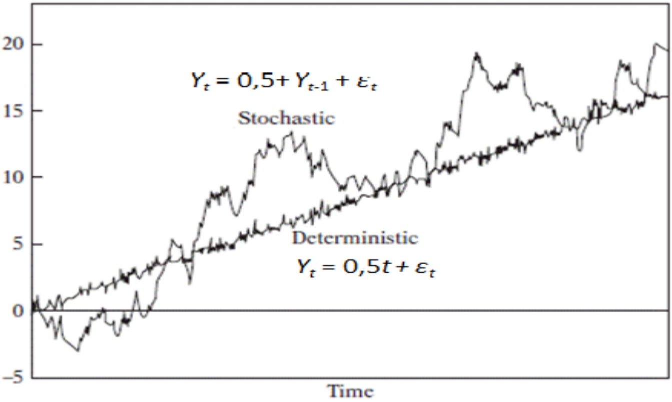
Processos Não Estacionários
Testes de Raiz Unitária

Bibliografia Básica:

- Enders, W. *Applied Econometric Time Series*. Cap. 4.
- Bueno, R. L. S. *Econometria de Séries Temporais*. Cap. 4.
- Box, G.E., Jenkins, G. M., Reinsel, G.C. & Ljung, G. M.(2016). *Time series analysis: forecasting and control*. Cap. 4.
- Morettin, P. A. *Análise de Séries Temporais.* Cap. 5.

Processos Não Estacionários

Processo não estacionário é um processo cuja média e/ou variância serão dependentes do tempo.



Processos Estocásticos Integrados

Processo estocástico integrado: processo estocástico que pode se tornar estacionário por meio da diferenciação.

Se um processo se torna estacionário ao se tomar sua primeira diferença dizemos que este processo é integrado de ordem um, ou I(1).

Se o processo somente se tornar estacionário na *d*-ésima diferença dizemos que este é um processo integrado de ordem d, ou I(d).

Um processo que não precisa ser diferenciado para se tornar estacionário é integrado de ordem zero, ou I(0).

Modelo *Random Walk* ou Passeio Aleatório

Um passeio aleatório é dado por:

$$Y_t = \delta + Y_{t-1} + \epsilon_t$$

em que $\epsilon_t \sim RB(0, \sigma^2_\epsilon)$. Se:

- $\delta \neq 0$: passeio aleatório com *drift* (ou **deslocamento** ou **tendência estocástica**). Neste caso, valor médio e variância são dependentes do tempo.
- $\delta = 0$: **tendência estocástica pura**. Neste caso, variância é dependente do tempo, ou seja,

$$E(Y_t) = E(Y_{t-1} + \epsilon_t) = \ldots = E(Y_0 + \sum_{t=1}^{T} \epsilon_t) = Y_0$$

  Temos que a esperança é constante e independe do tempo.

$$Var(Y_t) = E\{(Y_t - E(Y_t))^2\} = \ldots = \sum_{t=1}^{T} \sigma^2_\epsilon = T\sigma^2_\epsilon$$

  A variância é, portanto, uma função crescente do tempo, logo o passeio aleatório não é um processo estacionário.

Como identificar se a série é ou não estacionária?

1. Análise gráfica

- Inspeção visual raramente permite distingui-la como de tendência estocástica ou tendência determinística

2. Função de autocorrelação:

- Séries não estacionárias: fortes correlações seriais

3. Teste de Raiz Unitária.

Teste de Raiz Unitária ou Teste de Estacionariedade

Considere o modelo:

$$Y_t = \rho Y_{t-1} + \epsilon_t$$

Raiz unitária: Quando temos que $\rho = 1$ aparece o problema de raiz unitária que, basicamente, é sinônimo de não estacionaridade.

