixo.

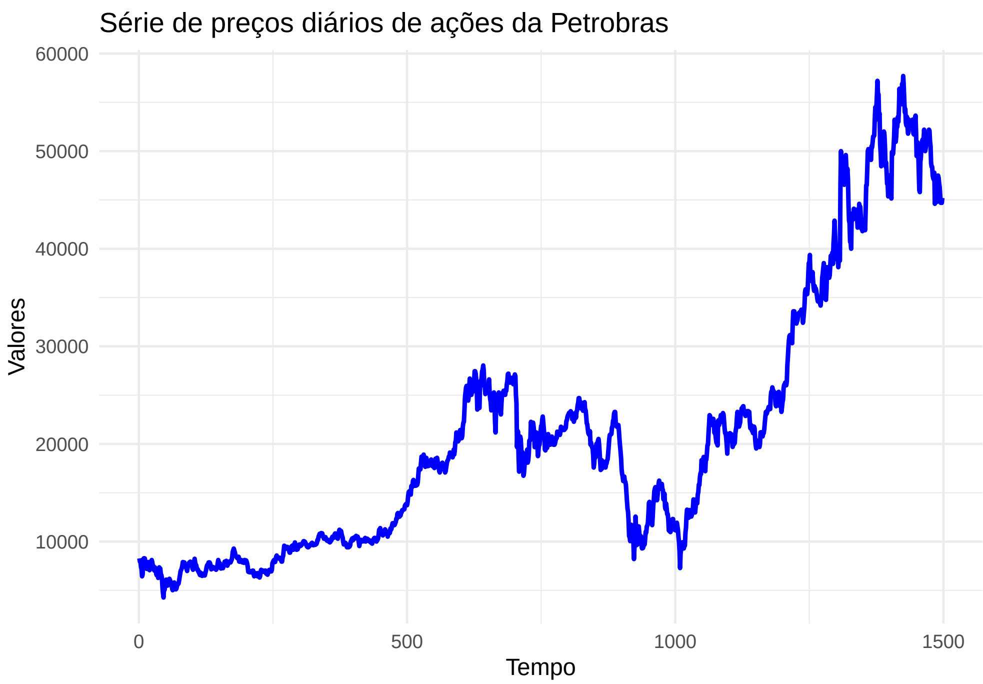




Série de preços diários de ações da Petrobras

```r
library(httr)
library(dplyr)
library(tseries)
library(urca)
url <- "https://www.ime.usp.br/~pam/D-PETRO"
dados <- GET(url)
petro <- read.csv(text = content(dados, "text")) |>
  rename(precos = X8780.295)

ggplot(petro, aes(x = 1:length(petro$precos), y = petro$precos)) +
  geom_line(color = "blue", size = 1) +
  labs(title = "Série de preços diários de ações da Petrobras", x = "Tempo", y = "Valores") +
  theme_minimal()
```



Para fazer o teste de Dickey-Fuller, usamos `type = 'trend'"`, pois a série apresenta intercepto e também parece apresentar tendência. Também foi usado `selecLags = "AIC"` para que o número de lags usado para corrigir a autocorrelação dos resíduos seja escolhido automaticamente com base no critério de AIC.

```
testedf <- urca::ur.df(y = petro$precos, type = "trend", selectlags = "AIC")
summary(testedf)
```

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4416.8  -263.3   -12.7   271.8  7723.9
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  8.387481  35.795304   0.234   0.8148
## z.lag.1     -0.005121   0.002436  -2.102   0.0357 *
## tt           0.160140   0.075403   2.124   0.0339 *
## z.diff.lag   0.119153   0.025703   4.636 3.87e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 688.3 on 1492 degrees of freedom
## Multiple R-squared:  0.01689,    Adjusted R-squared:  0.01492
## F-statistic: 8.546 on 3 and 1492 DF,  p-value: 1.256e-05
##
##
## Value of test-statistic is: -2.102 2.1316 2.4304
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau3 -3.96 -3.41 -3.12
## phi2  6.09  4.68  4.03
## phi3  8.27  6.25  5.34
```

No output, primeiro temos as informações sobre o modelo que foi usado para o teste, como o intercepto, coeficiente associado ao lag 1 (z.lag.1), o coeficiente da tendência linear (tt) e o coeficiente do lag da primeira diferença (z.diff.lag).

Obtemos que a estatística teste é -2.102 e o valor crítico para um nível de significância de 5% é -3.41. Como a estatística teste é maior que o valor crítico, aceitamos a hipótese nula e concluímos que a série não é estacionária.

A seguir, vamos verificar se o teste de Phillips Perron apresenta os mesmos resultados. Vamos usar `type = "Z-tau"` para definir a estatística teste usada e `model = trend` para informar que o modelo apresenta tendência.

