# *Jednotkový koreň (unit root), diferencovanie časového radu, unit root testy*

Beáta Stehlíková

Časové rady, FMFI UK, 2012/2013

# *Obsah*

- Čo je jednotkový koreň (unit root) a čo spôsobuje
- Ak má proces jednotkový koreň - ako dáta transformovať, aby sme s nimi mohli pracovať a použiť ARMA metodológiu
- Ako z dát zistiť, či má proces jednotkový koreň → unit root testy

# *Príklady*

- Majme proces $y_t = y_{t-1} + u_t$:
    - je to nestacionárny AR(1) proces s jednotkovým koreňom
    - pre jeho diferencie $\Delta y_t = y_t - y_{t-1}$ platí $\Delta y_t = u_t$
    - teda $\Delta y_t$ je stacionárny proces
- Majme nestacionárny proces s jednotkovým koreňom

$$(1 - \frac{1}{2}L)(1 - L)x_t = 1 + (1 - \frac{1}{3}L)u_t$$

Potom pre diferencie

$$\Delta y_t = y_t - y_{t-1} = (1 - L)y_t$$

platí

$$(1 - \frac{1}{2}L)\Delta y_t = 1 + (1 - \frac{1}{3}L)u_t,$$

teda $\Delta y_t$ je stacionárny proces

# Príklady

- Majme nestacionárny proces s dvojnásobným jednotkovým koreňom
$$(1-\frac{1}{2}L)(1-L)^2 x_t = 1 + (1-\frac{1}{3}L)u_t$$
Potom pre druhé diferencie
$$\Delta^2 y_t = \Delta(\Delta y_t) = (1-L)(1-L)y_t = (1-L)^2 y_t$$
platí
$$(1-\frac{1}{2}L)\Delta^2 y_t = 1 + (1-\frac{1}{3}L)u_t,$$
teda $\Delta^2 y_t$ je stacionárny proces
- Vo všeobecnosti:
Ak má proces jednotkový koreň násobnosti $k$ (a ostatné mimo jednotkového kruhu), tak jeho k-te diferencie sú stacionárne

# *ARIMA modely*

- Ak treba proces k-krát diferencovať, aby sme dostali stacionárny proces, nazýva sa integrovaný proces rádu $k$, označujeme $I(k)$
- Ak tie k-te diferencie sú ARMA(p,q), tak o pôvodnom procese hovoríme, že je ARIMA(p,k,q).
- Napríklad $x_t$, ak

$$(1 - \frac{1}{2}L)(1 - L)^2 x_t = 1 + (1 - \frac{1}{3}L)u_t,$$

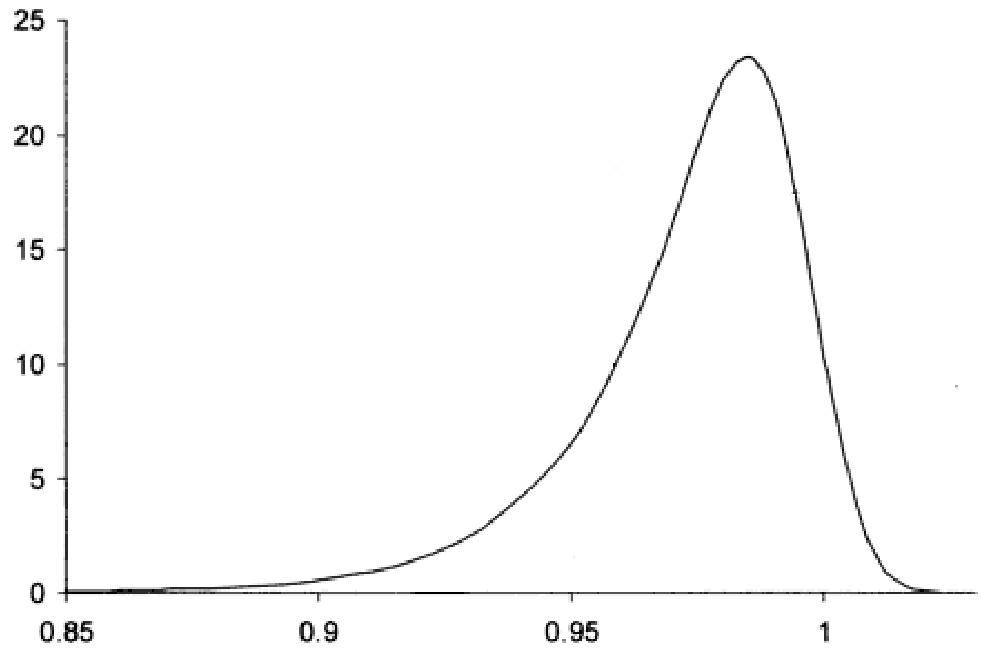
je proces ARIMA(1,2,1).

