Mittestatsionaarsed aegread

Brooks, C. „Introductory Econometrics for Finance“
Ch 8.1. Stationarity and Unit Root Testing
Ch 8.2. Tests for Unit Root in the Presence of Structural Breaks
Ch. 11.8. Panel Unit Root and Cointegration test

Enders, W. „Applied Econometric Time Series“ 4th ed.
Ch 4 „Models with trend“

Teemad

- Statsionaarsuse olulisus
- Ühikjuur
  - Ühikjuure protsess
  - Ühikjuure testid
    - Dickey-Fuller test DF
    - Phillips-Perroni test PP
    - KPSS test
  - Deterministlik ja stohhastiline trend
- Ühikjuure testimine ja struktuursed muutused
  - Keskväärtuse hüppeline muutus
  - Perroni protseduur
  - Zivot-Andrewsi test
- Ühikjuure testimine paneelandmete korral
  - Levin-Lin-Chu test
  - Im-Pesaran-Shin test
  - Hadri LM test

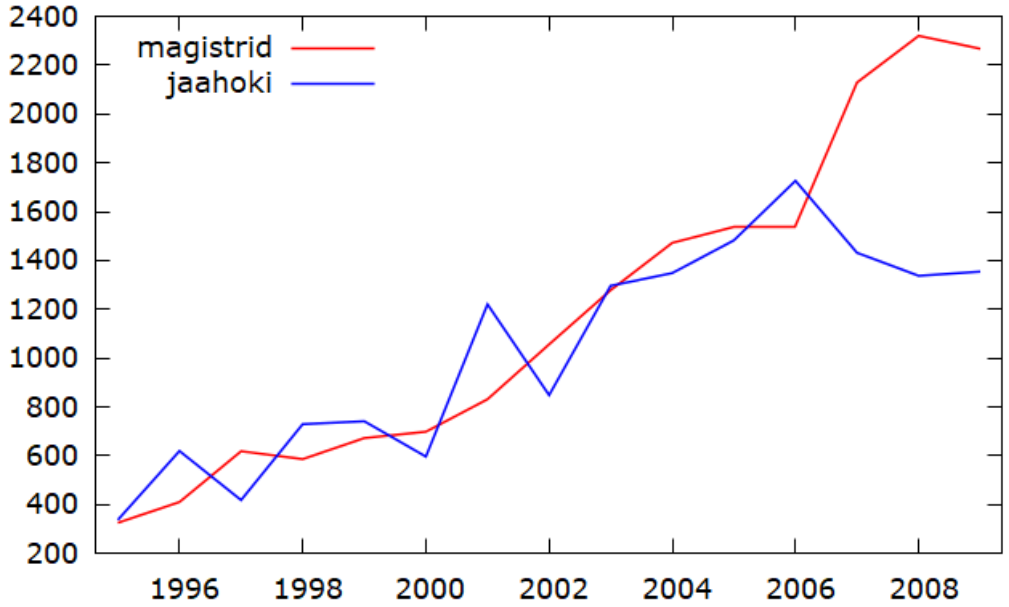
Miks on statsionaarsus oluline?

- Statsionaarsust tuleb kontrollida pikaajaliste seoste hindamisel.
  - Kui aegridu kasutatakse regressioonmudelite koostamisel, hoidumaks näivast regressioonist (*spurious regression*).
- Paljud testid, mida kasutatakse mudelite testimisel, eeldavad statsionaarsust.
- Statsionaarsus on oluline prognoosimise korral.
  - ARMA mudelid, mida kasutatakse aegridade modelleerimisel, on statsionaarsed.
- Statsionaarsed ja mittestatsionaarsed aegread reageerivad šokkidele erinevalt.
  - Statsionaarse protsessi korral ühe šoki u_t mõju aja möödudes väheneb.
  - Mittestatsionaarse protsessi korral šoki mõju aja möödudes ei vähene.

NÄIDE: magistriõpingute lõpetamine ja jäähokiga tegelemine Eestis

1995 – 2009
M – magistriõppe lõpetanud
J – jäähokiga tegutsejate arv

$$M_t = -111 + 1{,}25 J_t + u_t$$



	coefficient	std. error	t-ratio	p-value	
const	-111,168	277,067	-0,4012	0,6948	
J	1,25304	0,248444	5,044	0,0002	***

Mean dependent var	1182,133	S.D. dependent var	673,4129
Sum squared resid	2147240	S.E. of regression	406,4139
R-squared	0,661788	Adjusted R-squared	0,635771
F(1, 13)	25,43738	P-value(F)	0,000225

Statistiliselt oluline seos, jääkliikmete autokorrelatsioon puudub, heteroskedastiivsus puudub.

Phillipsi asümptootiline analüüs

Phillips (1986) analüüsis mittestatsionaarsete aegridade vahelise regressioonmudeli korral järgmiste statistikute käitumist:

- *t*-statistik,
- *F*-statistik,
- Durbin-Watsoni statistik *DW* (jääkliikmete autokorrelatsiooni testimine);
- Box-Ljungi *Q*-statistik (jääkliikmete autokorrelatsiooni testimine).

Tulemus piirjuhul, kui valimi maht T läheneb lõpmatusele ($T \to \infty$):
statistikutel puudub kindel asümptootiline jaotus.

Seega puuduvad korrektsed kriitilised väärtused, mille alusel viia läbi hüpoteeside testimist.

Siis ei ole korrektne hüpoteeside testimine võimalik.

Phillips, P. (1986), Understanding spurious regressions in econometrics. *Journal of Econometrics*, Vol. 33, 311 - 340

Statsionaarsus ja regressioonmudel

Soovime hinnata regressioonmudelit $y_t = a + bx_t + u_t$ kus y_t ja x_t on aegread.

Neli erinevat võimalust.

1. Nii y_t kui ka x_t on statsionaarsed. Kasutada võib klassikalist regressioonmudelit.

2. y_t ja x_t on üht ja sama järku integreeritud, vealiikmed u_t ei ole statsionaarsed: näiv regressioon. Tuleb hinnata regressioonmudelit statsionaarsete diferentside korral. Kui 1. järku diferentsid on statsionaarsed, siis

$$\Delta y_t = b_0 + b_1 \Delta x_t + \varepsilon_t, \quad \varepsilon_t = \Delta u_t$$

3. y_t ja x_t on üht ja sama järku integreeritud, vealiikmed u_t on statsionaarsed: kointegreeritud.

4. y_t ja x_t on erinevat järku integreeritud. Regressioon selliste suuruste vahel ei oma mõtet. Kui mõni aegrida 2. järku integreeritud, tasub seda logaritmida. Äkki logaritmide aegrida on 1. järku integreeritud.

ÜHIKJUUR

Juhuslik ekslemine kui ühikjuure protsess

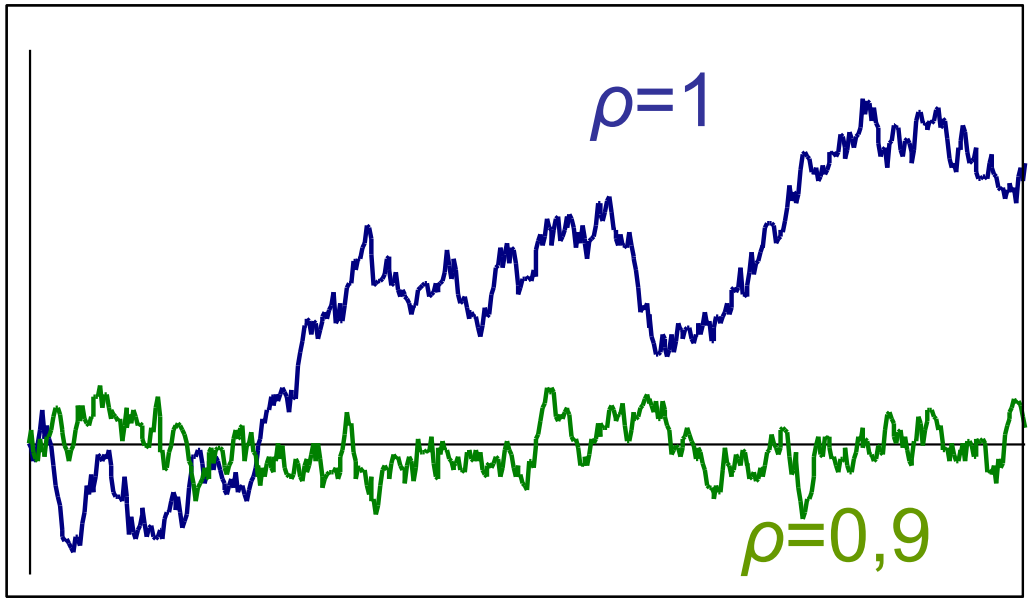
unit root process

Juhuslik ekslemine $y_t = y_{t-1} + u_t,$

Üldistame $y_t = \rho y_{t-1} + u_t, \qquad -1 < \rho < 1$

Demo „Ühikjuur..“.

Kui $|\rho|=1$, on tegemist puhta juhusliku ekslemisega.
Seda nimetatakse ühikjuure protsessiks, mis on mittestatsionaarne.



Kui $|\rho|< 1$, on protsess statsionaarne.

Kui $|\rho|> 1$, on plahvatuslik protsess (*explosive*) ja selliseid esineb majandusaegridade korral harva.

Ühikjuure testimine

Aegrea mudel $y_t = \rho y_{t-1} + u_t$

Et protsess oleks statsionaarne, peab $|\rho| < 1$.

Seda tuleb testida. Selleks tuleb viia läbi mudeli hindamine, leida ρ hinnang ja testida, kas
$H_0 : \rho=1$, $H_1 : \rho < 1$

või negatiivse ρ korral
$H_0 : \rho= -1$, $H_1 : \rho > -1$

PROBLEEM
Nullhüpoteesi kehtivuse korral on aegrida mittestatsionaarne ning aja *t* kasvades läheneb dispersioon lõpmatusele.

Kui nüüd kasutada ρ hindamiseks OLS-i, on hinnang nihkega.
Nihe on allapoole, st kui genereeriva protsessi ρ =1, siis ρ hinnang on <1.

Hüpoteesi ρ =1 testimiseks ei saa aegrea korral kasutada standardset *t*-testi.

Erinevad ühikjuure testid

- Dickey-Fuller DF
- Phillips-Perroni PP
- Kwiatkowski, Phillips, Schmidt ja Shin KPSS

Dickey-Fulleri ühikjuure test DF

$$y_t = \rho y_{t-1} + u_t \qquad H_0 : \rho = 1$$

Dickey ja Fuller: teisendame

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

$$\Delta y_t = \delta y_{t-1} + u_t,$$

$$\delta = \rho - 1 \qquad H_0 : \delta = 0$$

Testimiseks kasutatakse diferentside mudelit

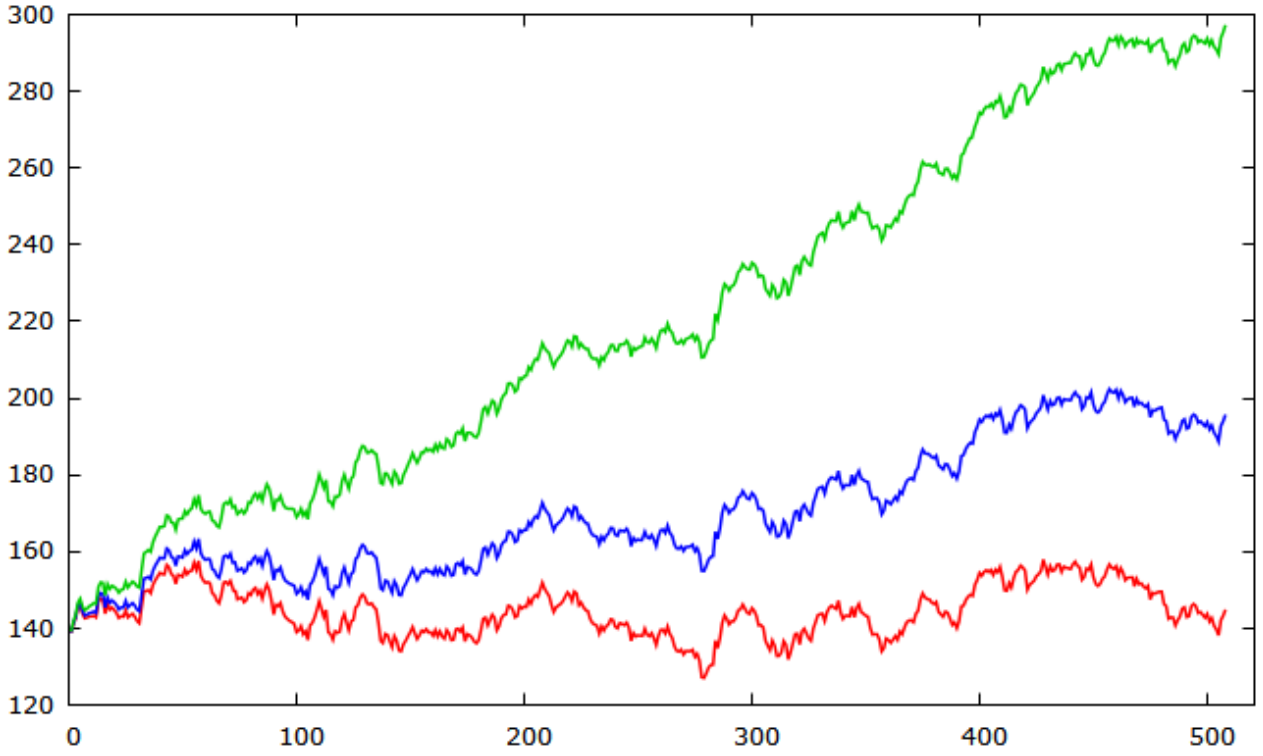
- Kui tuleb vastu võtta nullhüpotees: δ=0 (ρ=1), esineb ühikjuur.
- Kui tuleb vastu võtta sisukas hüpotees δ<0 (ρ<1), ühikjuurt pole.