Passeio aleatório, com ou sem drift, pode se tornar estacionário ao ser diferenciado, isto é:

$$Y_t = Y_{t-1} + \epsilon_t \Leftrightarrow \Delta Y_t = \epsilon_t$$

Exemplo

Considere a série de taxa de câmbio $Real/US\$$:

a) Faça o gráfico da série e construa a função de autocorrelação;

b) Tome a primeira diferença e construa a função de autocorrelação.

Testes de Raiz Unitária

Testes que abordaremos:

1. Testes Dickey Fuller
2. Teste Dickey Fuller Aumentado – ADF
3. Teste de Phillips Perron
4. Teste de KPSS
5. Teste de Dickey e Pantula
6. Teste com Quebra Estrutural - Quebra conhecida
7. Teste com Quebra Estrutural Desconhecida
8. Raízes Unitárias Sazonais

1. Teste Dickey Fuller – DF

Objetivo:

Testar a existência de 1 RU em Y_t quando o processo gerador da série for expresso por uma das expressões abaixo:

$$
\begin{array}{llll}
(1) & Y_t = \alpha + \beta t + \rho Y_{t-1} + \epsilon_t & \Rightarrow & \Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \epsilon_t \\
(2) & Y_t = \alpha + \rho Y_{t-1} + \epsilon_t & \Rightarrow & \Delta Y_t = \alpha + \gamma Y_{t-1} + \epsilon_t \\
(3) & Y_t = \rho Y_{t-1} + \epsilon_t & \Rightarrow & \Delta Y_t = \gamma Y_{t-1} + \epsilon_t
\end{array}
$$

em que α e βt são componentes determinísticos, denominados constante ou *drift* e tendência linear, respectivamente; ϵ_t é um ruído branco.

Teste DF

Hipóteses nula e alternativa:

$$H_0: \quad \rho = 1 \quad \Leftrightarrow \quad \gamma = 0 \quad \text{(1 RU)}$$

$$H_1: \quad \rho < 1 \quad \Leftrightarrow \quad \gamma < 0 \quad \text{(0 RU)}$$

- Teste Monocaudal à esquerda

- Estatísticas dos testes são chamadas de:

1. Modelo com constante e tendência determinística: τ_τ
2. Modelo com constante: τ_μ
3. Modelo sem termos determinísticos: τ

Teste DF

Sob a hipótese nula, a distribuição do teste não é convencional, ou seja, não é igual à distribuição t, dado que Y_t não é estacionário.

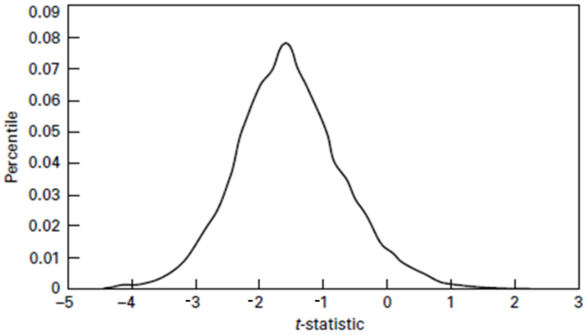


FIGURE 4.7 The Dickey–Fuller Distribution

Dickey & Fuller (1979) recalcularam os valores críticos: Tabela A, p. 488, Enders (2010)

Empirical Cumulative Distribution of τ

Probability of a Smaller Value Sample Size	0.01	0.025	0.05	0.10
No Constant or Time ($a_0 = a_2 = 0$)				τ
25	−2.66	−2.26	−1.95	−1.60
50	−2.62	−2.25	−1.95	−1.61
100	−2.60	−2.24	−1.95	−1.61
250	−2.58	−2.23	−1.95	−1.62
300	−2.58	−2.23	−1.95	−1.62
∞	−2.58	−2.23	−1.95	−1.62
Constant ($a_2 = 0$)				τ_μ
25	−3.75	−3.33	−3.00	−2.62
50	−3.58	−3.22	−2.93	−2.60
100	−3.51	−3.17	−2.89	−2.58
250	−3.46	−3.14	−2.88	−2.57
500	−3.44	−3.13	−2.87	−2.57
∞	−3.43	−3.12	−2.86	−2.57
Constant + time				τ_τ
25	−4.38	−3.95	−3.60	−3.24
50	−4.15	−3.80	−3.50	−3.18
100	−4.04	−3.73	−3.45	−3.15
250	−3.99	−3.69	−3.43	−3.13
500	−3.98	−3.68	−3.42	−3.13
∞	−3.96	−3.66	−3.41	−3.12

