Časové rady - úvod

Beáta Stehlíková

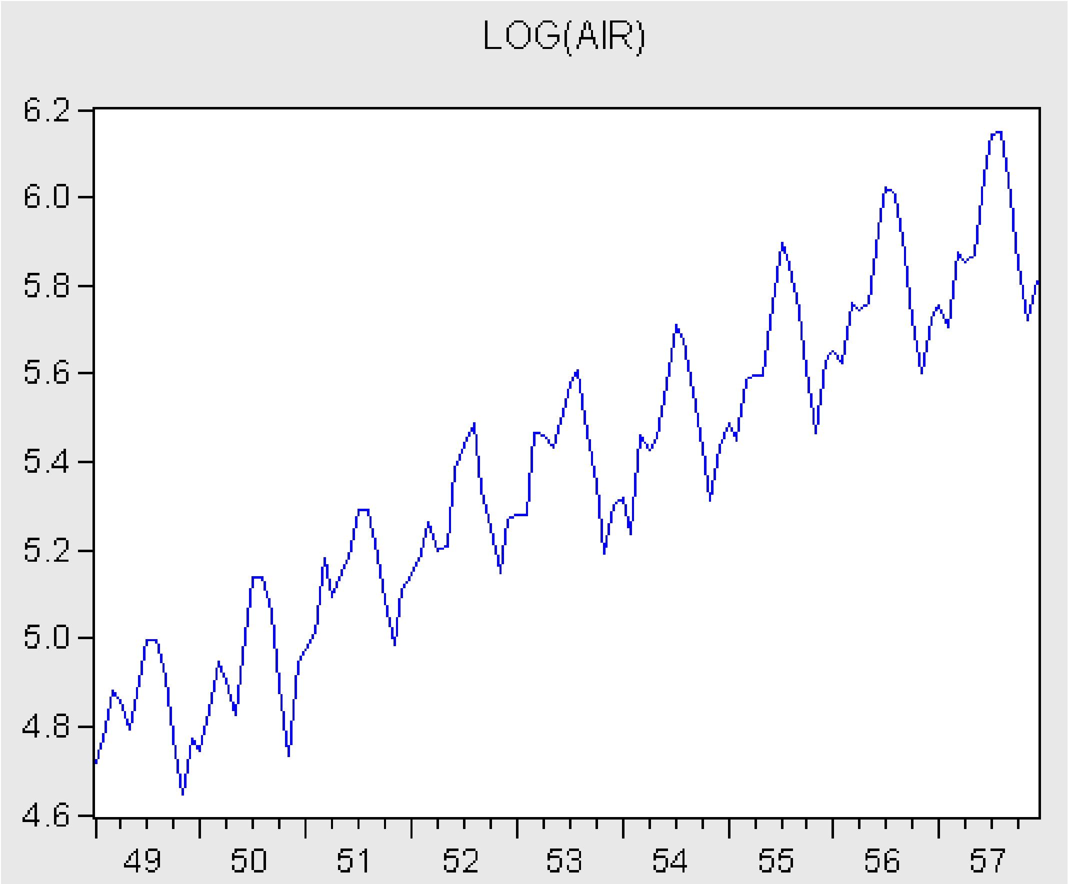
Časové rady, FMFI UK, 2012/2013

I.

Analýza časových radov - úvod

Analýza časových radov

- Máme mesačné dáta - počty cestujúcich aerolinkami:

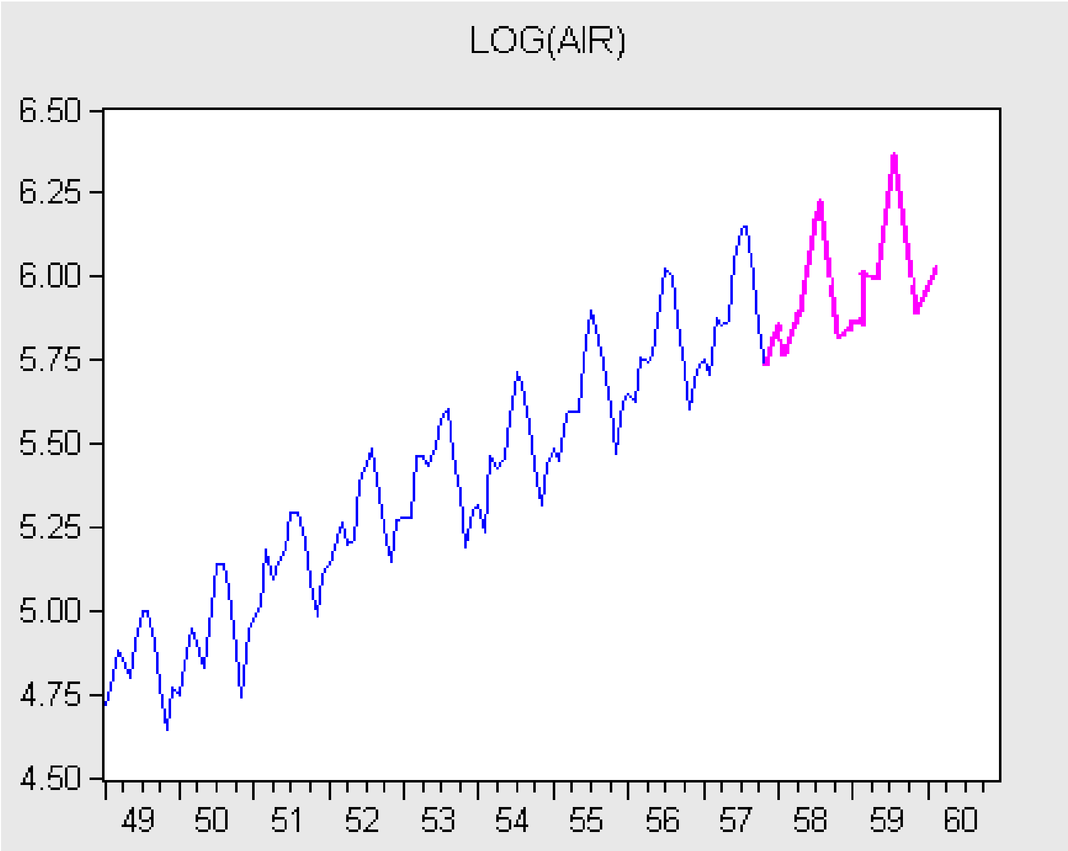


G. E. P. Box, G. M. Jenkins: **Time Series Analysis: Forecasting and Control**.

- Otázka: aký bude ďalší vývoj?

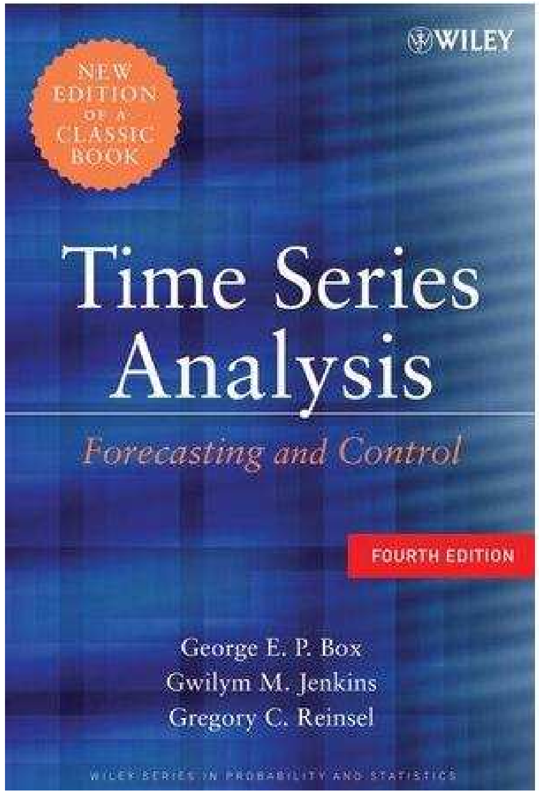
Analýza časových radov

- Intuitívne: zachová sa rastúci trend a sezónnosť (ak nenastane nejaký šok)



- Ako to vyjadriť kvantitatívne? Ako určiť presnosť odhadov, ako zostrojiť interval spoľahlivosti?