Juhusliku ekslemise variatsioonid

Lihtne juhuslik ekslemine $y_t = y_{t-1} + u_t$

Triiviga juhuslik ekslemine $y_t = \beta_0 + y_{t-1} + u_t$

Triivi ja deterministliku trendiga juhuslik ekslemine $y_t = \beta_0 + \beta_1 t + y_{t-1} + u_t$



Triivi ja deterministliku trendiga juhuslik ekslemine

Triiviga juhuslik ekslemine

Lihtne juhuslik ekslemine

Lisavõimalustega Dickey-Fuller test, ADF

Augmented Dickey-Fuller test

Juhuslik ekslemine võib olla erinev

Ilma konstandita, lihtne

$$y_t = y_{t-1} + u_t \quad \Longrightarrow \quad \Delta y_t = \delta y_{t-1} + u_t$$

Konstandiga (triiviga)

$$y_t = \beta_0 + y_{t-1} + u_t \quad \Longrightarrow \quad \Delta y_t = \beta_0 + \delta y_{t-1} + u_t$$

Konstandi (triiviga) ja deterministliku trendiga

$$y_t = \beta_0 + \beta_1 t + y_{t-1} + u_t \quad \Longrightarrow \quad \Delta y_t = \beta_0 + \beta_1 t + \delta y_{t-1} + u_t$$

Erinevate mudelite korral on kriitilised väärtused erinevad

T=100

$$\Delta y_t = \delta y_{t-1} + u_t$$

$$\Delta y_t = \beta_0 + \delta y_{t-1} + u_t$$

$$\Delta y_t = \beta_0 + \beta_1 t + \delta y_{t-1} + u_t$$

Olulisuse nivoo	Ktiitilised väärtused		
	Ilma konstandi ja deterministliku trendita	Konstandiga	Konstandi ja deterministliku trendiga
0,01	-2,60	-3,51	-4,04
0,05	-1,95	-2,89	-3,45
0,1	-1,61	-2,58	-3,15

Tuleb valida õige testitava mudeli kuju!

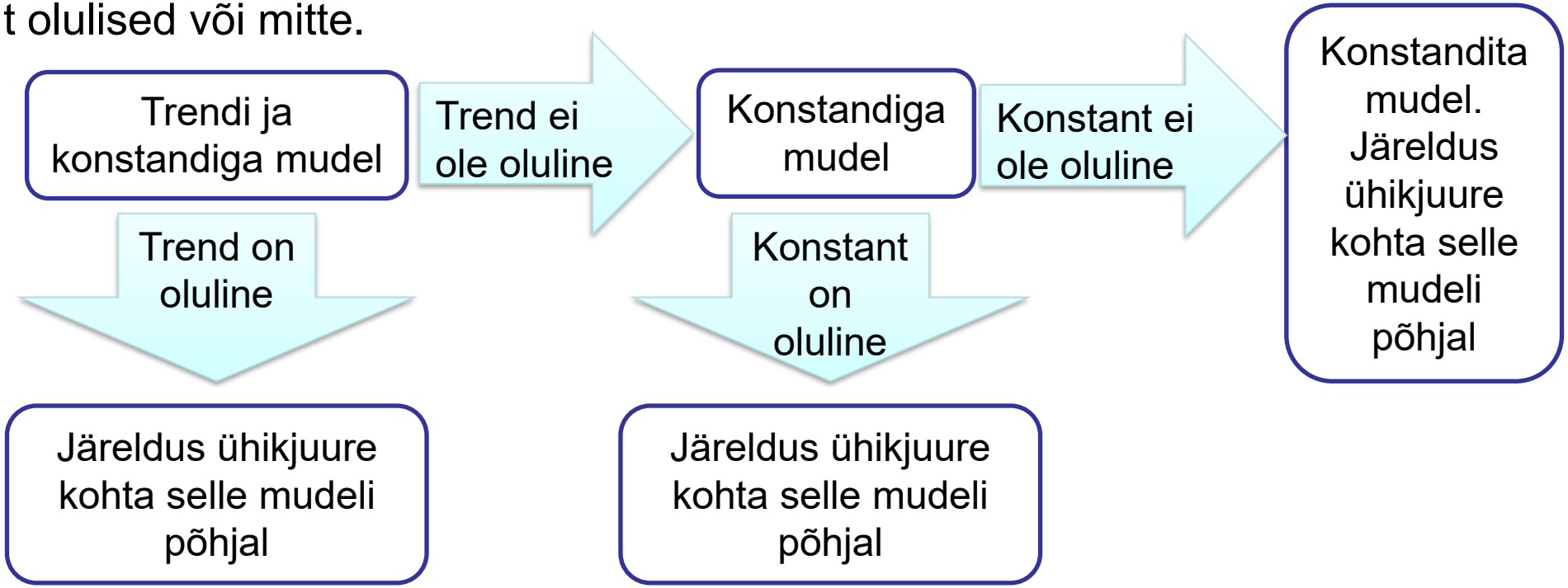
Programmides Gretl, Stata ja EViews on kasutusel MacKinnoni kriitilised väärtused.

MacKinnon, J. (1996) Numerical distribution functions for unit root and cointegration tests. *Journal of applied econometrics, 11*, 601-618.

Kuidas määrata testimiseks kasutatava võrrandi kuju

Ilma konstandita, konstandiga või konstandi ja trendiga?
Määramiseks on lihtsamaid ja keerulisemaid võimalusi.
Üks võimalus: visuaalne. Kui on silmnähtav trend, kasutada trendiga mudelit.

Teine võimalus: määrata, kas konstant ja trend on testimisel kasutatud regressioonmudelis statistiliselt olulised või mitte.



Märkus: kui ühikjuur puudub, vt deterministliku trendi määramist

NÄIDE: ADF, konstandi ja trendiga mudel

Eesti tööjõud 1991-2019 (mln). Testimiseks kasutame mudelit

$$\Delta y_t = \beta_0 + \beta_1 t + \delta y_{t-1} + u_t$$

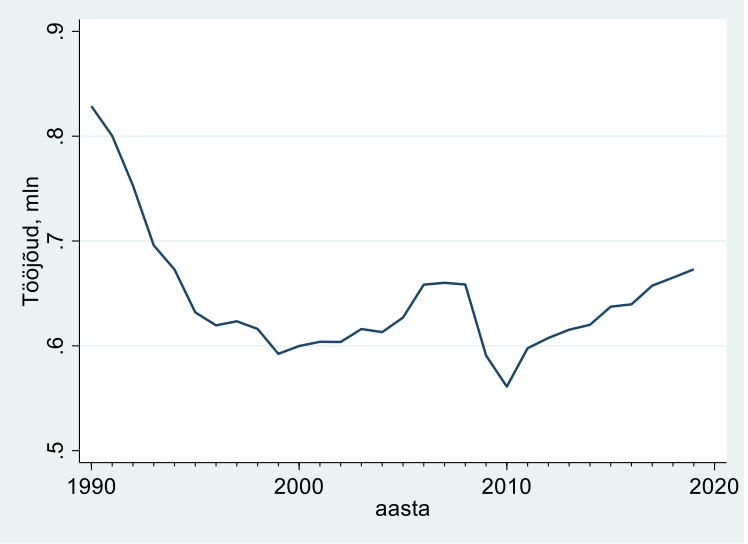
Aruanne programmis Stata

```
Dickey-Fuller test for unit root                   Number of obs   = 29
Variable: EMP                                      Number of lags  =  0

H0: Random walk with or without drift

                                         Dickey-Fuller
                   Test          ──────── critical value ────────
                statistic            1%           5%          10%
──────────────────────────────────────────────────────────────────
 Z(t)             -2.784        -4.343       -3.584       -3.230
──────────────────────────────────────────────────────────────────
MacKinnon approximate p-value for Z(t) = 0.2027.

Regression table
──────────────────────────────────────────────────────────────────────────────
        D.EMP | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
──────────────┼───────────────────────────────────────────────────────────────
          EMP |
δ         L1. |  -.1926368   .0691825    -2.78   0.010    -.3348435   -.0504301
            |
β1     _trend |   .0009121   .0004896     1.86   0.074    -.0000944    .0019185
β0      _cons |   .1049478   .0486307     2.16   0.040     .0049858    .2049097
──────────────────────────────────────────────────────────────────────────────
```



Et teha kindlaks, kas trend on oluline, vaatame testimiseks kasutatava regressioonmudeli hindamise tulemust.

Trend ei ole statistiliselt oluline, järgnevalt kasutame konstandiga mudelit.

NÄIDE: ADF, konstandiga mudel

Nüüd kasutame testimiseks konstandiga mudelit $\Delta y_t = \beta_0 + \delta y_{t-1} + u_t$

```
Dickey-Fuller test for unit root                   Number of obs   = 29
Variable: EMP                                      Number of lags  =  0

H0: Random walk without drift, d = 0

                                       Dickey-Fuller
                   Test      ---------- critical value ----------
               statistic           1%           5%          10%
--------------------------------------------------------------------
 Z(t)             -4.021       -3.723       -2.989       -2.625
--------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0013.

Regression table
------------------------------------------------------------------------------
       D.EMP | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         EMP |
δ        L1. |  -.2547052   .0633412    -4.02   0.000    -.3846707   -.1247397
β0     _cons |   .1585809   .0409437     3.87   0.001     .0745714    .2425904
------------------------------------------------------------------------------
```

Nullhüpotees on ümber lükatud, ühikjuurt ei esine.

Konstant on statistiliselt oluline, järeldus selle mudeli põhjal.

NÄIDE: ühikjuure testi tulemus võib sõltuda testimiseks kasutatava mudeli valikust

ADF test trendi ja konstandiga mudeli korral

H0: Random walk with or without drift

	Test statistic	Dickey-Fuller critical value 1%	5%	10%
Z(t)	**-2.784**	**-4.343**	**-3.584**	**-3.230**

MacKinnon approximate *p*-value for Z(t) = **0.2027.**

Ühikjuur on.

ADF test konstandiga mudeli korral, ilma trendita

H0: Random walk without drift, d = 0

	Test statistic	Dickey-Fuller critical value 1%	5%	10%
Z(t)	**-4.021**	**-3.723**	**-2.989**	**-2.625**

MacKinnon approximate *p*-value for Z(t) = **0.0013.**

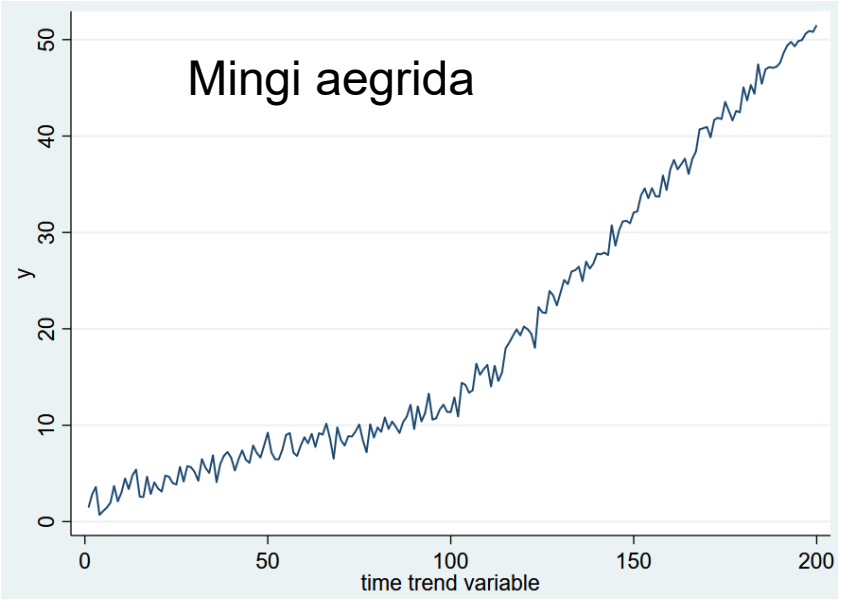
Ühikjuurt pole

Aga seda mudelit kasutada ei tohi,
sest trend polnud statistiliselt oluline.

Vale mudeli kasutamine testimisel võib anda vale tulemuse.

Seepärast tuleb alati analüüsida ka testimiseks kasutatava regressioonmudeli tulemusi (kas konstant on oluline, kas deterministlik trend on oluline).