Teste DF no R

Para realizar o teste de Dickey-Fuller, carregue o pacote `urca`.

Em seguida use a função `UR.DF`, a qual retorna, automaticamente, os valores críticos para o teste.

No R as estatísticas são denotadas por:

$$\tau_\tau = \tau_3$$

$$\tau_\mu = \tau_2$$

$$\tau = \tau_1$$

Significância dos Termos Determinísticos Individuais

EMPIRICAL DISTRIBUTION OF

Sample size n	0.90	0.95	0.975	0.99
	$\hat{\tau}_{\alpha\tau}$ FOR $(\alpha, \beta, \rho) = (0, 0, 1)$ IN $Y_t = \alpha + \beta t + \rho Y_{t-1} + e_t$.			
25	2.77	3.20	3.59	4.05
50	2.75	3.14	3.47	3.87
100	2.73	3.11	3.42	3.78
250	2.73	3.09	3.39	3.74
500	2.72	3.08	3.38	3.72
∞	2.72	3.08	3.38	3.71
	$\hat{\tau}_{\beta\tau}$ FOR $(\alpha, \beta, \rho) = (0, 0, 1)$ IN $Y_t = \alpha + \beta t + \rho Y_{t-1} + e_t$			
25	2.39	2.85	3.25	3.74
50	2.38	2.81	3.18	3.60
100	2.38	2.79	3.14	3.53
250	2.38	2.79	3.12	3.49
500	2.38	2.78	3.11	3.48
∞	2.38	2.78	3.11	3.46
	$\hat{\tau}_{\alpha\mu}$ FOR $(\alpha, \rho) = (0, 1)$ IN $Y_t = \alpha + \rho Y_{t-1} + e_t$.			
25	2.20	2.61	2.97	3.41
50	2.18	2.56	2.89	3.28
100	2.17	2.54	2.86	3.22
250	2.16	2.53	2.84	3.19
500	2.16	2.52	2.83	3.18
∞	2.16	2.52	2.83	3.18

Significância dos Termos Determinísticos

Análise da significância dos elementos determinísticos (constante e tendência linear) por meio de testes de hipóteses conjuntos, utilizando valores críticos simulados por DF.
Dickey & Fuller (1981) sugerem as estatísticas F denominadas ϕ_1, ϕ_2 e ϕ_3 para testar hipóteses conjuntas:

$$H_0 : \gamma = \alpha = 0 \Rightarrow \Phi_1$$

$$H_0 : \gamma = \alpha = \beta = 0 \Rightarrow \Phi_2$$

$$H_0 : \gamma = \beta = 0 \Rightarrow \Phi_3$$

As estatísticas relacionadas a essas hipóteses são:

$$\Phi_i = \frac{(SQRes_{restrita} - SQRes_{irrestrita})/r}{SQRes_{irrestrita}/(T-k)} \quad i = 1, 3$$

sendo r o número de restrições, T número de observações, k número de parâmetros estimados no modelo irrestrito.