# *Jednotkový koreň a t-štatistika*

- Skúsme použiť testovanie hypotéz o koeficientoch regresného modelu známe z ekonometrie.
- [Kirchgässner, Wolters], example 5.1:
  - proces $y_t = y_{t-1} + u_t$, počet pozorovaní $T = 200$
  - vygenerujeme 100 000 realizácií
  - MNŠ odhadujeme model $y_t = \alpha + \rho y_{t-1} + u_t$ a zaznamenávame
    1. odhad $\hat{\rho}$
    2. hodnotu t-štatistiky zodpovedajúcej nulovej hypotéze $\rho = 1$ (ktorá platí)

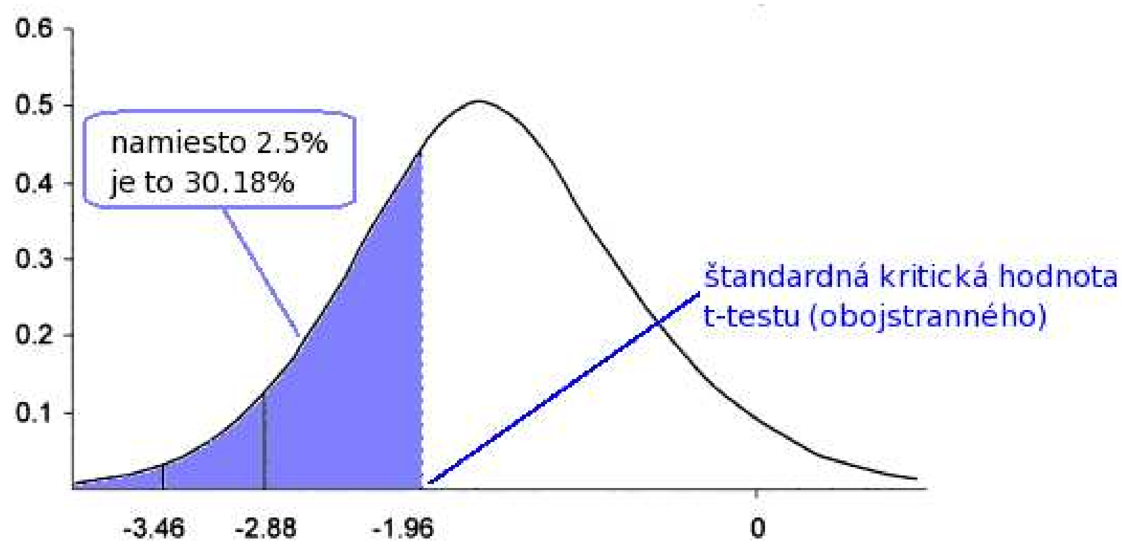
# *Jednotkový koreň a t-štatistika*

- Odhad $\hat{\rho}$ - mal by mať normálne rozdelenie so strednou hodnotou rovnou jeho skutočnej hodnote $\rho = 1$
- Rozdelenie odhadu zo simulácií je však nasledovné:



# *Jednotkový koreň a t-štatistika*

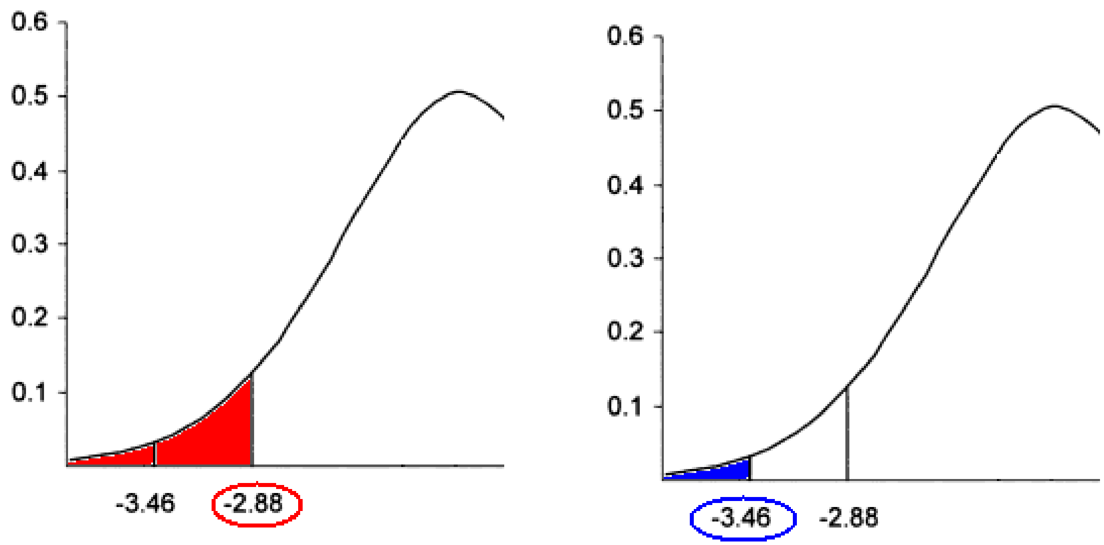
- Rozdelenie t-štatistiky zo simulácií:




- Štandardné kritické hodnoty sa nedajú použiť → použijeme t-štatistiku, ale s inými kritickými hodnotami

# *Jednotkový koreň a t-štatistika*

However, if we use the critical values of J.G. MACKINNON (1991), which, in this situation, are -2.876 at the 5 percent level and -3.465 at the 1 percent level, with rejection rates of 4.99 percent and 0.99 percent, the significance levels are almost exactly realised in our simulations.




- Kritické hodnoty:
  - ◇ od čoho závisia
  - ◇ ako ich vypočítať
  - ◇ čo ak nemáme AR(1) proces, ale všeobecnejší

# *Testovanie jednotkového koreňa*

- AR(1) proces:

  (1) $$y_t = \rho y_{t-1} + u_t$$

  jednotkový koreň znamená, že $\rho = 1$.
- Ekvivalentne:

  $$\Delta y_t = (\rho - 1) y_{t-1} + u_t$$

  a zaujíma nás t-štatistika zo signifikancie koeficienta pri $y_{t-1}$ - ale s inou kritickou hodnotou
- Tá kritická hodnota
  - závisí od počtu dát
  - zmení sa, ak rovnica (1) obsahuje konštantu a/alebo lineárny trend
- Vo všeobecnosti: $\Delta y_t = \alpha + \beta t + (\rho - 1) y_{t-1} + u_t$