NÄIDE: trendiga või ilma?



Visuaalne hinnang näitab, et aegreal esineb trend. Ajahetkel 100 see muutub.

ADF testi regressioonmudeli aruanne aga näitab, et trend ei ole statistiliselt oluline nivool 0,05.
See on tingitud sellest, et kuna trend muutub, siis selle kordaja hindamise standardviga tuleb suur ja *t*-statistik väike.

ADF test

```
Augmented Dickey-Fuller test for unit root

Variable: y                                  Number of obs   = 195
                                             Number of lags  =   4

H0: Random walk with or without drift

                                      Dickey-Fuller
                   Test        ———— critical value ————
                statistic          1%           5%          10%
———————————————————————————————————————————————————————————————
 Z(t)             -0.586       -4.008       -3.437       -3.137
———————————————————————————————————————————————————————————————
MacKinnon approximate p-value for Z(t) = 0.9797.

Regression table
———————————————————————————————————————————————————————————————
         D.y | Coefficient  Std. err.      t    P>|t|     [95% conf
———————————————————————————————————————————————————————————————
           y |
         L1. |  -.0103264   .0176328    -0.59   0.559      -.04511
         LD. |  -.7053988   .0731071    -9.65   0.000    -.8496144
        L2D. |  -.4958468   .0865521    -5.73   0.000    -.6665849
        L3D. |  -.2289336   .0859881    -2.66   0.008    -.3985591
        L4D. |  -.1438549   .0709615    -2.03   0.044    -.2838381
             |
      _trend |   .0080987   .0046418     1.74   0.083    -.0010581
       _cons |   .0299188   .1932033     0.15   0.877    -.3512061
———————————————————————————————————————————————————————————————
```

Antud juhul tuleks ühikjuure testimisel lähtuda ikka trendiga mudelist.

NÄIDE: Eelneva aegrea testimine ilma trendita

Konstandiga mudel

```
Augmented Dickey-Fuller test for unit root

Variable: y                              Number of obs   = 197
                                         Number of lags  =   2

H0: Random walk without drift, d = 0

                                   Dickey-Fuller
                   Test      ──────── critical value ────────
               statistic          1%           5%          10%
──────────────────────────────────────────────────────────────
 Z(t)            2.816        -3.478       -2.884       -2.574
──────────────────────────────────────────────────────────────
MacKinnon approximate p-value for Z(t) = 1.0000.

Regression table
──────────────────────────────────────────────────────────────────────────────
        D.y │ Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
────────────┼─────────────────────────────────────────────────────────────────
          y │
        L1. │   .0151937   .0053959     2.82   0.005     .0045512    .0258362
        LD. │  -.6470046   .0678921    -9.53   0.000    -.7809103   -.5130989
       L2D. │   -.358951   .0674515    -5.32   0.000    -.4919878   -.2259143
            │
      _cons │   .1934564   .1299971     1.49   0.138     -.062941    .4498538
──────────────────────────────────────────────────────────────────────────────
```

Ilma konstandita mudel

```
Augmented Dickey-Fuller test for unit root

Variable: y                              Number of obs   = 197
                                         Number of lags  =   2

H0: Random walk without drift, a = 0, d = 0

                                   Dickey-Fuller
                   Test      ──────── critical value ────────
               statistic          1%           5%          10%
──────────────────────────────────────────────────────────────
 Z(t)            6.117        -2.587       -1.950       -1.616
──────────────────────────────────────────────────────────────

Regression table
──────────────────────────────────────────────────────────────────────────────
        D.y │ Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
────────────┼─────────────────────────────────────────────────────────────────
          y │
        L1. │   .0213345   .0034877     6.12   0.000     .0144559    .0282131
        LD. │  -.6441971    .068078    -9.46   0.000    -.7784651    -.509929
       L2D. │  -.3539287   .0675776    -5.24   0.000    -.4872098   -.2206476
──────────────────────────────────────────────────────────────────────────────
```

Probleem: teststatistik Z(t) on positiivne

Järelikult ilma trendita mudelid ühikjuure testimiseks ei sobi ja tuleb kasutada trendiga mudelit (eelnev slaid).

Kui ADF teststatistik positiivne

Kui ADF teststatistik tuleb positiivne, siis kindlasti vastava kujuga mudel ühikjuure testimiseks EI SOBI.

ADF test, testitakse kordajat δ

$$\Delta y_t = \delta y_{t-1} + u_t, \quad \delta = \rho - 1$$

Kui teststatistik $Z(t) = \dfrac{\delta}{se(\delta)} > 0$ siis $\delta > 0$

ja järelikult $\rho > 1$

Siis peaks $y_t = \rho y_{t-1} + u_t$ olema plahvatuslik protsess, mida majandusaegridades üldiselt ei eksisteeri.

Testimine koos viitaegadega

$$\Delta y_t = \delta y_{t-1} + u_t$$

Testimine sellisel kujul annab õige tulemuse siis, kui u_t on valge müra (autokorrelatsioon puudub).

Kui jääkliikmetel u_t esineb autokorrelatsioon, soovitasid Dickey ja Fuller lisada testmudelisse diferentside viitajad Δy_{t-1}, Δy_{t-2}, …

$$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^{p} \theta_i \Delta y_{t-1} + u_t$$

AR(p) protsess

Kõik variandid kokku

$$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^{p} \theta_i \Delta y_{t-1} + \beta_m d_{m\,t} + u_t$$

Deterministlik osa

$m = 1: \quad d_1 = 0$ ilma konstandita võrrand

$m = 2: \quad d_2 = 1$ konstandiga võrrand

$m = 3: \quad d_{3\,t} = \{1, t\}$ konstandi ja deterministliku trendiga võrrand

ADF test ja viitaegade arv

$$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^{p} \theta_i \Delta y_{t-1} + \beta_m d_{m\,t} + u_t$$

Liiga vähe viitaegu: esineb jääkliikmete autokorrelatsioon, teststatistik vale, on nihkunud nullist eemale: I liiki vea tõenäosus suur.

Ühikjuur on (kehtib H_0), aga test näitab, et ei ole.

Liiga palju viitaegu: *t*-statistik kaldub 0 poole ning testi võimsus väike: II liiki vea tõenäosus suur.

Ühikjuurt ei ole (kehtib H_1), aga test näitab, et on.

Sobiva viitaegade arvu *p* määramine.

1. Võib vaadata, kas viitajad on statistiliselt olulised. Alustatakse maksimaalsest viitaegade arvust ja hakatakse neid eemaldama, kuni kõik on statistiliselt olulised (nt Stata).
2. Võib kasutada informatsioonikriteeriumeid (nt Gretl).

Schwert (1989) leidis simulatsioonide abil, et maksimaalne viitaegade arv võiks olla

$$p_{\max} = \left[12 \times \left(\frac{T}{100} \right)^{0,25} \right]$$

[ ] tähendab täisosa

Näiteks

T	10	50	100
p max	6	10	12

Schwert, G. W. (1989). Tests for Unit Roots: A Monte Carlo Investigation. *Journal of Business & Economic Statistics*, *7*(2), 5–17.

Phillips-Perroni test

ADF test $$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^{p} \theta_i \Delta y_{t-1} + \beta_m d_{m\,t} + u_t$$

Phillips-Perroni test PP $$\Delta y_t = \delta y_{t-1} + \beta_m d_{m\,t} + u_t$$

kohandatud standardvead

Võimaliku autokorrelatsiooni arvestamine

Phillips-Perroni (PP) testi korral on lähenemine jääkliikmete autokorrelatsioonile erinev: kordaja δ standardvead on kohandatud (Newey-West standardvead).
Need võtavad arvesse võimalikku jääkliikmete autokorrelatsiooni.
Aga nende arvutamisel kasutatakse etteantud viitaegade arvu.

Viimasel ajal on populaarseks muutunud just Phillips-Perroni test: pole vaja viitaegade arvu üle muret tunda.

Phillips, P. C. B. and Perron, P. (1988) Testing for a unit root in time series regression. Biometrika, 75, 335-346.

Ühikjuure testide võimsus, I

- Testi võimsus (*power*) näitab, kui suure tõenäosusega lükatakse tagasi mittekehtiv nullhüpotees (võimsus =1 – II liiki vea tõenäosus).
- Hea võimsusega ühikjuure test lükkab tagasi nullhüpoteesi, kui aegreal tegelikult ühikjuur puudub.
- Ühikjuure testide probleem on madal võimsus: võivad näidata ühikjuurt, kui aegreal tegelikult ühikjuur puudub.
  - Põhjus: lõpliku valimi korral on võimalik hästi lähendada deterministliku trendiga protsessi ühikjuure protsessiga ja vastupidi.
- Monte Carlo simulatsioonid on näidanud, et nii Dickey-Fulleri kui ka Phillips-Perroni testi võimsus on madal.
  - St, kui ühikjuurt ei esine (nullhüpotees ei kehti), võib testi tulemus olla, et esineb (mittekehtiva nullhüpoteesi vastuvõtmine), st II liiki vea tõenäosus on suur.

Ühikjuure testide võimsus, II

1. Testi võimsus sõltub vaadeldava ajavahemiku pikkusest.
   30 vaatlust 30 aasta jooksul vs 100 vaatlust 100 päeva jooksul

   II liiki vea tõenäosus väiksem

2. Kui $\rho \approx 1$, aga mitte täpselt 1, võib ühikjuure test anda tulemuseks ühikjuure.
   Kinnituseks võib kasutada KPSS testi.

3. Probleemiks võivad olla struktuursed muutused.
   Lahenduseks modelleerida struktuurseid muutusi.

KPSS test

Kwiatkowski, Phillips, Schmidt ja Shin (KPSS) pakkusid välja testi, mille korral nullhüpoteesiks on ühikjuure puudumine:

H_0 aegrida on (trend-)statsionaarne

H_1 aegrida ei ole (trend-)statsionaarne, esineb ühikjuur

Aegrida esitatakse kujul

$$y_t = \beta t + r_t + e_t$$

deterministlik trend

juhuslik ekslemine

statsionaarne juhuslik komponent

kus juhusliku ekslemise komponent $r_t = r_{t-1} + u_t, \quad u_t \sim iid(0, \sigma^2)$

Testitakse nullhüpoteesi, et dispersioon Var(u_t) =0. Kui nii, siis

$$r_t = r_{t-1} = r_{t-2} = \ldots = r_0$$

Sellisel juhul ei ole r_t juhuslik ekslemine ja tegemist on lihtsa trend-statsionaarse mudeliga

$$y_t = \beta t + r_0 + e_t \qquad H_0$$

Dispersiooni testimiseks kasutatakse LM (*Lagrange multiplier*) teststatistikut.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Y. Shin. 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, 1992, 159-178

KPSS test kui täiendus ADF või PP testile

KPSS test on täienduseks ADF või PP testile, sobib kinnituse leidmiseks.

ADF/PP	KPSS	Lõplik järeldus
H_0, ühikjuur	H_1, ühikjuur	Ühikjuur esineb
H_1, ei ole ühikjuurt	H_0, ei ole ühikjuurt	Ei ole ühikjuurt
H_0, ühikjuur	H_0, ei ole ühikjuurt	Aegrida ei sisalda piisavalt informatsiooni, et otsustada, kas on deterministliku või stohhastilise trendiga.
H_1, ei ole ühikjuurt	H_1, ühikjuur	

Programmis Stata tuleb KPSS testi jaoks installeerida lisamoodul kpss *)

```
ssc install kpss
```

*) Baum, C.F. (2000), "KPSS: Stata module to compute Kwiatkowski-Phillips-Schmidt-Shin test for stationarity," Statistical Software Components S410401, Boston College Department of Economics, revised 13 May 2018.

Deterministliku või stohhastilise trendi kindlakstegemine

Kui trendi ja konstandiga mudeli korral

- trend on statistiliselt oluline

  ja

- trendiga mudeli korral ühikjuur puudub,

siis vaadata ühikjuure testi ilma trendita mudeli korral.

Kui seal ühikjuur on, siis deterministliku trendi lisamine mudelisse kõrvaldas ühikjuure.

Järelikult on deterministliku trendiga aegrida ehk trend-statsionaarne TS.

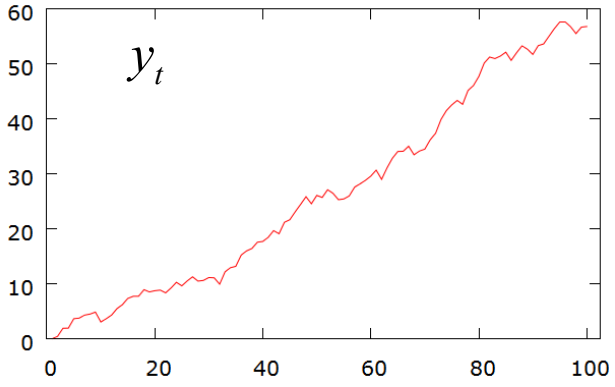
Kui trendi ja konstandiga mudeli korral trend on statistiliselt oluline ja tuleb vastu võtta nullhüpotees, st ühikjuur on, siis stohhastiline trend.
Ka diferents-statsionaarne DS.