EMPIRICAL DISTRIBUTION OF

Sample size n	0.90	0.95	0.975	0.99
Φ_1 FOR $(\alpha, \rho) = (0, 1)$ IN $Y_t = \alpha + \rho Y_{t-1} + e_t$				
25	4.12	5.18	6.30	7.88
50	3.94	4.86	5.80	7.06
100	3.86	4.71	5.57	6.70
250	3.81	4.63	5.45	6.52
500	3.79	4.61	5.41	6.47
∞	3.78	4.59	5.38	6.43
Φ_2 FOR $(\alpha, \beta, \rho) = (0, 0, 1)$ IN $Y_t = \alpha + \beta t + \rho Y_{t-1} + e_t$				
25	4.67	5.68	6.75	8.21
50	4.31	5.13	5.94	7.02
100	4.16	4.88	5.59	6.50
250	4.07	4.75	5.40	6.22
500	4.05	4.71	5.35	6.15
∞	4.03	4.68	5.31	6.09
Φ_3 FOR $(\alpha, \beta, \rho) = (\alpha, 0, 1)$ IN $Y_t = \alpha + \beta t + \rho Y_{t-1} + e_t$				
25	5.91	7.24	8.65	10.61
50	5.61	6.73	7.81	9.31
100	5.47	6.49	7.44	8.73
250	5.39	6.34	7.25	8.43
500	5.36	6.30	7.20	8.34
∞	5.34	6.25	7.16	8.27

2. Testes Dickey Fuller Aumentado – ADF

Resíduos $\hat{\epsilon}_t$ obtidos dos modelos (1) a (3) são ruídos brancos?
Teste ADF: adiciona as defasagens da variável dependente, ou seja, supõe-se que a série é gerada por um processo auto-regressivo de ordem p.

$$(4) \quad \Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum_{i=1}^{p-1} \delta_i \Delta Y_{t-i} + \epsilon_t$$

$$(5) \quad \Delta Y_t = \alpha + \gamma Y_{t-1} + \sum_{i=1}^{p-1} \delta_i \Delta Y_{t-i} + \epsilon_t$$

$$(6) \quad \Delta Y_t = \gamma Y_{t-1} + \sum_{i=1}^{p-1} \delta_i \Delta Y_{t-i} + \epsilon_t$$

Inclusão de termos autorregressivos não altera a convergência das estatísticas τ, τ_μ e τ_τ. Portanto, usa-se nos testes ADF os mesmo valores críticos utilizados nos testes DF.

Teste ADF

Para identificar o número de atrasos p de forma que $\hat{\epsilon}_t \sim RB(0, \sigma^2_\epsilon)$, usa-se:

1) análise das autocorrelações dos resíduos do modelo sem termos de aumento;
2) critérios de informação: Akaike (AIC) ou Bayesian (BIC):

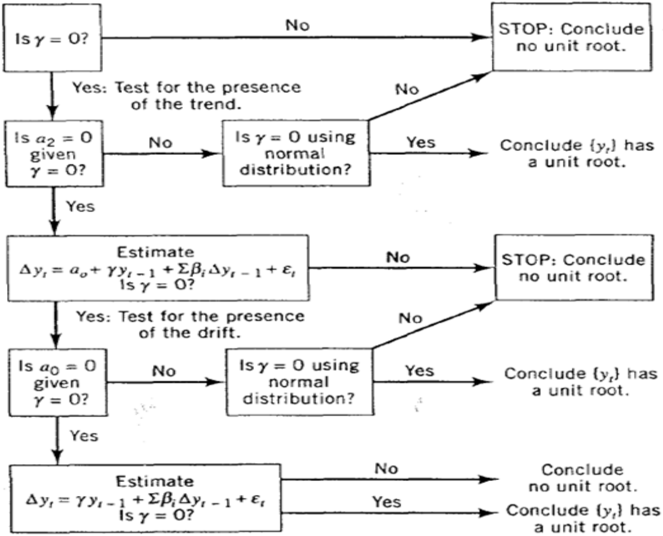
$$AIC(p) = ln(\sigma^2_\epsilon) + p\frac{2}{T}$$

$$BIC(p) = ln(\sigma^2_\epsilon) + p\frac{ln(T)}{T}$$

em que T é o número de observações e p número de parâmetros.