Box a Jenkins



Time Series Analysis: Forecasting and Control, 4th Edition
George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel

ISBN: 978-0-470-27284-8

Hardcover
784 pages
July 2008

"A modernized new edition of one of the most trusted books on time series analysis. Since publication of the first edition in 1970, Time Series Analysis has served as one of the most influential and prominent works on the subject."

http://eu.wiley.com

Box a Jenkins

"The first paper you wrote with Jenkins has been considered as a breakthrough in statistics. How do you become interested in time series?"

Rozhovor s G. E. P. Boxom po oslave jeho 80. narodenín

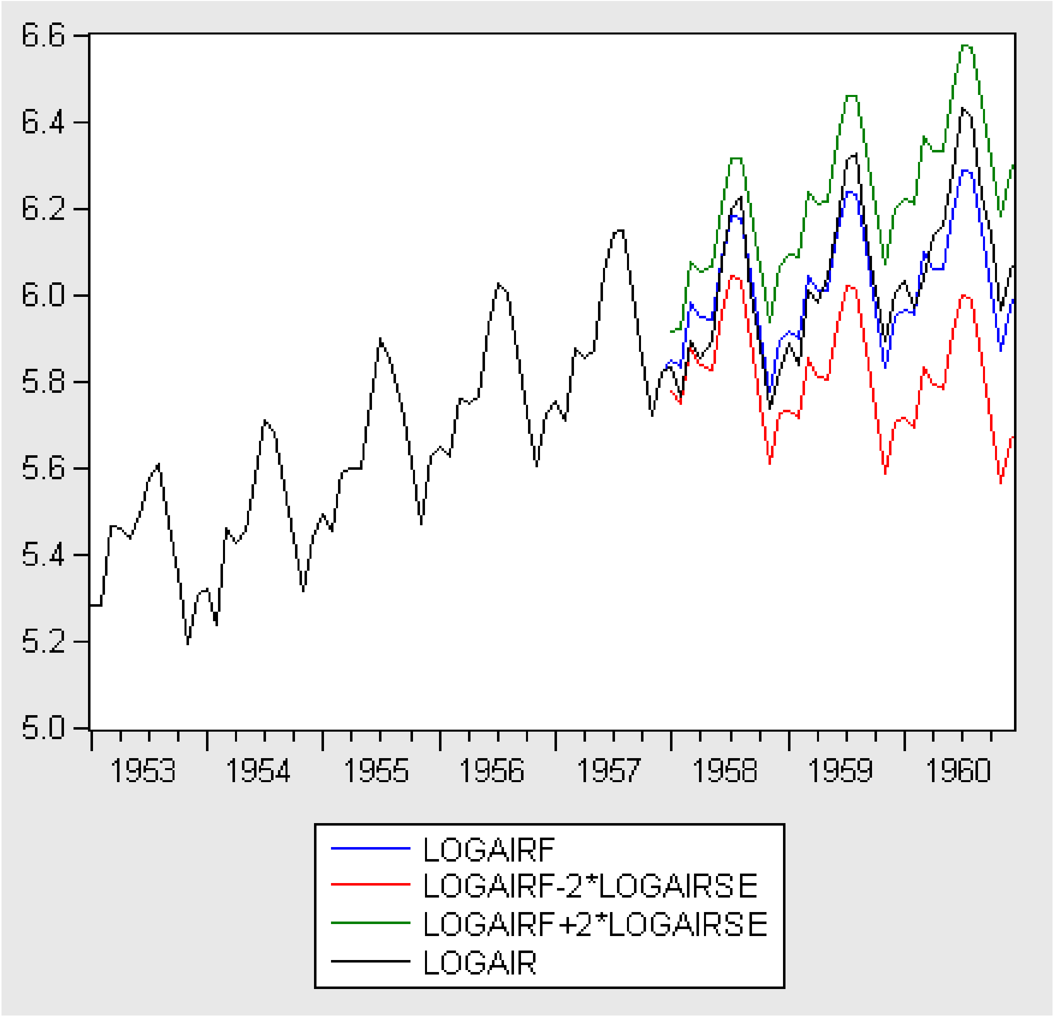
http://halweb.uc3m.es/esp/Personal/personas/dpena/articles/boxIJFinter4.PDF