```r
testepp <- urca::ur.pp(petro$precos, type = "Z-tau", model = "trend")
summary(testepp)
```

```
##
## ##################################
## # Phillips-Perron Unit Root Test #
## ##################################
##
## Test regression with intercept and trend
##
##
## Call:
## lm(formula = y ~ y.l1 + trend)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4521.5  -265.3   -11.3   262.5  7695.8
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.208e+02  5.390e+01   2.241   0.0252 *
## y.l1        9.954e-01  2.449e-03 406.434   <2e-16 ***
## trend       1.519e-01  7.579e-02   2.005   0.0452 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 692.8 on 1494 degrees of freedom
## Multiple R-squared:  0.9973, Adjusted R-squared:  0.9973
## F-statistic: 2.786e+05 on 2 and 1494 DF,  p-value: < 2.2e-16
```

```
##
##
## Value of test-statistic, type: Z-tau  is: -1.945
##
##            aux. Z statistics
## Z-tau-mu             2.0727
## Z-tau-beta           2.0373
##
## Critical values for Z statistics:
##                     1pct      5pct     10pct
## critical values -3.969401 -3.415306 -3.129519
```

Primeiro são retornados informações do modelo usado, como o intercepto, o coeficiente associado ao primeiro lag e a tendência. Abaixo, temos que o valor da estatística foi -1.9529 e o valor crítico para um nível de significância de 5% foi -3.415306. Como a estatística teste é maior que o valor crítico, aceitamos a hipótese nula, ou seja, há uma raíz unitária e a série não é estacionária.

Série de preços diários de ações do Banespa

```r
library(httr)
library(dplyr)
library(tseries)

url <- "https://www.ime.usp.br/~pam/D-BANESPA"
dados <- GET(url)
banespa <- read.csv(text = content(dados, "text")) |>
  rename(preco = X7.438023)

ggplot(banespa, aes(x = 1:length(banespa$preco), y = banespa$preco)) +
  geom_line(color = "blue", size = 1) +
  labs(title = "Série de preços diários de ações do Banespa", x = "Tempo", y = "Valores") +
  theme_minimal()
```

Assim como anteriormente, usamos `type = 'trend'"`, pois a série apresenta intercepto e também parece apresentar tendência e `selecLags = "AIC"` para que o número de lags usado para corrigir a autocorrelação dos resíduos seja escolhido automaticamente com base no critério de AIC.

```r
testedf <- urca::ur.df(y = banespa$preco, type = "trend", selectlags = "AIC")
summary(testedf)
```

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.7181 -0.3809 -0.0873  0.3142 11.4398
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.0341826  0.0845592  -0.404  0.68609
## z.lag.1     -0.0083196  0.0031644  -2.629  0.00865 **
## tt           0.0004594  0.0001982   2.318  0.02059 *
## z.diff.lag   0.1035742  0.0257486   4.023 6.05e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 1.555 on 1492 degrees of freedom
## Multiple R-squared:  0.01446,    Adjusted R-squared:  0.01248
## F-statistic: 7.295 on 3 and 1492 DF,  p-value: 7.407e-05
##
##
## Value of test-statistic is: -2.6291 2.5306 3.4561
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau3 -3.96 -3.41 -3.12
## phi2  6.09  4.68  4.03
## phi3  8.27  6.25  5.34
```

No output, primeiro temos as informações sobre o modelo que foi usado para o teste, como o intercepto, coeficiente associado ao lag 1 (z.lag.1), o coeficiente da tendência linear (tt) e o coeficiente do lag da primeira diferença (z.diff.lag).