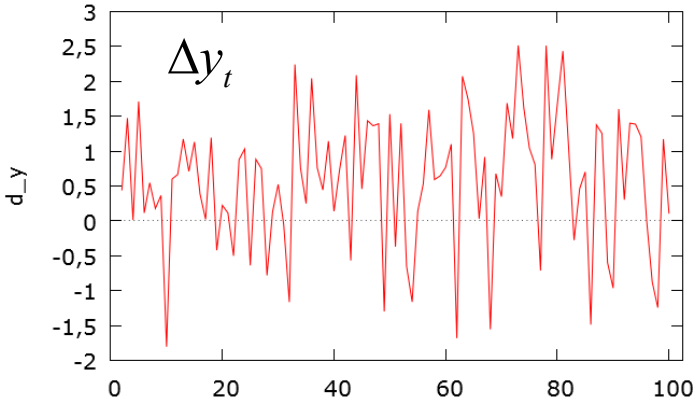
Stohhastilise trendi eemaldamine

Stohhastilise trendiga mudel $y_t = \mu + y_{t-1} + u_t$

Leiame 1. järku diferentsid $\Delta y_t = y_t - y_{t-1} = \mu + u_t$

Stohhastiline trend on eemaldatud



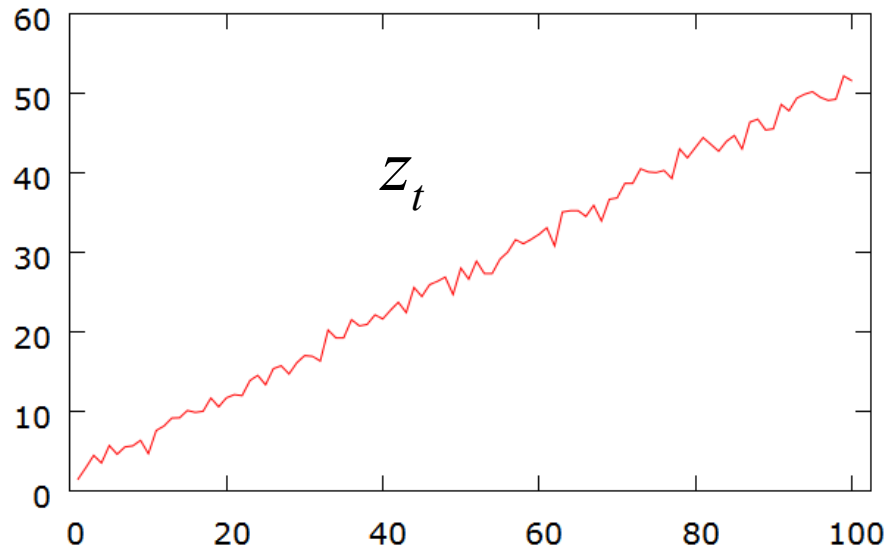


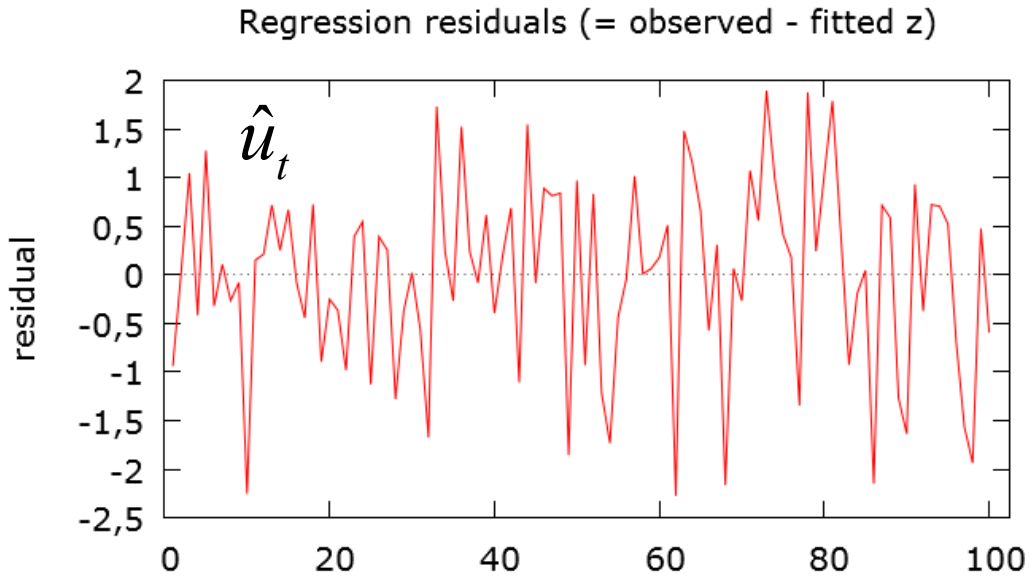
Stohhastilise trendi eemaldamiseks.

1. Leida 1. järku diferentsid.
2. Testida nende statsionaarsust. Kui on statsionaarsed, on stohhastiline trend eemaldatud.
3. Kui 1. järku diferentsid ei ole statsionaarsed, diferentsida veelkord. Seni kuni on saavutatud statsionaarsus.

Stohhastilise trendi eemaldamiseks tuleb diferentsida.

Deterministliku trendi eemaldamine





Trendi eemaldamiseks

1. Hinnata regressioonmudelit

$$z_t = \alpha + \beta t + u_t$$

2. Salvestada selle mudeli jääkliikmed $\hat{u}_t$
3. Nende jääkliikmete aegrida peab olema statsionaarne. Testida ühikjuure testiga, ilma konstandita mudel.

Regressioon aegridade vahel ning stohhastiline ja deterministlik trend

$$y_t = b_1 + b_2 x_{2t} + b_3 x_{3t} + b_4 x_{4t} + u_t$$

1. Kui kõik aegread stohhastilise trendiga, sama järku integreeritud: regressioon vastavate diferentside vahel. Näiteks kui 1. järku integreeritud, siis

$$\Delta y_t = b_2 \Delta x_{2t} + b_3 \Delta x_{3t} + b_4 \Delta x_{4t} + \varepsilon_t$$

2. Kui kõik aegread deterministliku trendiga, siis lisada mudelisse deterministlik trend *t*. Näiteks lineaarne trend

$$y_t = b_1 + b_2 x_{2t} + b_3 x_{3t} + b_4 x_{4t} + b_5 t + u_t$$

3. Kui mõned aegread stohhastilise trendiga, aga mõned deterministliku trendiga, on probleem. Mõned autorid soovitavad siis lisada mudelisse deterministlik trend *t* ja väidavad, et sellest piisab, et kõrvaldada näiva regressiooni oht.

$$y_t = b_1 + b_2 x_{2t} + b_3 x_{3t} + b_4 x_{4t} + b_5 t + u_t$$

A. Noriega and D. Ventosa-Santaulària, (2005) “Spurious regression under deterministic and stochastic trends,” Universidad de Guanajuato, Department of Economics and Statistics, Department of Economics and Finance Working Papers, MPRA Paper No. 58772.

Deterministliku või stohhastilise trendiga mudel?

- Kumb on parem?
- Poleemika selle üle on kestnud majandusteoorias aastakümneid. Ühest vastust pole.
- Mõlemad lähtuvad konstantse kasvu mudelist.
- Erinevus on juhuslikkuse modelleerimisel.

Erinevus pole mitte matemaatiline, vaid sisuline: kuidas käituvad majandussuurused?

Deterministlik vs stohhastiline trend

Deterministlik trend $$y_t = a_0 + a_1 t + u_t$$

$$\Delta y_t = y_t - y_{t-1} = (a_0 + a_1 t + u_t) - (a_0 + a_1(t-1) + u_{t-1}) = a_1 + u_t - u_{t-1}$$

Tinglik keskväärtus $E[\Delta y_t \mid y_{t-1}] = E\left[a_1 + u_t - u_{t-1}\right] = a_1$

Iga kord kui *t* suureneb 1 võrra, muutub keskväärtus konstandi a_1 võrra.

Juhuslik ekslemine $$y_t = y_{t-1} + u_t$$

Kui y_{t-1} on antud, ei tea me, kas hetkel *t* liigub ülespoole või allapoole, tõenäosused on ühesugused.
Stohhastiline trend.

$$\Delta y_t = u_t$$

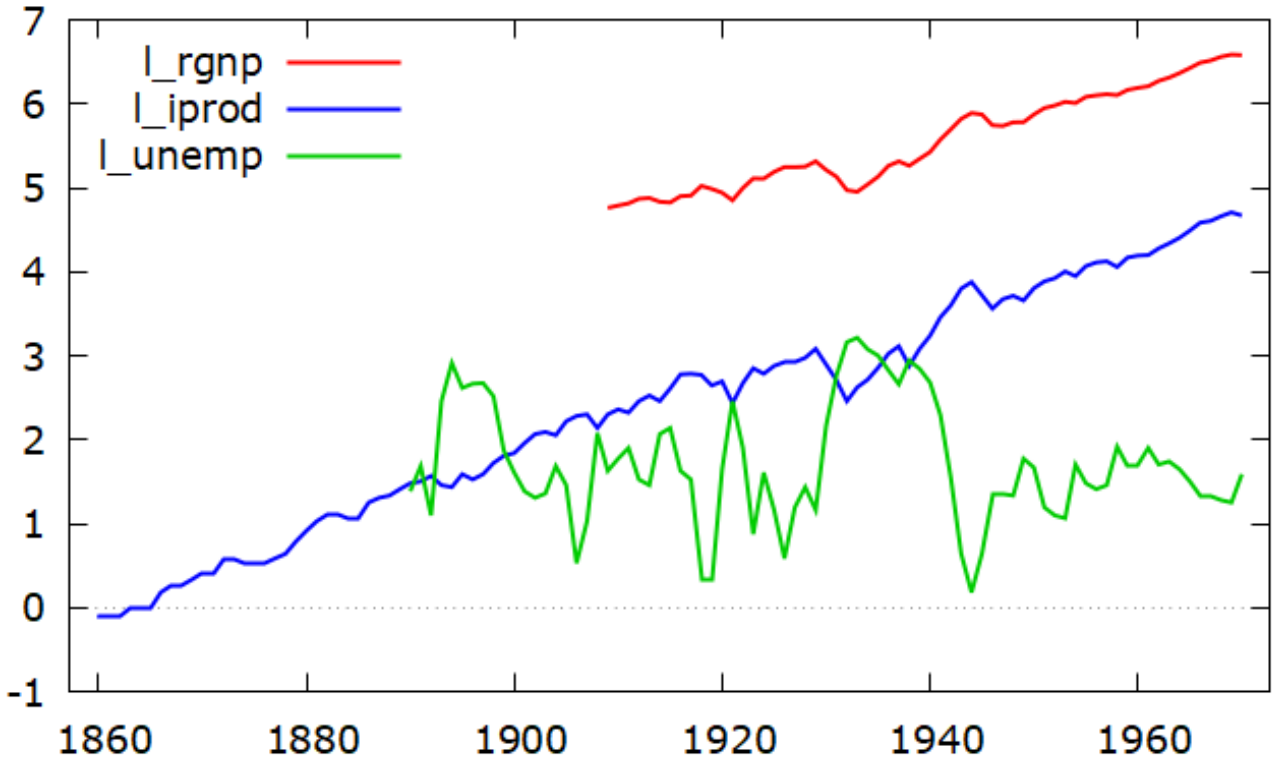
Tinglik keskväärtus $E[\Delta y_t \mid y_{t-1}] = E\left[u_t\right] = 0$

Nelsoni ja Plosseri tulemused

Nelson ja Plosser analüüsisid mitmeid USA makroökonoomika aegridu 1860-1970.
Kasutasid logaritmitud väärtusi.
Kas stohhastiline trend (ST) või deterministlik trend (DT)?

Tulemused:

Real GNP	ST
Nominal GNP	ST
Industrial production	ST
Unemployment rate	DT



Nelson, C. R. & Plosser, C. R. Trends and random walks in macroeconomic time series: Some evidence and implications
Journal of Monetary Economics, 1982, 10, 139 - 162

Perroni tulemused

Perron (1989) samade andmete põhjal: tegemist on deterministliku trendi muutumisega, ühikjuurt ei ole. Struktuurne muutus, mille põhjustas börsikrahh 1929.

Reaalpalga logaritmi aegrea mudel

$$y_t = \mu_1 + \beta t + (\mu_2 - \mu_1) D_L + u_t$$

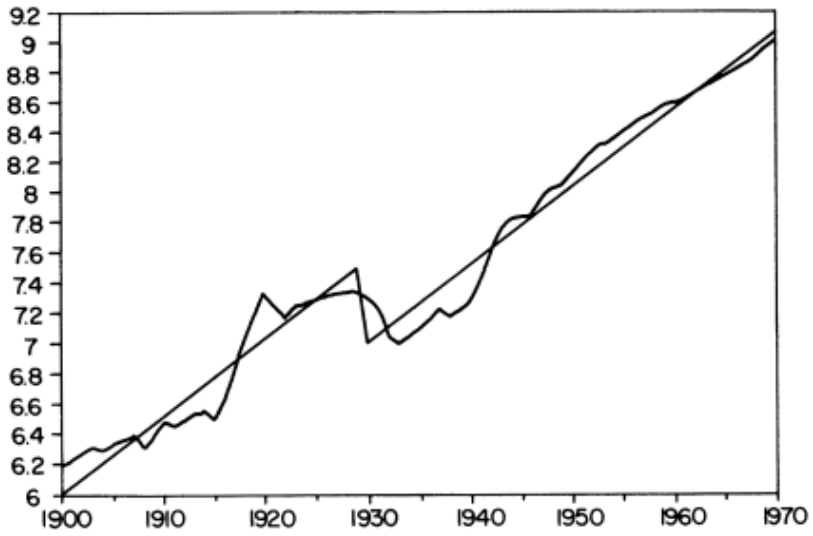


FIGURE 1.—Logarithm of "Nominal Wages."