Historický vývoj

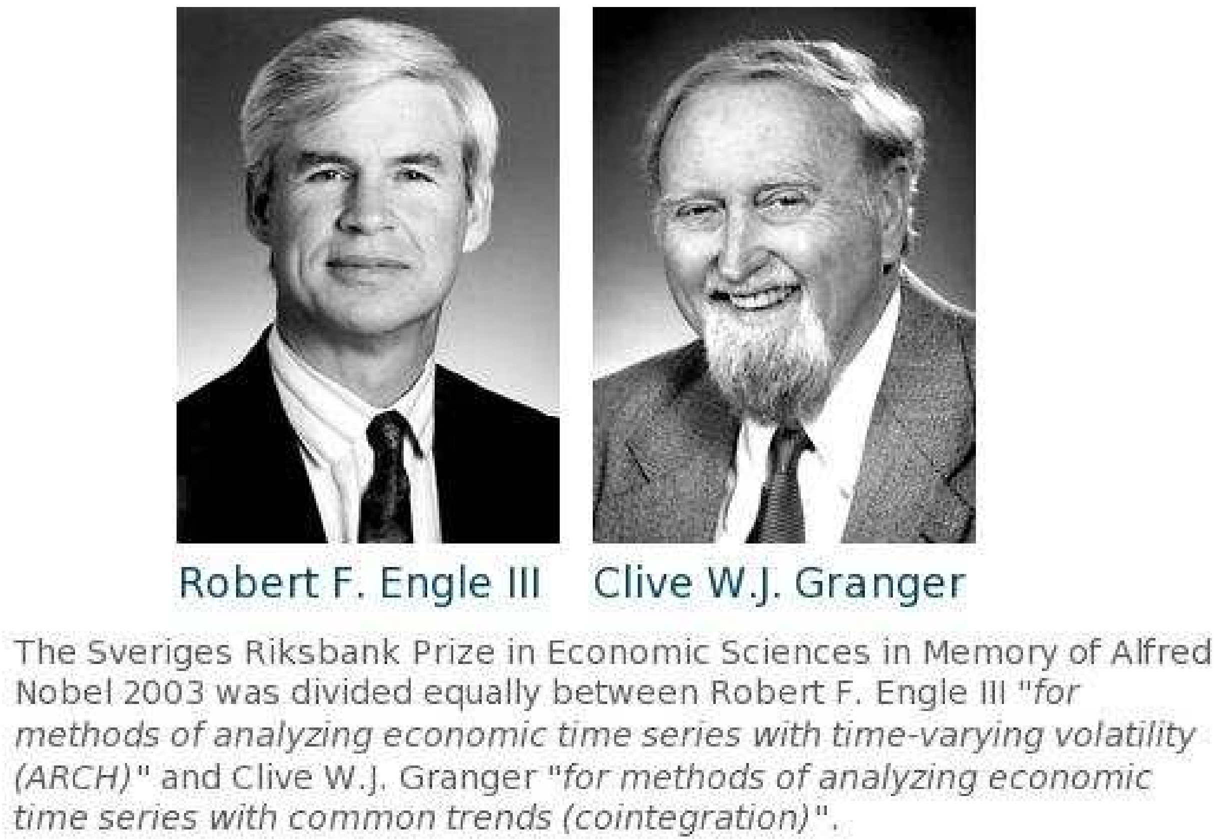
- Časové rady v ekonometrii, bez zohl'adnenia závislosti pozorovaní
- Cochrane & Orcutt (1949): dôsledok korelovaných rezíduí v regresii
- Durbin & Watson (1950-51): test na identifikáciu autokorelácie prvého rádu v rezíduách
- Box & Jenkins (1970) : jednorozmerné modely pre časové rady - OBSAH TOHTO KURZU
- Granger & Newbold (1975): tieto jednorozmerné modely dávali často lepšie predikcie ako vel'ké ekonometrické modely (niekol'ko stoviek rovníc)