# Testovanie jednotkového koreňa

- AR(p) proces: $y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots \alpha_p y_{t-p} + u_t$
  jednotkový koreň $\rightarrow$ $\alpha_1 + \ldots \alpha_p = 1.$
- Upravíme do tvaru:

  $$y_t = \rho y_{t-1} + \theta_1 \Delta y_{t-1} + \theta_2 \Delta y_{t-2} + \ldots + \theta_{p-1} \Delta y_{t-p+1} + u_t,$$

  kde $\rho = \sum_{j=1}^{p} \alpha_j$, $\theta_i = -\sum_{j=i+1}^{p} \alpha_j$ pre $i = 1, \ldots, p-1$
- Ekvivalentne:

  $$\Delta y_t = (\rho - 1) y_{t-1} + \theta_1 \Delta y_{t-1} + \theta_2 \Delta y_{t-2} + \ldots + \theta_{p-1} \Delta y_{t-p+1} + u_t$$

  a zaujíma nás t-štatistika z koeficienta pri $y_{t-1}$
- Vo všeobecnosti: $y$ môže mat' trend a/alebo intercept
  $\Rightarrow \Delta y_t = \alpha + \beta t + (\rho - 1) y_{t-1} + \theta_1 \Delta y_{t-1} + \theta_2 \Delta y_{t-2} + \ldots$
  $+ \theta_{p-1} \Delta y_{t-p+1} + u_t$

# *Augmented Dickey-Fuller test (ADF)*

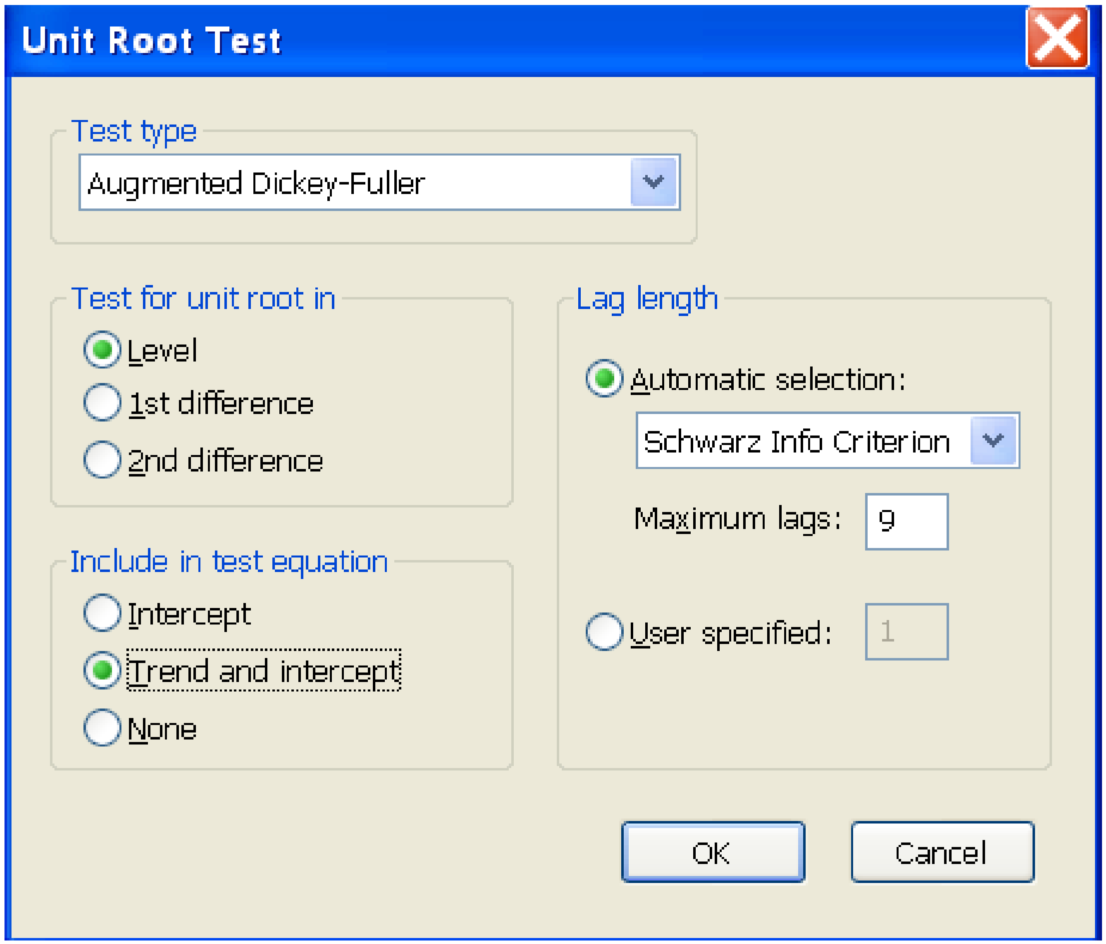
Wayne A. Fuller (1976)