Figure 4.7 A procedure to test for unit roots.

Estimate $\Delta y_t = a_o + \gamma y_{t-1} + a_2 t + \Sigma\beta_i \Delta y_{t-i} + \varepsilon_t$



Exemplo

Verificar a existência de RU para:

1. IPCA dessazonalizada (Fonte: IPEADATA).
2. Taxa de Câmbio médio *Real*/*US*$ (Fonte: IPEADATA).

Poder do Teste de Dickey-Fuller

Erro Tipo I: probabilidade de rejeitar a hipótese nula quando esta hipótese é verdadeira.

Erro Tipo II: probabilidade de não rejeitar a hipótese nula quando a hipótese alternativa é verdadeira.

Poder do Teste: é calculado como 1 menos a probabilidade de se cometer um erro do tipo II, isto é, a probabilidade de se rejeitar a hipótese nula quando ela efetivamente é falsa.

Poder do Teste de Dickey-Fuller

Poder dos testes DF e ADF: em geral, é pequeno devido a alta persistência e/ou tendência presente nas variáveis macroeconômicas.

Consequência: muitas vezes tendemos a aceitar a existência de uma raíz unitária quando na verdade deveríamos rejeita-lá.

Teste DF assume que os resíduos são não correlacionados, sendo que uma maneira de garantir isso é a de incluir defasagens das diferenças da variável dependente na regressão de teste, isto é, aplicar o teste ADF.

Ao se incluir um número maior de regressores que não estão presentes no processo da série, somente agrava o problema.

Teste de Phillips-Perron (PP)

Phillips & Perron (1988) propuseram uma estatística de teste não paramétricas em que a hipótese nula de uma raiz unitária, explicitamente permite dependência fraca e heterogeneidade do processo de erro.

O teste é baseado nas respectivas estatísticas:

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \epsilon_t \quad \rightarrow \quad z_{\tau,\beta}$$

$$\Delta Y_t = \alpha + \gamma Y_{t-1} + \epsilon_t \quad \rightarrow \quad z_{\tau,\mu}$$

$$\Delta Y_t = \gamma Y_{t-1} + \epsilon_t \quad \rightarrow \quad z_{\tau}$$

Teste PP

Obtenção das estatísticas z_τ, $z_{\tau,\mu}$ $z_{\tau,\beta}$ depende do cálculo da variância de longo prazo dos resíduos:

$$\hat{\upsilon}^2 = \hat{\sigma}^2 + \frac{2}{T}\sum_{j=1}^{M}\omega\left(\frac{j}{M+1}\right)\sum_{t=j+1}^{T}\hat{\epsilon}_t\hat{\epsilon}_{t-j}$$

sendo $\hat{\sigma}^2$ variância populacional estimada; M tamanho do lag; T tamanho da amostra; $\omega\left(\frac{j}{M+1}\right)$ função de ponderação.

Teste PP

- Cálculo da função de ponderação: proposta por Bartlett, ou Parzen ou Quadrática;

- Perron (1990) recomenda o uso da janela de Parzen.

- Grande parte dos trabalhos empíricos emprega ponderação proposta por Bartlett.

- Determinação de M: critério de Newey-West (1994) ou Andrews (1991).

Teste PP

Phillips & Perron (1988) também definiram testes sobre os coeficientes do modelo, em vez de usar a estatística t, chamados de z_α. Neste caso, os testes são realizados sobre a distribuição dos coeficientes.

Do ponto de vista prático, não há diferença entre a análise associada às estatísticas ou sobre a distribuição.

Teste PP no R

Pacote `urca`

- Usa a função `ur.pp()`

```
ur.pp(x, type = c('Z-alpha', 'Z-tau'), model =
c('constant', 'trend'),
lags = c('short', 'long'), use.lag = NULL)
```

`Z-alpha` teste sobre os coeficientes do modelo e `Z-tau` teste sobre as estatísticas do modelo (na prática não há diferença);

`short`: truncamento é dado por $trunc(4 * (n/100)^{1/4})$;

`long`: truncamento é dado por $trunc(12 * (n/100)^{1/4})$;

`use.lag`: usuário fornece o número de defasagens;

Para correção do termo erro é usado a ponderação de Bartlett.