Aktsiahindade logaritmi aegrea mudel

$$y_t = \mu_1 + \beta_1 t + (\mu_2 - \mu_1) D_L + (\beta_2 - \beta_1) D_T t + u_t$$

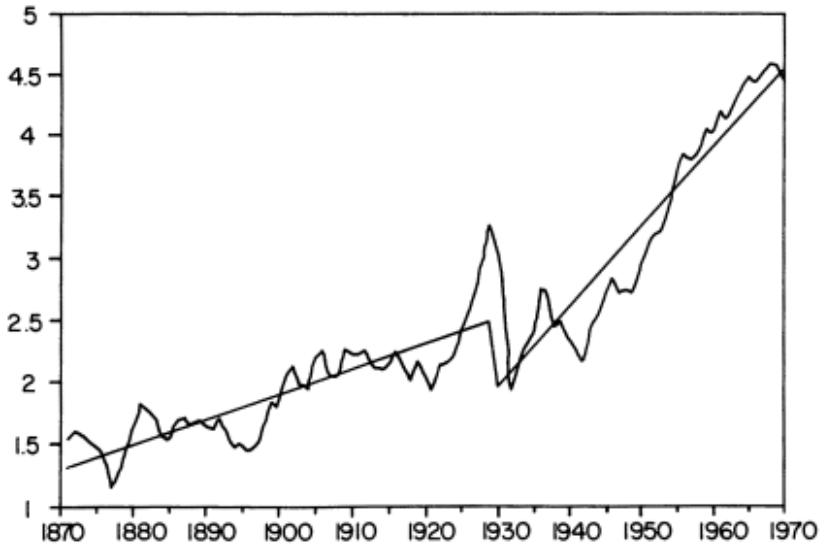


FIGURE 3.—Logarithm of "Common Stock Prices."

Perron, P. (1989), The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis *Econometrica, The Econometric Society,* Vol. 57 , 1361-1401.

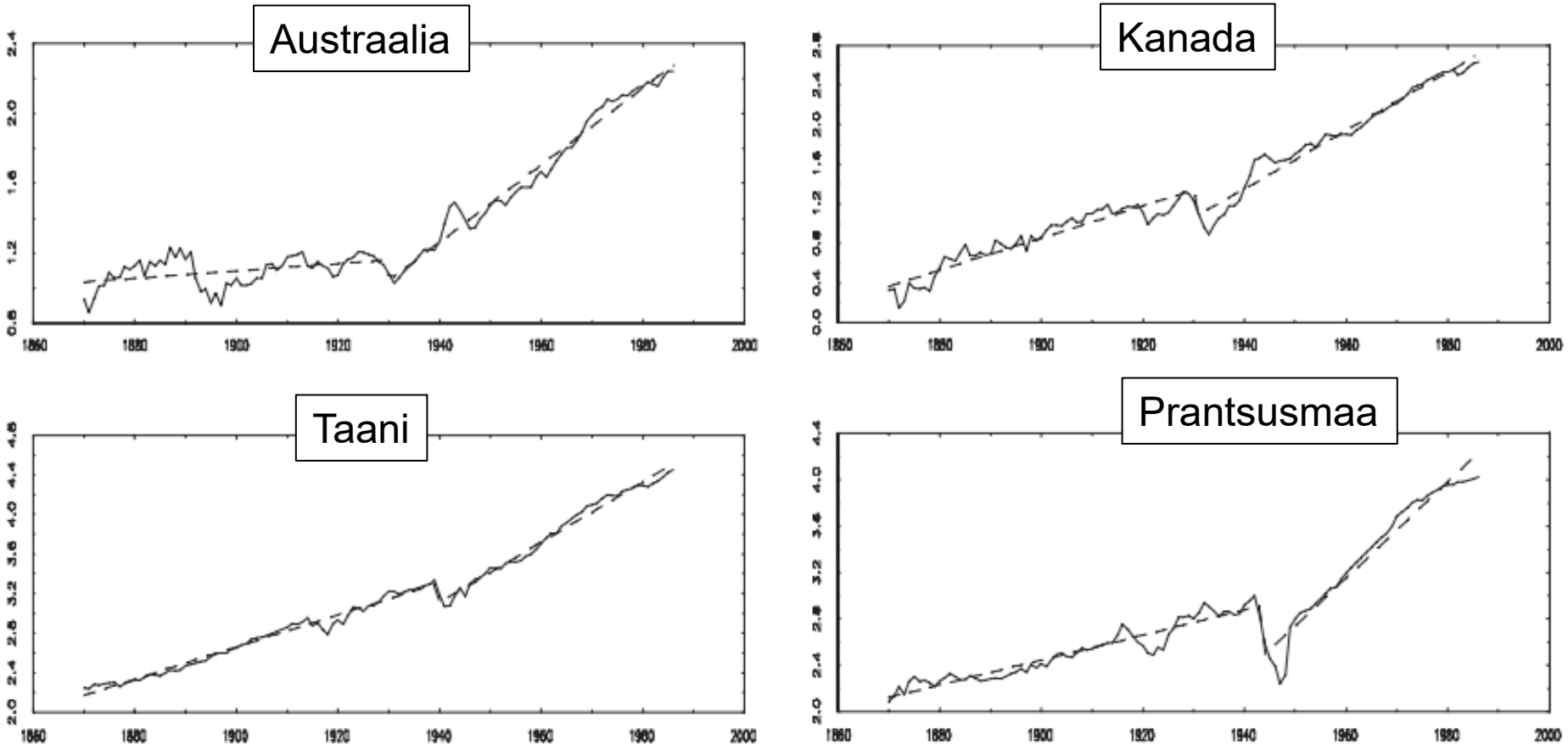
Ühikjuure testimine, kokkuvõte

- Ei ole korrektne rääkida statsionaarsuse testimisest: ühikjuure testiga määratakse ka kindlaks, kas trendiga aegreal on stohhastiline või deterministlik trend.
  - Deterministliku trendiga aegrida ei ole statsionaarne.
- Analüüsida, millise kujuga mudelit testimiseks kasutada:
  - konstandi ja trendiga;
  - konstandiga;
  - ilma konstandita.
- Erinevad testid:
  - ADF testi korral tuleb valida sobiv viitaegade arv;
  - PP testi korral pole viitaegade arvu määramine vajalik;
  - et tulemuses kindel olla, kasutada kontrollimiseks KPSS testi.
- Kui aegreal on olemas trend, tuleb testimise abil välja selgitada, kas trend on stohhastiline või deterministlik.
  - Kui trendiga testimisel ühikjuur puudub (nullhüpotees ümber lükatud), aga ilma trendita mudeli korral ühikjuur esineb, on trend deterministlik.
  - Kui trendiga testimisel ühikjuur esineb (võtame vastu nullhüpoteesi), on trend stohhastiline.
- Kui on kahtlusi, et võivad esineda struktuursed muutused, kasutada ühikjuure testimist koos struktuursete muutustega (Perroni protseduur või Zivot-Andrewsi test).

ÜHIKJUURE TEST JA STRUKTUURSED MUUTUSED

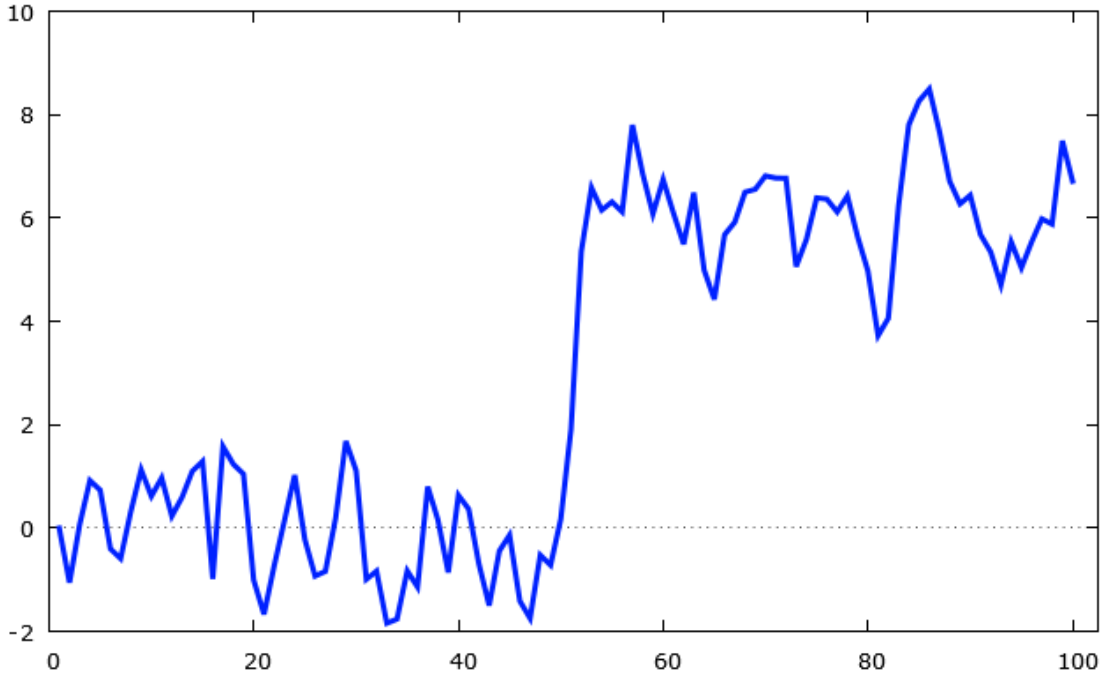
Aegridades on muutuv trend

Fig. 1. Annual log per capita gross domestic product (GDP): 1870–1986.



Perron, P. & Zhu, X. (2005), Structural breaks with deterministic and stochastic trends *Journal of Econometrics* , Vol. 129 , 65 - 119

Statsionaarne protsess ja keskväärtuse hüppeline muutus



$$t = 0,\ldots,50$$

$$y_t = 0,5y_{t-1} + u_t$$

$$t = 51,\ldots,100$$

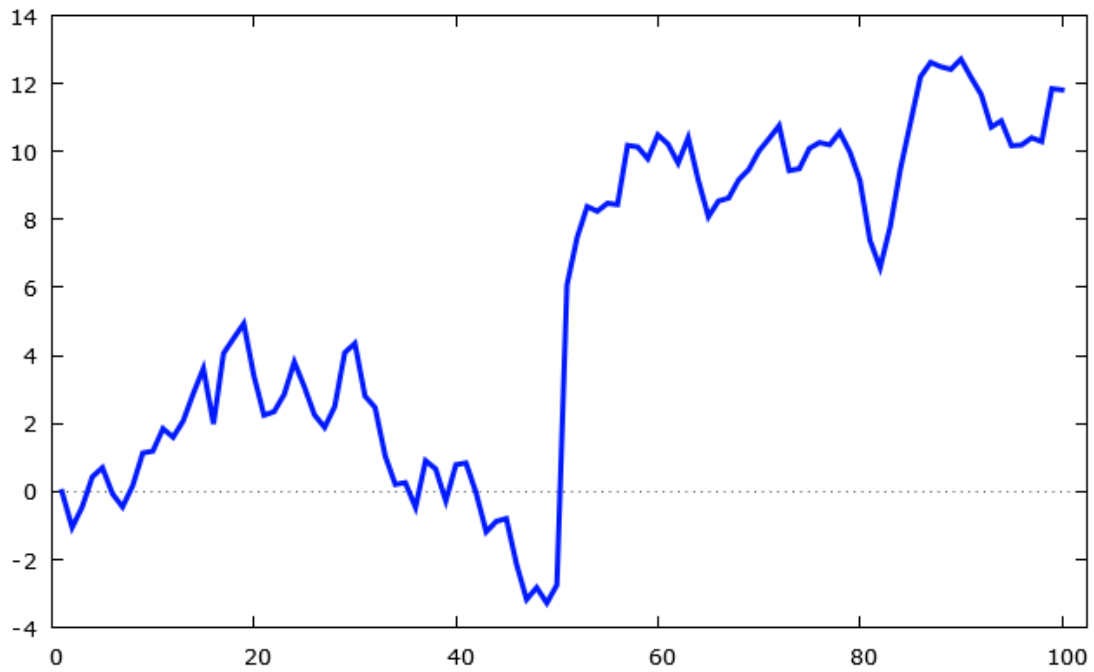
$$y_t = 3 + 0,5y_{t-1} + u_t$$

Kummaski alamvalimis statsionaarne protsess, keskväärtus muutub hüppeliselt.