David A. Dickey, Wayne A. Fuller (1979, 1981)

- Odhadujeme rovnicu
  $\Delta y_t = \alpha + \beta t + (\rho - 1) y_{t-1} + \theta_1 \Delta y_{t-1} + \ldots + \theta_k \Delta y_{t-k} + u_t$
  pričom musíme
  - rozhodnúť, či zahrnúť konštantu $\alpha$ a/alebo lineárny trend $\beta$ (podľa toho, či ich obsahuje proces $y$)
  - určiť $k$
- Zaujíma nás potom t-štatistika zo signifikancie koeficienta pri $y_{t-1}$, ale so správnymi kritickými hodnotami

# *Augmented Dickey-Fuller test (ADF)*

- ADF test v EViews:



# *ADF test - kritické hodnoty*

- James G. MacKinnon (1991) - dostupné ako súčasť doplnenej verzie z roku 2010:

  James G. MacKinnon: **Critical Values for Cointegration Tests**. Queen's Economics Department Working Paper No. 1227, 2010..

  Dostupné online: http://ideas.repec.org/p/qed/wpaper/1227.html

- Simulačne získané hodnoty:

**Table 1. Response Surface Estimates of Critical Values**

| $N$ | Variant | Level | Obs. | $\beta_\infty$ | (s.e.) | $\beta_1$ | $\beta_2$ |
|---|---|---|---|---|---|---|---|
| 1 | no constant | 1% | 600 | −2.5658 | (0.0023) | −1.960 | −10.04 |
| | | 5% | 600 | −1.9393 | (0.0008) | −0.398 | |
| | | 10% | 560 | −1.6156 | (0.0007) | −0.181 | |
| 1 | no trend | 1% | 600 | −3.4336 | (0.0024) | −5.999 | −29.25 |
| | | 5% | 600 | −2.8621 | (0.0011) | −2.738 | −8.36 |
| | | 10% | 600 | −2.5671 | (0.0009) | −1.438 | −4.48 |
| 1 | with trend | 1% | 600 | −3.9638 | (0.0019) | −8.353 | −47.44 |
| | | 5% | 600 | −3.4126 | (0.0012) | −4.039 | −17.83 |
| | | 10% | 600 | −3.1279 | (0.0009) | −2.418 | −7.58 |

# ADF test - kritické hodnoty

- Ak v regresii používame $T$ dát, kritická hodnota je $\beta_\infty + \beta_1/T + \beta_2/T^2$
- To sú tie hodnoty, ktoré sa spomínajú v príklade v knihe [Kirchgässner, Wolters]:

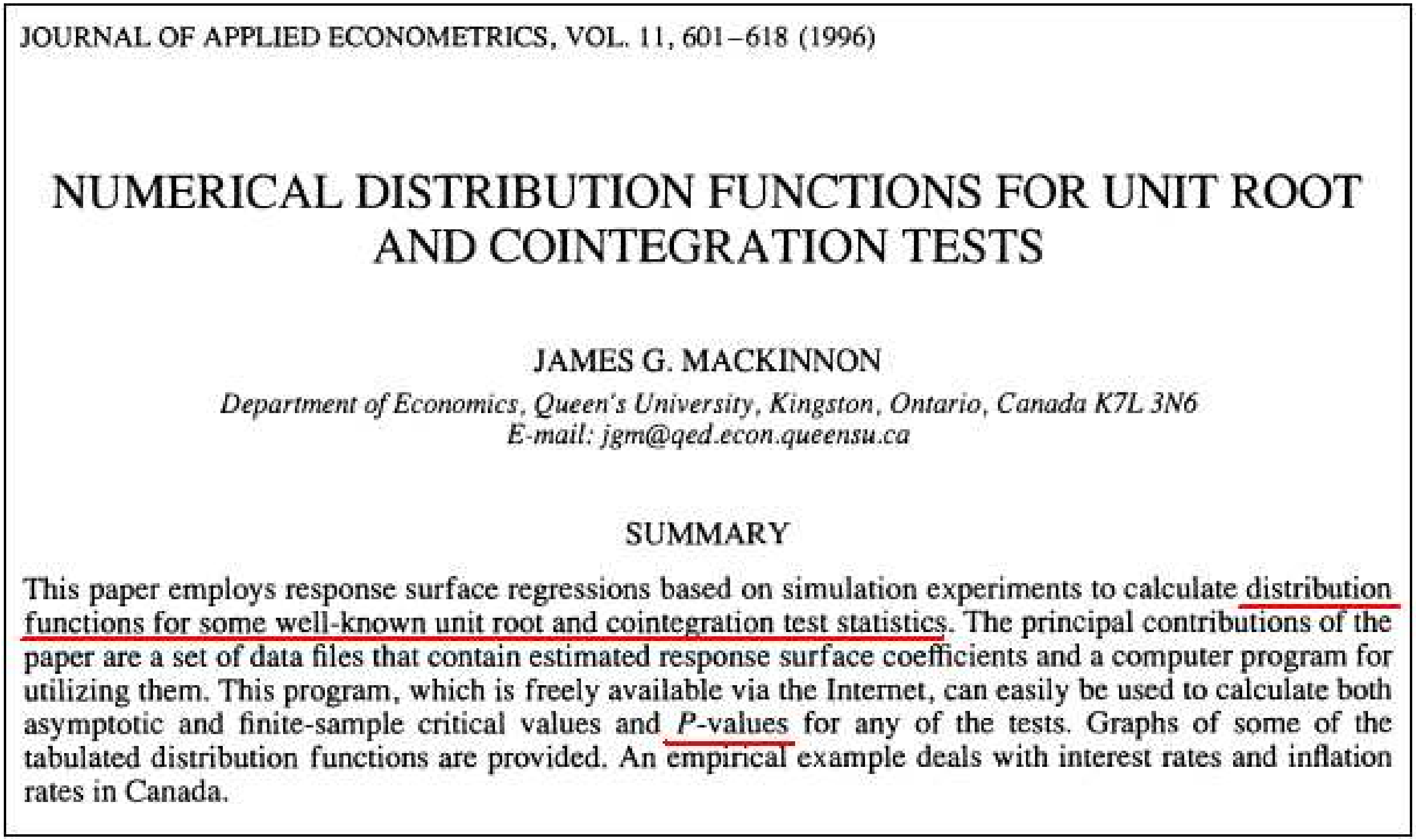
However, if we use the critical values of J.G. MACKINNON (1991) which, in this situation, are -2.876 at the 5 percent level and -3.465 at the 1 percent level, with rejection rates of 4.99 percent and 0.99 percent, the significance levels are almost exactly realised in our simulations.