Obtemos que a estatística teste é -2.6291 e o valor crítico para um nível de significância de 5% é -3.41. Como a estatística teste é maior que o valor crítico, aceitamos a hipótese nula e concluímos que a série não é estacionária.

Agora, vamos usar o teste de Phillips Perron. Novamente, usamos `type = "Z-tau"` para definir a estatística teste usada e `model = trend` para informar que o modelo apresenta tendência.

```r
testepp <- urca::ur.pp(banespa$preco, type = "Z-tau", model = "trend")
summary(testepp)
```

```
##
## ##################################
## # Phillips-Perron Unit Root Test #
## ##################################
##
## Test regression with intercept and trend
##
##
## Call:
## lm(formula = y ~ y.l1 + trend)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.0655 -0.3570 -0.0849  0.3312 11.4384
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.2885277  0.1130007   2.553   0.0108 *
## y.l1        0.9924193  0.0031723 312.842   <2e-16 ***
## trend       0.0004205  0.0001987   2.117   0.0344 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.562 on 1494 degrees of freedom
## Multiple R-squared:  0.9967, Adjusted R-squared:  0.9967
## F-statistic: 2.238e+05 on 2 and 1494 DF,  p-value: < 2.2e-16
##
```

```
##
## Value of test-statistic, type: Z-tau  is: -2.5502
##
##             aux. Z statistics
## Z-tau-mu                2.0980
## Z-tau-beta              2.2585
##
## Critical values for Z statistics:
##                        1pct      5pct     10pct
## critical values -3.969401 -3.415306 -3.129519
```

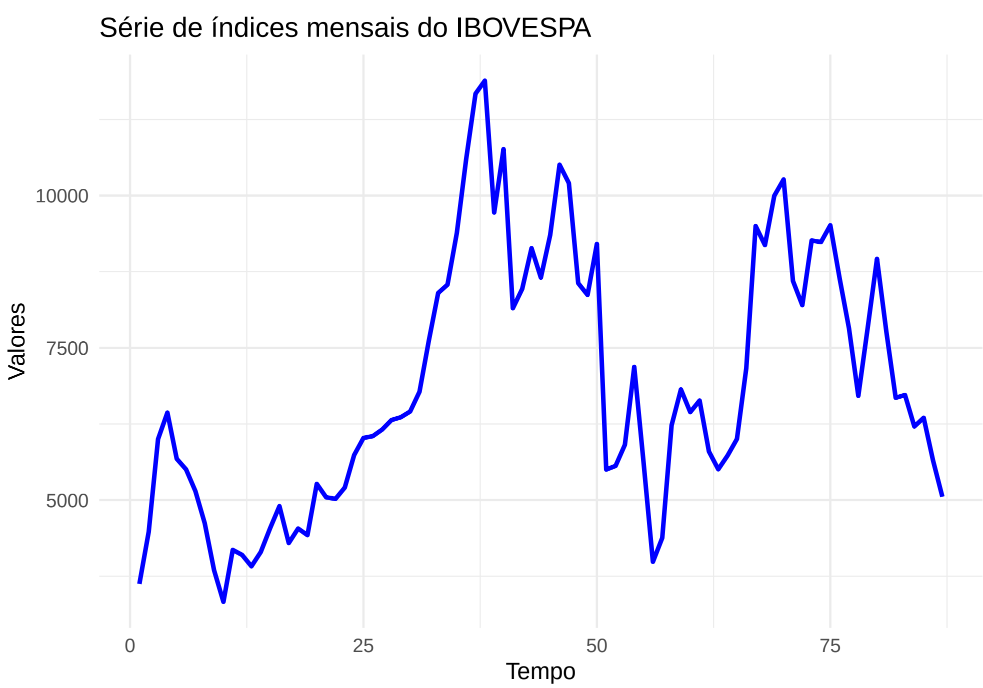
Primeiro são retornados informações do modelo usado, como o intercepto, o coeficiente associado ao primeiro lag e a tendência. Em seguida, temos que o valor da estatística foi -2.5502 e o valor crítico para um nível de significância de 5% foi -3.415306.