$$y_t = 3D + 0,5y_{t-1} + u_t$$

fiktiivne tunnus $D = \begin{cases} 0, & t \le 50 \\ 1, & t > 50 \end{cases}$

Ühikjuure protsess ja ühekordne hüpe



$$y_t = y_{t-1} + D_p + u_t$$

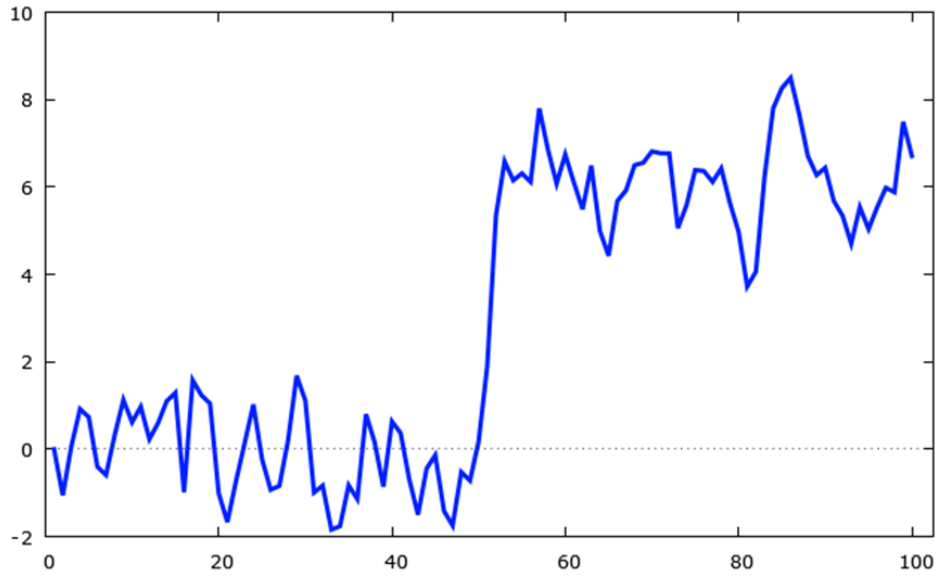
Ühekordne impulss

$$D_p = \begin{cases} 0, & t \neq 51 \\ 10, & t = 51 \end{cases}$$

Ühikjuure protsessi korral on ühekordsel impulsil jääv mõju kõigile järgnevate y_t väärtustele, st mõjutab y_t keskväärtust, kui $t \geq p$

$$y_{t \geq p} = y_0 + \sum_{i=0}^{t-1} u_{t-i} + D_p$$

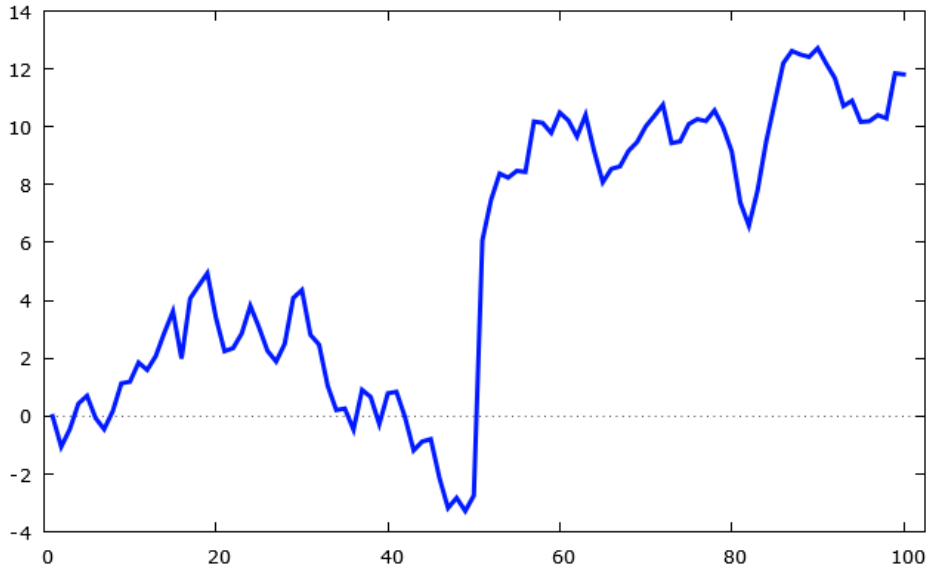
Kumb on kumb?



Statsionaarne, ühekordselt muutuv keskväärtus

?

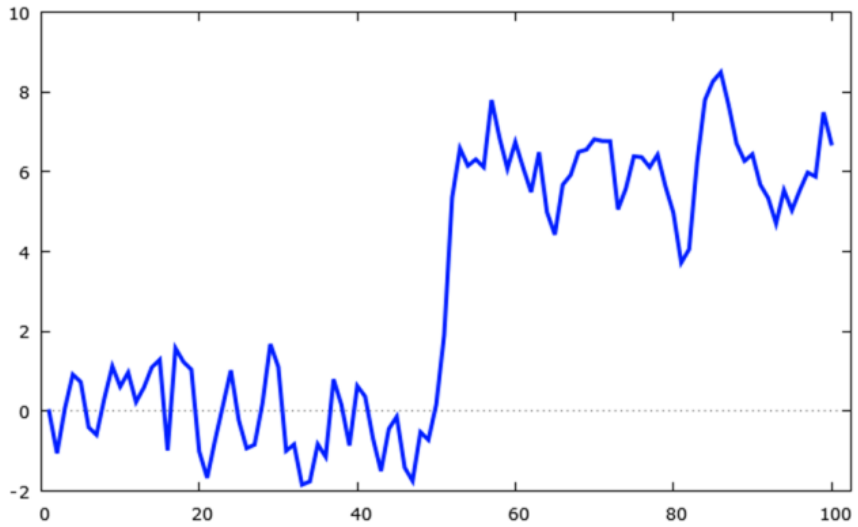
Mittestatsionaarne, ühekordse impulsiga



Kuidas testida?

Keskväärtuse hüppeline muutus ja ühikjuure testimine

Genereeriv protsess, ühikjuurt pole. $y_t = 3D + 0{,}5 y_{t-1} + u_t$ $\quad D = \begin{cases} 0, & t \le 50 \\ 1, & t > 50 \end{cases}$



Testime tervet aegrida

$t = 0, \ldots, 100$

Testimine (koos trendi ja konstandiga) näitab, et on ühikjuur.
Kuigi tegelikult pole, on hüpe.

Testime eraldi alamvalimeid

$t = 0, \ldots, 50$ Testimine näitab, et ei ole ühikjuurt

$t = 51, \ldots, 100$ Testimine näitab, et ei ole ühikjuurt

NÄIDE: piima hind

Wang ja Tomek: piima hind Illinoisis, kuised väärtused.
Ühikjuure testimiseks trendiga mudel.

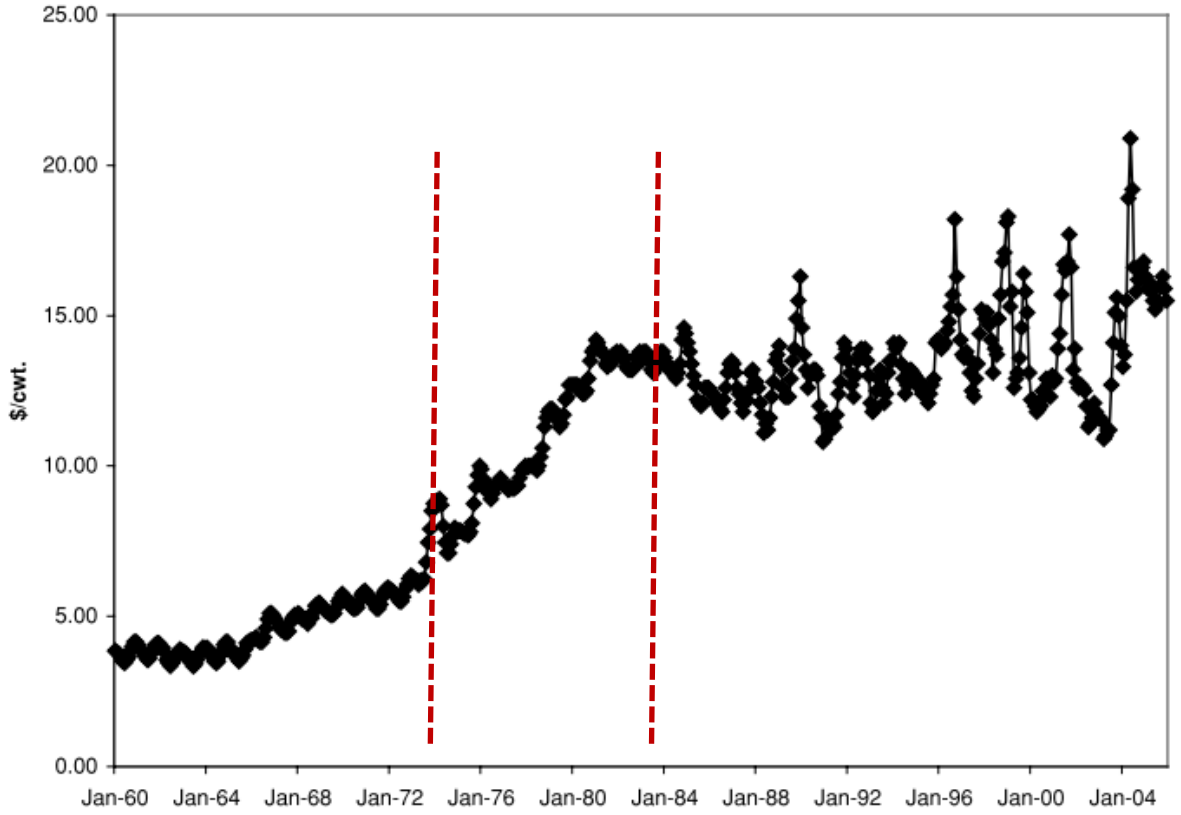


Figure 4. Monthly milk price ($/cwt), January 1960–December 2005, Illinois

Tulemused

Terve periood
Jaan 1960 – dets. 2005:
mittestatsionaarne (on ühikjuur)

Kolmeks jagatud
Jaan. 1960 – dets. 1972:
on ühikjuur

Jaan. 1974 – dets. 1980:
on ühikjuur

Jaan. 1982 – dets. 2005:
ei ole ühikjuurt

Wang, D. & Tomek, W. G. (2007), Commodity Prices and Unit Root Tests *American Journal of Agricultural Economics,* Vol. 89, 873-889

Struktuurseid muutusi tuleb arvestada

Kui me ei arvesta võimalikku struktuurset muutust (keskväärtuse muutus, trendi muutumine) ja kasutame ühikjuure testimiseks vale kujuga mudelit, võime tulemuseks saada ühikjuure, kuigi aegrida on mõlema alamperioodi kestel ilma ühikjuureta.

Struktuursed muutused ja Perroni protseduur

1. Jagada aegrida kaheks ja mõlemas osas testida ühikjuurt eraldi, kasutades ADF või PP testi.
   - Kaheks jagamisel on mõlema regressioonmudeli vabadusastmete arv väike.
   - Lühikese aegrea korral on parem testida tervet aegrida.

Perron näitas, et standardne ühikjuure testimine ei suuda ümber lükata nullhüpoteesi, kui deterministliku trendi korral esineb hüpe.

2. Kasutada Perroni poolt pakutud protseduuri, mis võimaldab hüppe esinemise korral kindlaks teha: kas diferents-statsionaarne DS või trend-statsionaarne TS.

Perroni protseduur võimaldab testida kolme erinevat situatsiooni:

- hüppeline eksogeenne muutus aegrea tasemes (keskväärtuse muutus) (variant A);
- eksogeenne trendi muutus (variant B);
- muutus nii aegrea tasemes kui trendis (variant C).

Perron, P. (1989), The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis *Econometrica, The Econometric Society,* Vol. 57 , 1361-1401

Perroni protseduur, variant A, I

Testimine juhul, kui on võimalik hüpe aegrea tasemes. 2 võimalikku mudelit.

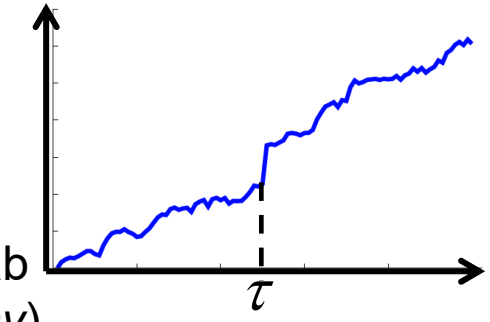
Nullhüpotees

$$H_{0A}: \quad y_t = a_0 + y_{t-1} + \mu_1 D_P + u_t$$

Nullhüpoteesi vastu võtmisel esineb stohhastiline trend ja ühekordne hüpe.

$$D_P = \begin{cases} 1, & t = \tau + 1 \\ 0, & t \neq \tau + 1 \end{cases}$$

D_p on fiktiivne tunnus, mis kirjeldab ühekordset impulssi (*pulse dummy*).