**Table 1. Response Surface Estimates of Critical Values**

| $N$ | Variant | Level | Obs. | $\beta_\infty$ | (s.e.) | $\beta_1$ | $\beta_2$ |
|---|---|---|---|---|---|---|---|
| 1 | no constant | 1% | 600 | −2.5658 | (0.0023) | −1.960 | −10.04 |
| | | 5% | 600 | −1.9393 | (0.0008) | −0.398 | |
| | | 10% | 560 | −1.6156 | (0.0007) | −0.181 | |
| 1 | no trend | 1% | 600 | −3.4336 | (0.0024) | −5.999 | −29.25 |
| | | 5% | 600 | −2.8621 | (0.0011) | −2.738 | −8.36 |
| | | 10% | 600 | −2.5671 | (0.0009) | −1.438 | −4.48 |
| 1 | with trend | 1% | 600 | −3.9638 | (0.0019) | −8.353 | −47.44 |
| | | 5% | 600 | −3.4126 | (0.0012) | −4.039 | −17.83 |
| | | 10% | 600 | −3.1279 | (0.0009) | −2.418 | −7.58 |

# *ADF test - kritické hodnoty*

- James G. MacKinnon (1996) - nielen kritické hodnoty, ale aj P hodnoty pre konkrétnu hodnotu testovacej štatistiky:



JOURNAL OF APPLIED ECONOMETRICS, VOL. 11, 601–618 (1996)

NUMERICAL DISTRIBUTION FUNCTIONS FOR UNIT ROOT AND COINTEGRATION TESTS

JAMES G. MACKINNON

*Department of Economics, Queen's University, Kingston, Ontario, Canada K7L 3N6*
*E-mail: jgm@qed.econ.queensu.ca*

SUMMARY

This paper employs response surface regressions based on simulation experiments to calculate distribution functions for some well-known unit root and cointegration test statistics. The principal contributions of the paper are a set of data files that contain estimated response surface coefficients and a computer program for utilizing them. This program, which is freely available via the Internet, can easily be used to calculate both asymptotic and finite-sample critical values and *P*-values for any of the tests. Graphs of some of the tabulated distribution functions are provided. An empirical example deals with interest rates and inflation rates in Canada.

Program: http://econ.queensu.ca/faculty/mackinnon/numdist/

# ADF test - ukážka použitia

Null Hypothesis: LOGPCOCOA has a unit root
Exogenous: None
Lag Length: 2 (Automatic based on SIC, MAXLAG=18)

| | | t-Statistic | Prob.* |
|---|---|---|---|
| Augmented Dickey-Fuller test statistic | | 0.481897 | 0.8188 |
| Test critical values: | 1% level | -2.569504 | |
| | 5% level | -1.941445 | |
| | 10% level | -1.616282 | |

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LOGPCOCOA)
Method: Least Squares
Date: 10/18/11 Time: 12:46
Sample (adjusted): 1960M04 2002M09
Included observations: 510 after adjustments

| | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| LOGPCOCOA(-1) | 0.000332 | 0.000689 | 0.481897 | 0.6301 |
| D(LOGPCOCOA(-1)) | 0.352113 | 0.044181 | 7.969833 | 0.0000 |
| D(LOGPCOCOA(-2)) | -0.117298 | 0.044191 | -2.654332 | 0.0082 |

| | | | |
|---|---|---|---|
| R-squared | 0.110808 | Mean dependent var | 0.002585 |
| Adjusted R-squared | 0.107301 | S.D. dependent var | 0.066382 |
| S.E. of regression | 0.062719 | Akaike info criterion | -2.694426 |
| Sum squared resid | 1.994400 | Schwarz criterion | -2.669517 |
| Log likelihood | 690.0786 | Hannan-Quinn criter. | -2.684660 |
| Durbin-Watson stat | 1.983803 | | |

# *ADF test - ukážka použitia*

- Otázky k výstupu:
  - Aká hypotéza sa testuje a aký je záver (na základe P hodnoty - zamietame nulovú hypotézu alebo nie?) Čo to znamená pre modelovanie premennej $LOGPCOCOA$?
  - Ako bola zostavená regresia uvedená vo výstupe? Kde sa v nej nachádza testovacia štatistika - vysvetlite, prečo nás zaujíma práve táto hodnota.