Como a estatística teste é maior que o valor crítico, aceitamos a hipótese nula, ou seja, há uma raíz unitária e a série não é estacionária.

Série de índices mensais do IBOVESPA

```r
url <- "https://www.ime.usp.br/~pam/M-IBV-SP"
dados <- GET(url)
ibov <- read.table(text = content(dados, "text"), header = T) #/>
  #rename(preco = X7.438023)

ggplot(ibov, aes(x = 1:length(ibov$IBOV), y = ibov$IBOV)) +
  geom_line(color = "blue", size = 1) +
  labs(title = "Série de índices mensais do IBOVESPA", x = "Tempo", y = "Valores") +
  theme_minimal()
```



Dessa vez iremos usar `type = 'drift'"`, pois a série apresenta intercepto, mas não parece apresentar tendência e `selecLags = "AIC"` para que o número de lags usado para corrigir a autocorrelação dos resíduos seja escolhido automaticamente com base no critério de AIC.

```
testedf <- urca::ur.df(y = ibov$IBOV, type = "drift", selectlags = "AIC")
summary(testedf)
```

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3513.8  -523.4    39.3   555.3  2258.1
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 828.23428  365.09285   2.269   0.0259 *
## z.lag.1      -0.11838    0.05048  -2.345   0.0214 *
## z.diff.lag    0.08366    0.10891   0.768   0.4446
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 937.9 on 82 degrees of freedom
## Multiple R-squared:  0.06379,    Adjusted R-squared:  0.04095
## F-statistic: 2.794 on 2 and 82 DF,  p-value: 0.06704
##
##
## Value of test-statistic is: -2.345 2.7513
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau2 -3.51 -2.89 -2.58
## phi1  6.70  4.71  3.86
```

No output, primeiro temos as informações sobre o modelo que foi usado para o teste, como o intercepto, coeficiente associado ao lag 1 (z.lag.1) e o coeficiente do lag da primeira diferença (z.diff.lag).

Obtemos que a estatística teste é -2.345 e o valor crítico para um nível de significância de 5% é -2.89. Como a estatística teste é maior que o valor crítico, aceitamos a hipótese nula e concluímos que a série não é estacionária.

Mais uma vez, vamos executar o teste de Phillips Perron. Aqui usamos `type = "Z-tau"` para definir a estatística teste usada e `model = constant` para informar que o modelo apresenta intercepto, mas não tendência.

```r
testepp <- urca::ur.pp(ibov$IBOV, type = "Z-tau", model = "constant")
summary(testepp)
```

```
## 
## ################################## 
## # Phillips-Perron Unit Root Test # 
## ################################## 
## 
## Test regression with intercept 
## 
## 
## Call:
## lm(formula = y ~ y.l1)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -3459.3  -533.8    26.7   536.1  2347.2 
## 
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 810.38921  349.20537   2.321   0.0227 *  
## y.l1          0.88525    0.04835  18.309   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 931.5 on 84 degrees of freedom
## Multiple R-squared:  0.7996, Adjusted R-squared:  0.7972 
## F-statistic: 335.2 on 1 and 84 DF,  p-value: < 2.2e-16
## 
## 
## Value of test-statistic, type: Z-tau  is: -2.4127 
## 
##          aux. Z statistics
## Z-tau-mu            2.3575
## 
## Critical values for Z statistics: 
##                      1pct      5pct     10pct
## critical values -3.507211 -2.895068 -2.584427
```

Primeiro são retornados informações do modelo usado, como o intercepto e o coeficiente associado ao primeiro lag. Em seguida, temos que o valor da estatística foi -2.4127 e o valor crítico para um nível de significância de 5% foi -2.895068

Como a estatística teste é maior que o valor crítico, aceitamos a hipótese nula, ou seja, há uma raíz unitária e a série não é estacionária.