Exemplo

Verificar a existência de RU para:

1. IPCA dessazonalizada (Fonte: IPEADATA).
2. Taxa de Câmbio médio $Real/US\$$ (Fonte: IPEADATA).

Teste de KPSS

Kwiatkowski, Phillips, Schmidt & Shin (1992)[1] desenvolveram um teste de raiz unitária denominado KPSS.

Neste teste as hipóteses são formuladas por:

$$H_0 : \quad \{y_t\} \text{ é } I(0) \qquad \Rightarrow \quad \text{A série é estacionária}$$

$$H_1 : \quad \{y_t\} \text{ é } I(1) \qquad \Rightarrow \quad \text{A série é não estacionária}$$

Rejeitar H_0 significa que a série tem uma raiz unitária.

[1] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992), "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?", *Journal of Econometrics*, 54, 159–178.

Teste de KPSS

O teste KPSS é baseado no seguinte modelo:

$$y_t = \delta t + r_t + \epsilon_t \tag{1}$$

$$r_t = r_{t-1} + u_t \tag{2}$$

em que r_t é um passeio aleatório e ϵ_t é um ruído branco com média zero e variância constante. O valor r_0 é fixo e corresponde ao nível da série.

Se $\delta = 0$ o modelo é escrito como uma constante (μ) como termo determinístico.

Se $\delta \neq 0$ o modelo tem uma constante e uma tendência determinística.

Passos do Teste de KPSS

1. Estima-se o modelo representado pela equação (1)
2. Obtém-se os resíduos $\hat{\epsilon}_t$.
3. Calcula-se a soma parcial dos resíduos: $S_t = \sum_{j=1}^{t} \hat{\epsilon}_t$
4. A estatística do teste é dada por: $KPSS = \frac{\sum_{t=1}^{T} S_t^2}{\hat{\sigma}_\epsilon^2}$

   sendo T o tamanho da amostra, $\hat{\sigma}_\epsilon^2$ a estimativa da variância de ϵ:

$$\hat{\sigma}_\epsilon^2 = \frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}_t^2 \quad + \quad \underbrace{\frac{2}{T}\left(\sum_{j=1}^{M}\omega\left(\frac{j}{M+1}\right)\sum_{t=j+1}^{T}\hat{\epsilon}_t\hat{\epsilon}_{t-j}\right)}_{\textbf{correção da autocorrelação residual}}$$

Se KPSS for maior que valor crítico: rejeitar H_0.

Teste KPSS no R – Pacote `urca`

- Função `ur.kpss()`

  ```
  ur.kpss(y, type = c("mu", "tau"),
  lags = c("short", "long", "nil"), use.lag = NULL)
  ```

  em que `mu` indica o uso do termo constante e `tau` o emprego da constante e tendência.

  Para a escolha do número de termos para a correção da autocorrelação dos resíduos temos as opções:

a) `short` que representa incluir até a defasagem dada pelo valor inteiro de: $4 * \left(\frac{T}{100}\right)^{1/4}$;

b) `long` que representa incluir até a defasagem dada pelo valor inteiro de: $12 * \left(\frac{T}{100}\right)^{1/4}$;

c) `nil` significa não fazer nenhuma correção no termo de erro.

O termo `use.lag` permite que o próprio usuário fornece o número de defasagens para a correção da autocorrelação dos resíduos.

Exemplo

Verificar a existência de RU para:

1. IPCA dessazonalizada (Fonte: IPEADATA).
2. Taxa de Câmbio médio $Real/US\$$ (Fonte: IPEADATA).