Ukážka výstupu modelu

- Úvodný príklad o počte cestujúcich - predikcie, intervaly spoľahlivosti a porovnanie s realitou:



Historický vývoj



http://www.nobelprize.org

- ARCH model a jeho zovšeobecnenia, kointegrácia - na výberových predmetoch

Zaujímavosť

- Manželka R. Engla sa narodila v Prešove, v r. 2005 spolu navštívili Prešov

Rob, Marianne, Jordan and Lindsey in Williamstown 2002.

http://www.nobelprize.org/nobel_prizes/economics/laureates/2003/engle-autobio.html

http://www.presov.sk/portal/?c=12&id=3590

II.

Časové rady - základné pojmy

Základné pojmy - obsah prednášky

- Časový rad, momenty
- Stacionarita, ergodicita
- Biely šum
- Woldova reprezentácia
- Autokorelačná funkcia, testy o autokorelačnej funkcii
- Operátor posunu

Časový rad, momenty

- Stochastický proces $x_1, \ldots, x_T$ - úplne je charakterizovaný T-rozmernou distribučnou funkciou
- Obvykle sa zameriavame len na prvé dva momenty:
  - stredná hodnota $E[x_t]$
  - variancia $Var[x_t]$
  - kovariancie $Cov[x_t, x_s]$, nazývajú sa autokovariancie

Stacionarita a ergodicita

- Väčšinou máme len jeden časový rad - jednu realizáciu náhodného procesu $\rightarrow$ aby sa dala robiť štatistická inferencia, potrebujeme dodatočné predpoklady
- Napríklad: na to, aby sme odhadli strednú hodnotu, ... potrebujeme viac ako jednu realizáciu tejto náhodnej premennej
- Ergodický proces - výberové momenty počítané z časového radu s T pozorovaniami konvergujú pre $T \rightarrow \infty$ k zodpovedajúcim momentom
- Tento koncept má zmysel iba ak predpokladáme, že $E[x_t] = \mu$, $Var[x_t] = \sigma^2, \ldots$ pre $\forall t$

Stacionarita a ergodicita

- Silná stacionarita : združená distribučná funkcia sa nemení pri posune v čase
- Obvykle sa pracuje so slabším predpokladom → slabá stacionarita :