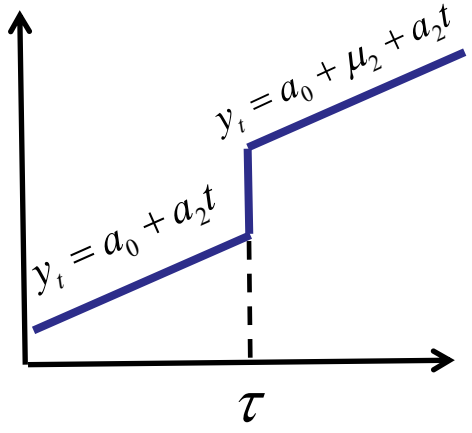
Sisukas hüpotees

$$H_{1A}: \quad y_t = a_0 + a_2 t + \mu_2 D_L + u_t$$

Sisuka hüpoteesi vastu võtmisel esineb deterministlik trend, mis hüppab.

$$D_L = \begin{cases} 1, & t > \tau \\ 0, & t \leq \tau \end{cases}$$

D_L on fiktiivne tunnus, mis mõjutab sirge vabaliiget (*level dummy*).



Tuleb kindlaks teha, kumb mudel kirjeldab aegrea muutumist paremini.

Erinevalt DF testist ei ole nullhüpoteesile vastav mudel saadav sisukale hüpoteesile vastavast mudelist y_{t-1} kordajale kitsenduse panekuga.

Perroni protseduur, variant A, II

Seepärast tuleb hüpoteesidele H_{0A} ja H_{1A} vastavad mudelid panna kokku.

$$H_{0A}: \quad y_t = a_0 + y_{t-1} + \mu_1 D_P + u_t \qquad H_{1A}: \quad y_t = a_0 + a_2 t + \mu_2 D_L + u_t$$

1. Hinnata kombineeritud mudelit $$y_t = a_0 + a_1 y_{t-1} + a_2 t + \mu_1 D_P + \mu_2 D_L + u_t$$

2. Kui jääkliikmetes esineb autokorrelatsioon, siis lisada diferentside viitajad, st hinnata mudelit

$$y_t = a_0 + a_1 y_{t-1} + a_2 t + \mu_1 D_P + \mu_2 D_L + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + u_t$$

3. Arvutada nullhüpoteesile a_1 =1 vastav *t*-statistik $$t = \frac{\hat{a}_1 - 1}{se(\hat{a}_1)}$$

ja võrrelda seda Perroni poolt leitud kriitiliste väärtustega.

Kui *t*-statistik on nullist kaugemal kui kriitiline, on nullhüpotees ümber lükatud. Siis on hüppega deterministlik trend (H_{1A}).
Kui tuleb vastu võtta nullhüpotees, siis on stohhastiline trend ja üksik hüpe .

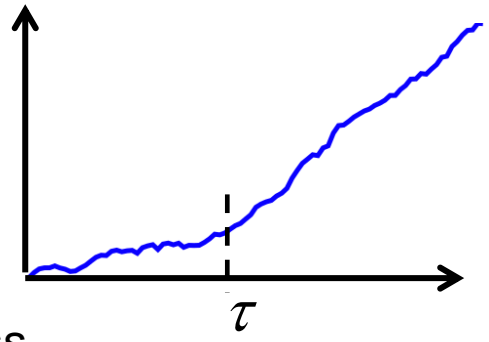
Perroni protseduur, variant B

Teine hüpoteeside paar: kas on ühikjuure protsess muutuva triiviga või deterministlik trend, mille tõus muutub (*Changing Growth*).

Nullhüpotees

$$H_{0B}: \quad y_t = a_0 + y_{t-1} + \mu_2 D_L + u_t$$

$$D_L = \begin{cases} 1, & t > \tau \\ 0, & t \le \tau \end{cases}$$



Kui see kehtib, siis ajahetkeni τ $\quad y_t = a_0 + y_{t-1} + u_t$

Peale ajahetke τ $\quad y_t = a_0 + \mu_2 + y_{t-1} + u_t$

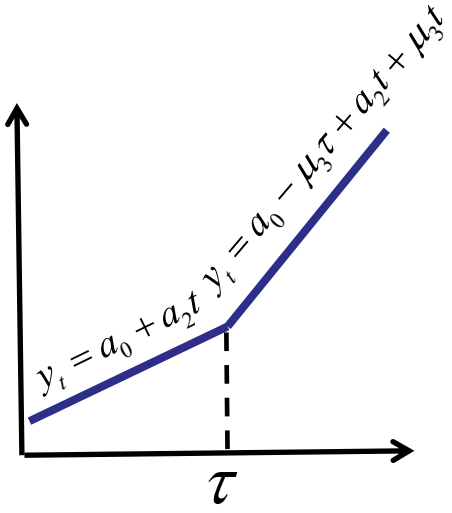
Triiviga ühikjuure protsess.
Ühikjuure protsess, triiv muutus.

Sisukas hüpotees

$$H_{1B}: \quad y_t = a_0 + a_2 t + \mu_3 D_T + u_t$$

Deterministlik trend muutub

$$D_T = \begin{cases} t - \tau, & t > \tau \\ 0, & t \le \tau \end{cases}$$



Mudel, mida tuleb hinnata ja mille kordajat a_1 testida (kas a_1 =1).

$$y_t = a_0 + a_1 y_{1-1} + a_2 t + \mu_2 D_L + \mu_3 D_T + u_t$$

Perroni protseduur, variant C

Kolmas hüpoteeside paar lubab korraga mõlemat eelmist (*Mixed model*).

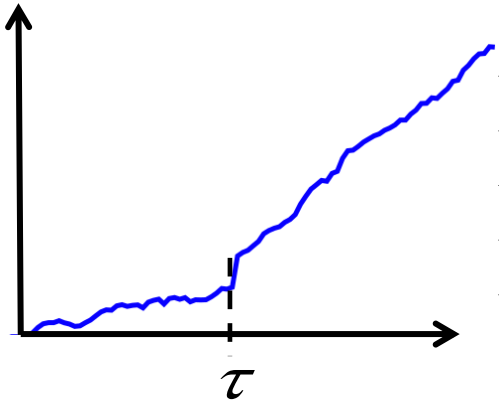
Nullhüpotees

$$H_{0C}: \quad y_t = a_0 + y_{t-1} + \mu_1 D_P + \mu_2 D_L + u_t$$

Ühikjuure protsess ühekordse impulsi ja muutuva triiviga.

$$D_P = \begin{cases} 1, & t = \tau + 1 \\ 0, & t \neq \tau + 1 \end{cases}$$

$$D_L = \begin{cases} 1, & t > \tau \\ 0, & t \leq \tau \end{cases}$$

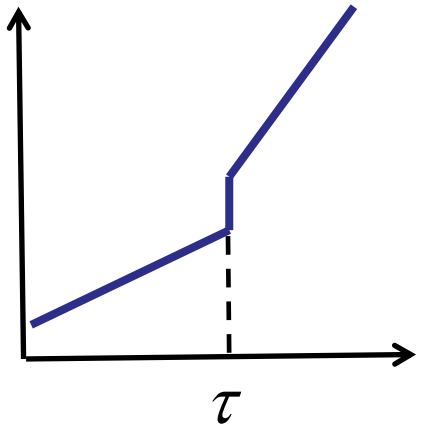


Sisukas hüpotees

$$H_{1C}: \quad y_t = a_0 + a_2 t + \mu_2 D_L + \mu_3 D_T + u_t$$

Deterministlik trend, mis korraga hüppab ja muudab tõusu.

$$D_T = \begin{cases} t - \tau, & t > \tau \\ 0, & t \leq \tau \end{cases}$$



Mudel, mida tuleb hinnata ja mille kordajat a_1 testida (kas a_1 =1)

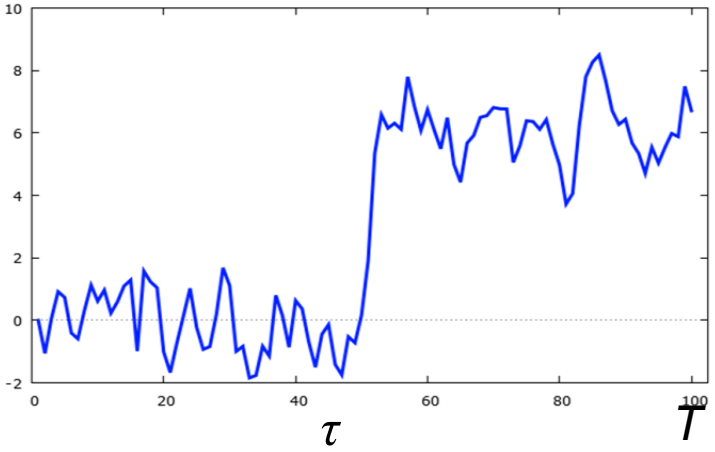
$$y_t = a_0 + a_1 y_{1-1} + a_2 t + \mu_1 D_P + \mu_2 D_L + \mu_3 D_T + u_t$$

Perroni protseduuri kriitilised väärtused

Perron viis läbi simulatsioonid, genereerimaks kriitilisi väärtusi, ja näitas, et a_1 väärtuste jaotus sõltub sellest, kus toimus hüpe.
Seda iseloomustab suhe

$$\lambda = \frac{\tau}{T}$$

τ aegrea pikkus kuni hüppeni
T terve aegrea pikkus



Kui *λ*=0 või *λ*=1, siis struktuurne muutus puudub ja kriitilised väärtused langevad kokku ADF testi kriitiliste väärtustega.

Tabelis on *t*-statistiku kriitilised väärtused mõningate λ väärtuste korral olulisuse nivool 5%. Täpsemalt vt Perroni artikkel tabelid IV.B, V.B ja VI.B.

λ	A	B	C
0,2	-3,77	-3,80	-3,99
0,5	-3,76	-3,96	-4,24
0,7	-3,80	-3,85	-4,18

Kui statistiku $$t = \frac{\hat{a}_1 - 1}{se(\hat{a}_1)}$$

väärtus on nullist kaugemal kui vastav kriitiline, on ühikjuurele vastav H_0 ümber lükatud.

Perron, P. (1989), The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis *Econometrica, The Econometric Society,* Vol. 57 , 1361-1401

Kui murdepunkt pole teada

- Perroni protseduuri kasutamine eeldab, et struktuurse muutuse toimumise hetk (murdepunkt) on teada.
- Võib hinnata visuaalselt, vaadates aegrea diagrammi, või tugineda ajaloolistele sündmustele.
- On olemas ka meetodeid murdepunkti statistiliseks määramiseks. Näiteks järgmistelt autoritelt
  - Amsler, C. and Lee, J. “An LM Test for a Unit Root in the Presence of Structural Change.” *Econometric Theory* 11, (June 1995), 359-68.
  - Zivot, E. and Andrews, D.W.K., 1992, “Further Evidence on the Great Crash, the Oil-Price Shock and the Unit Root Hypothesis.” *Journal of Business and Economic Statistics* 10, 251-270 (*Statas olemas lisamoodul*)
  - Perron, P. "Further Evidence on Breaking Trend Functions in Macroeconomic Variables.” *Journal of Econometrics* 80 (1997), 355 – 85.
  - Lee, J. and Strazicich, M. “Lagrange Multiplier Unit Root Test with Two Structural Breaks.” *Review of Economics and Statistics* 85, (Nov., 2003), 1082-1089.

Zivot-Andrewsi test, I

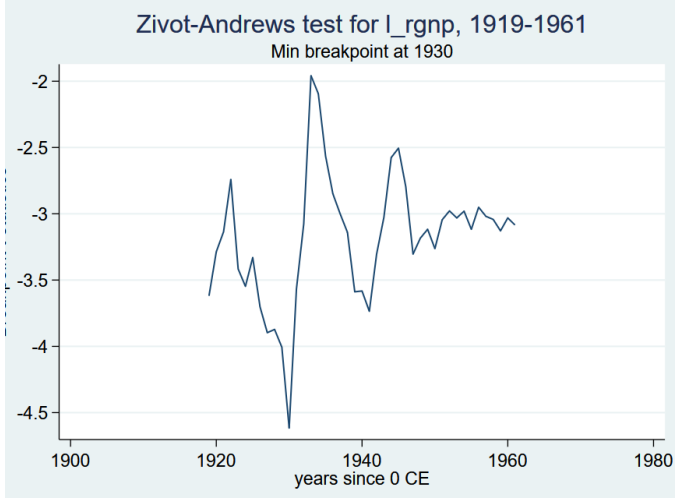
Erinevalt Perroni protseduurist, kus murdepunktis toimuv struktuurne muutus on eksogeenne, võimaldasid Zivot ja Andrews struktuurset muutust kui **endogeenset sündmust.**

- Murdepunkti otsimine käib aegrea andmete alusel. Otsitakse sellist kohta, kus nullhüpoteesi esinemine on kõige väiksema tõenäosusega, st teststatistik on nullist kõige kaugemal.