$$E[x_t] = \mu \ \forall t \tag{1}$$

$$Cov[x_t, x_s] = \gamma(|t-s|) \ \forall t, s \tag{2}$$

z (2) vyplýva: $Var[x_t] = const.$ pre $\forall t$

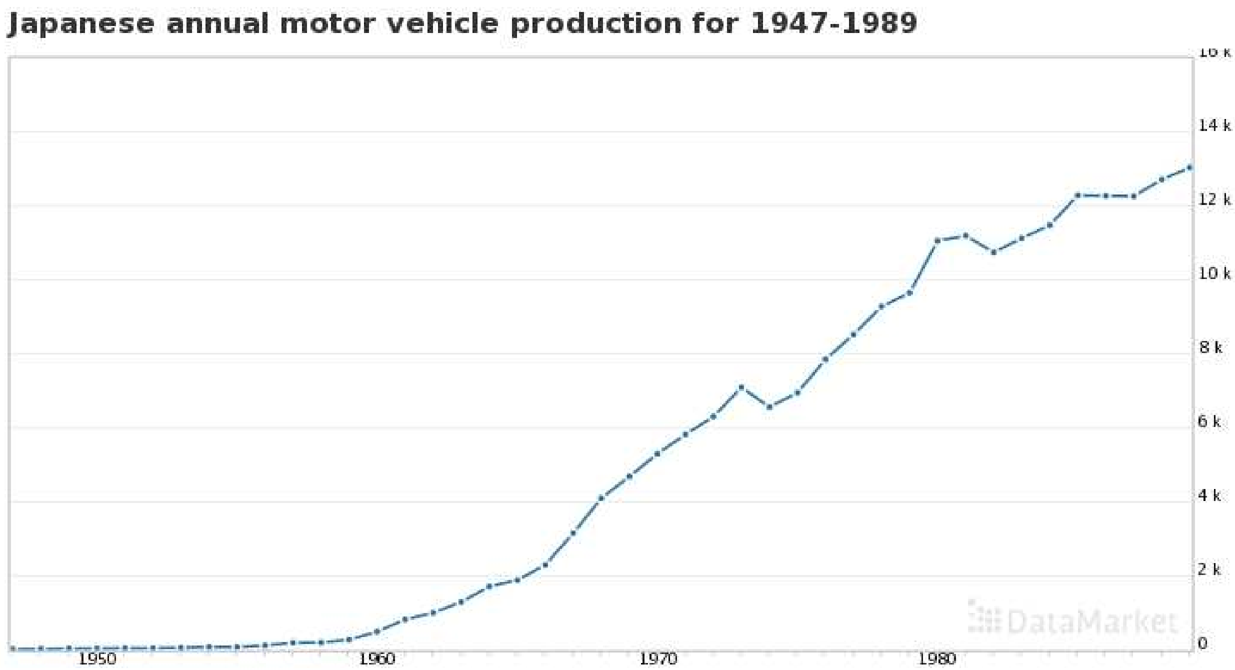
- Ďalej budeme pod stacionaritou rozumieť slabú stacionaritu

Stacionarita - dáta

- Stacionárny časový rad: dáta sú priťahované k určitej rovnovážnej hodnote, okolo ktorej oscilujú
- Nestacionárny časový rad: napríklad trend