- Perroni protseduuri situatsioonidele A, B ja C vastavaid mudeleid hinnatakse erinevate murdepunktide korral vahemikus $t = 0 + r$ kuni $t = T - r$.
  - Vaikimisi r = 0,15T

- Iga potentsiaalse murdepunkti korral leitakse vastav *t*-statistik.

Näiteks Perroni artiklis kasutatud USA reaalse RKP logaritmi aegrea testimine perioodil 1909-1970. Testitud situatsiooni A. Esitatud on *t*-statistiku väärtuste graafik erinevate potentsiaalsete murdepunktide korral. Minimaalne on -4,6 aastal 1930.



Zivot, E. and D. W. K. Andrews, 1992, “Further Evidence on the Great Crash, the OilPrice Shock and the Unit Root Hypothesis.” *Journal of Business and Economic Statistics* 10, 251-270

Zivot-Andrewsi test, II

Leitakse koht, kus *t*-statistiku väärtus on minimaalne.
Kuvatakse ka kriitilised väärtused nivoodel 10%, 5% ja 1%.

Kui *t*-statistik on nullist kaugemal kui kriitiline, võtta vastu sisukas hüpotees: ühikjuur puudub.

Viitaegade arvu võib ette anda või lasta otsida optimaalset viitaegade arvu kas viitaegade statistilist olulisust vaadates (*t*-test) või informatsioonikriteeriumi AIC või BIC alusel.

Programmis Stata tuleb installeerida Baumi poolt loodud lisamoodul *zandrews*

Vaikimisi kasutatakse situatsiooni A testimist: break(intercept)
Saab testida ka situatsiooni B: break(trend)
või situatsiooni C: break(both)

Christopher F Baum, 2004. "ZANDREWS: Stata module to calculate Zivot-Andrews unit root test in presence of structural break," Statistical Software Components S437301, Boston College Department of Economics, revised 31 Jul 2015.

ÜHIKJUURE TESTIMINE PANEELANDMETE KORRAL

Paneelandmed ja ühikjuur

- Mikro-paneelandmete (nt isikud) korral on tavaliselt objektide arv *N* suur ja aegridade pikkus *T* väike.
- Kui objektide arv *N* on suur ja aegridade pikkus *T* väike, võib kasutada standardset paneelandmete hindamise protseduuri, ei ole vaja muretseda ühikjuure pärast.
  - Levin, Lin ja Chu simulatsioonid näitasid, et sellisel juhul alluvad teststatistikud piisavalt hästi jaotusseadustele, isegi väikese valimi korral.
- Erinevad makroökonoomika analüüsid on hakanud kasutama makro-paneelandmeid, kus aegridade pikkus *T* on suur:
  - majanduskasvu konvergents;
  - ostujõu pariteet;
  - rahvusvahelised teadus- ja arendustegevuse kõrvalmõjud.
- Sellest tulenevalt on tekkinud probleem võimalikust ühikjuure esinemisest paneelandmetes ja selle testimisest.
- Mittestatsionaarsuse tõttu ei pruugi erinevad teststatistikud alluda standardsetele jaotusseadustele.
  - Entorf (1997) leidis, et kui $T \to \infty$ ja N on fikseeritud, võib FE mudeli *t*-statistikute põhjal otsuse vastuvõtmine olla ekslik: näiv regressioon.

Miks paneelandmete ühikjuure testid?

Paneelandmete korral igal objektil oma aegread. Aegridade arv = objektide arv × tunnuste arv.

Võib ju testida iga objekti jaoks eraldi (ADF, PP, KPSS).

AGA
lõpliku valimi korral on paneelandmete ühikjuure testidel võimsus suurem kui objektide kaupa (individuaalselt), sest aegrea varieerumisele lisandub varieerumine üle objektide.

See võimaldab leida täpsemaid parameetrite hinnanguid.

Võimsus suurem, st II liiki vea tõenäosus väiksem.
II liiki viga: mittekehtiva nullhüpoteesi vastuvõtmine (testi tulemus ühikjuur, kuigi tegelikult pole).

Taylor, M.P. and Sarno, L., 1998. The behavior of real exchange rates during the post-Bretton Woods period. *Journal of International Economics*, *46*(2), pp.281-312.

Ühikjuure testid paneelandmete korral

Esimene generatsioon.

Eeldavad ristkorrelatsiooni puudumist.

- **Levin-Lin-Chu** ehk LLC
- **Im-Pesaran-Shin** ehk IPS
- Harris-Tzavalis
- Breitung
- Fisheri tüüpi (*Fisher type*)
- **Hadri** (KPSS testi üldistus)

Järgnevalt vaatame lähemalt esimest kahte ja viimast.

Levin-Lin-Chu test

Lähtutakse ADF testist, mis iga *i*-nda objekti jaoks

$$\Delta y_{i,t} = \alpha_i + \theta_t + \beta_i t + \delta_i y_{i,t-1} + \sum_{j=1}^{p_i} \gamma_j \Delta y_{i,t-j} + v_{i,t}, \quad t = 1,2,\ldots,T; i = 1,2,\ldots,N$$

Peale teisendusi hinnatakse üht, ühendatud regressiooni.
Seega eeldatakse, et autoregressiivne kordaja on homogeenne üle erinevate objektide *i :*

$$\delta_1 = \delta_2 = \cdots = \delta_N = \delta$$

Peamine puudus

$H_0: \quad \delta_i \equiv \delta = 0 \quad (\rho_i \equiv \rho = 1) \quad \forall i$ kõikidel aegridadel on ühikjuur

$H_1: \quad \delta_i \equiv \delta < 0 \quad (\rho_i \equiv \rho < 1) \quad \forall i$ kõikidel aegridadel ei ole ühikjuurt

- Eeldatakse, et puudub korrelatsioon ristandmete vahel.
- Saab kasutada vaid balansseeritud paneelandmete korral.
- Sobib kasutada, kui aegread on pikad, aga objekte vähe, sest eeldab, et suhe $\frac{N}{T} \to 0$

Levin, A., Lin, C.-F., and J. Chu, J. (2002) “Unit root tests in panel data: asymptotic and finite-sample properties”, *Journal of Econometrics* 108: 1-24.

Im, Pesaran ja Shin test, I

Ei nõuta, et autoregressiivne parameeter δ oleks kõikide objektide aegridades ühine, homogeenne.

Lähtutakse ADF regressioonist kujul

$$\Delta y_{i,t} = \delta_i y_{i,t-1} + z_{i,t}\gamma_i + \varepsilon_{i,t}$$

iid, on lubatud dispersioonide heterogeensus üle objektide

On erinevatel objektidel erinev

1) Iga objekti *i* korral
   - Hinnatakse seda mudelit eraldi.
   - Leitakse parameetri δ_i *t*-statistik t_i
2) Leitakse *t*-statistikute keskmine üle kõikide objektide.

IPS testi saab kasutada ka balansseerimata paneelandmete korral.

Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics,* 115(1), 53-74.

Im, Pesaran ja Shin test, II

Nullhüpotees: kõikide objektide aegridadel on ühikjuur, st $\delta_i = 0 \quad \forall i$ korral.

Sisukas hüpotees: osade objektide aegridadel ühikjuur puudub.

$$H_1 = \begin{cases} \delta_i < 0 & i = 1, 2, \ldots N_1 \\ \delta_i = 0 & i = N_1 + 1, \ldots N \end{cases}$$

St neid aegridu, kus ühikjuur puudub, on nullist erinev arv.

Kui sisuka hüpoteesi vastuvõtmisel tahame teada, milliste objektide aegridadel ühikjuur puudub, tuleb aegridu eraldi testida (ADF, PP, KPSS).

Erinevad teststatikud

t-bar $\quad \bar{t} = \frac{1}{N} \sum_{i=1}^{N} t_{\delta_i}$

kus iga individuaalse aegrea *i* korral on leitud nullhüpoteesi $\delta_i = 0$ testimiseks leitud teststatistik t_{δ_i}.
See statistik eeldab, et *N* ja *T* on fikseeritud.

Lisaks leitakse korrigeeritud statistikud:

t-tilde-bar $\quad \bar{t}$ Arvutamisel kasutatakse DF regressiooni vealiikmete dispersioonist erinevat hinnangut.

Z-t-tilde-bar $\quad Z_{\bar{\tilde{t}}}$ Allub standardiseeritud normaaljaotusele.

NÄIDE: raha nõudlust mõjutavad tegurid

6 riiki, aastased andmed 1975-2002.

Enne mudeli hindamist viidi läbi IPS testid.

Diferentsimata suuruste korral ühikjuur esines, H_0

Diferentside Δ korral ühikjuurt ei esinenud, H_1

Otsus: paneelandmete mudelis kasutas diferentse.

TABLE 1. The IPS Unit Root Test Results

Variables	Intercept only	q (optimal lag length)	Both intercept and trend	q (optimal lag length)
$Ln(M2_i/P_i)_t$	0.512 [0.696]	1	-0.345 [0.365]	4
$\Delta Ln(M2_i/P_i)_t$	-11.183 [0.000]	0	-9.779 [0.000]	3
$Ln(Y_i)_t$	-0.882 [0.189]	4	0.723 [0.765]	4
$\Delta Ln(Y_i)_t$	-10.180 [0.000]	0	-9.266 [0.000]	3
$(RD_i - RL_i)_t$	-1.120 [0.131]	2	-0.827 [0.204]	2
$\Delta(RD_i - RL_i)_t$	-11.479 [0.000]	1	-8.908 [0.000]	1
$\Delta Ln(P_i)_t$	-6.098 [0.000]	1	-5.102 [0.000]	1
$Ln(REER_i)_t$	-0.727 [0.234]	1	0.498 [0.691]	1
$\Delta Ln(REER_i)_t$	-5.761 [0.000]	0	-4.451 [0.000]	3
$Ln(R_{US})_t$	-0.451 [0.326]	0	0.761 [0.777]	0
$\Delta Ln(R_{US})_t$	-3.761 [0.000]	0	-2.564 [0.005]	0

Valadkhani, A. (2006). What Determines the Demand for Money in the Asian-Pacific Countries? An Empirical Panel Investigation, Department of Economics, University of Wollongong, https://ro.uow.edu.au/commwkpapers/147

Hadri LM test

Kui Im-Pesaran-Shin testi korral on nullhüpotees ümber lükatud, siis vähemalt ühe objekti aegreal ühikjuurt ei ole. Aga kuidas testida, kas kõikidel ühikjuur puudub?

KPSS testi arendus paneelandmete korral: Hadri test

Objektide aegread kujul $y_{it} = \beta_i t + r_{it} + e_{it}$

deterministlik trend

juhuslik ekslemine

statsionaarne juhuslik komponent

kus juhusliku ekslemise komponent $r_{it} = r_{it-1} + u_{it}$

Testitakse nullhüpoteesi, et dispersioon Var(u_{it}) =0.

H_0 korral ei ole r_t juhuslik ekslemine ja tegemist on lihtsa trend-statsionaarse mudeliga

$$\mathrm{H}_0: \quad y_{it} = \beta_i t + r_{i0} + e_{it} \qquad \forall i$$

H_0: kõikide objektide aegread on (trend-)statsionaarsed.
H_1: vähemalt ühe objekti aegrida ei ole (trend-)statsionaarne, esineb ühikjuur.

Dispersiooni testimiseks kasutatakse LM (*Lagrange multiplier*) teststatistikut.

Hadri, K. (2000). Testing for stationarity in heterogeneous panel data. *Econometrics Journal* 3: 148-161.

Ühikjuure testid paneelandmete korral kui esineb ristkorrelatsioon

Teine generatsioon ei eelda ristkorrelatsiooni puudumist.

- CIPS test
  - Pesaran, M.H. (2007). A simple panel unit root test in the presence of cross section dependence. *Journal of Applied Econometrics* 27: 265–312.
  - STATA-s vajalik lisamoodul xtcips

- PANIC test
  - J. Bai and S. Ng, (2004) A PANIC attack on unit roots and cointegration, *Econometrica* 72, 1127-1177.
  - STATA-s vajalik lisamoodul xtpanicca

Näiv regressioon paneelandmete korral

Paneelandmete korral on näiva regressiooni oht väiksem.

- Kui $N \to \infty$ ja $T \to \infty$, saadakse mittestatsionaarsete paneelandmete korral mõjusad hinnangud.
  - Sest hinnangud leitakse üle objektide keskmiste.
- Kui aga N on fikseeritud ja $T \to \infty$, ehk suhe $\frac{N}{T} \to \infty$, on olemas näiva regressiooni oht.
  - Siis tuleb kindlasti pöörata tähelepanu aegridade statsionaarsusele ja viia läbi ühikjuure testimine.
- Kui ajaperioodide arv T väike ja objektide arv N suur, võib kasutada tavalist paneelandmete protseduuri ja statsionaarsuse testimist pole vaja (vt Baltagi).

Kao, C. (1999). Spurious regression and residual-based tests for cointegration in panel data. *Journal of econometrics*, *90*(1), 1-44.
Phillips, P. C., & Moon, H. R. (1999). Linear regression limit theory for nonstationary panel data. *Econometrica*, *67*(5), 1057-1111.
Baltagi (2005) „Econometric Analysis of Panel Data“, lk 241.