Stacionarita - dáta

- PRÍKLAD 1

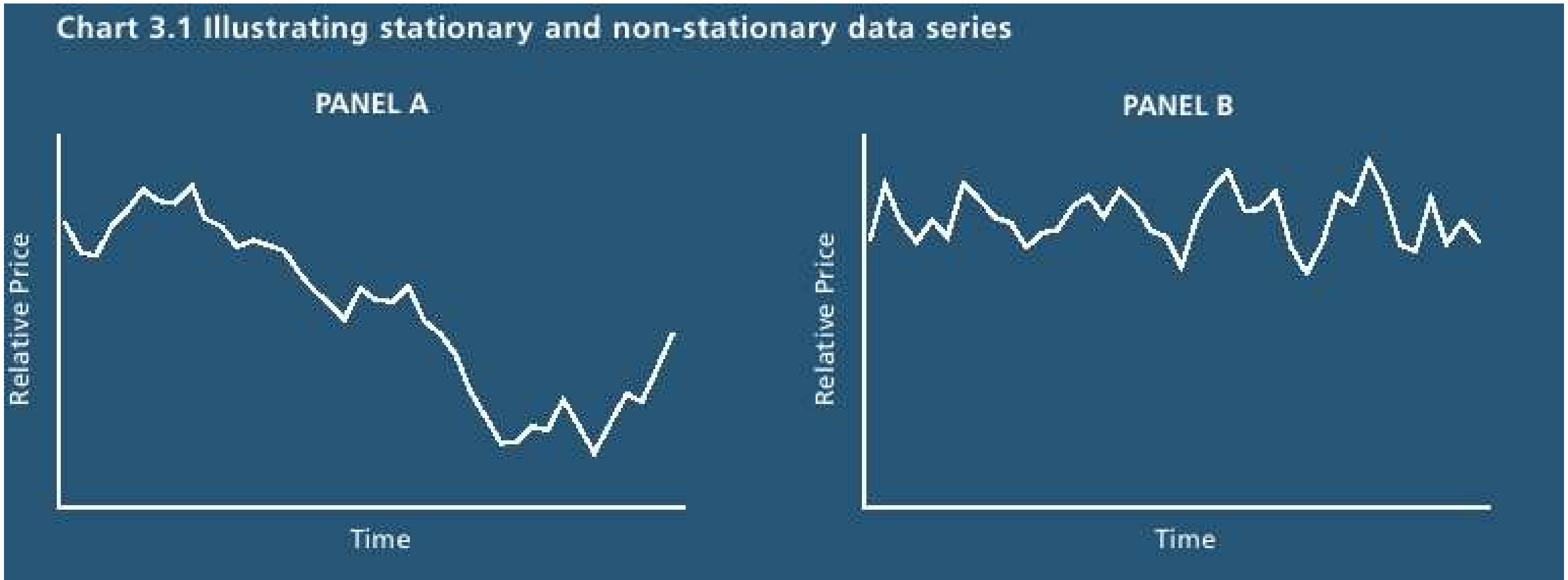


http://data.is/RVcjbL

Vidíme:
rastúci trend ⇒ stredná hodnota nie je konštantná ⇒ časový rad nie je stacionárny

Stacionarita - dáta

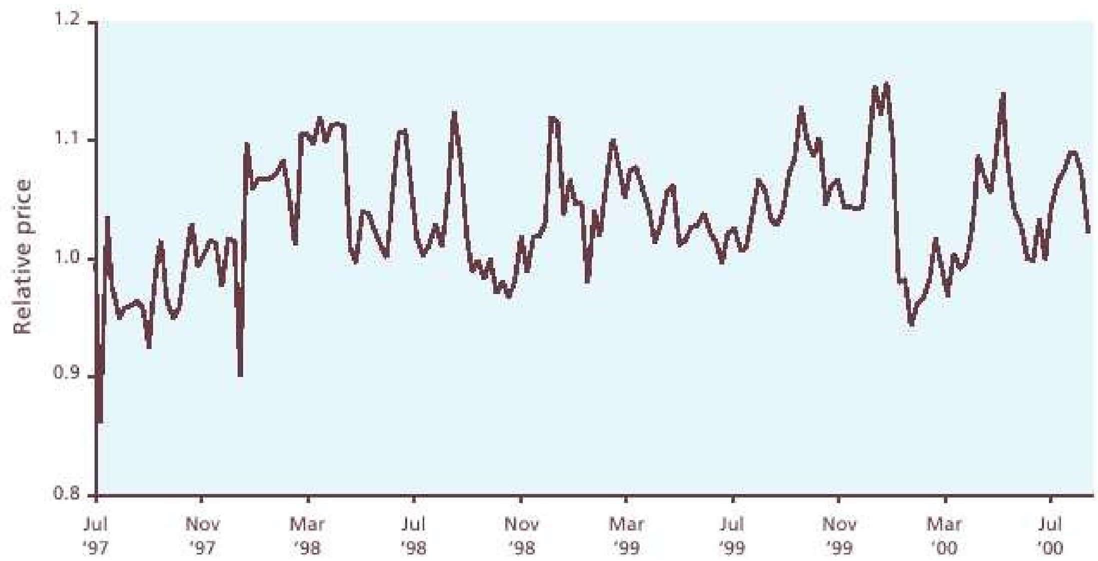
- PRÍKLAD 2
  Relatívna cena: ak sa dva tovary predávajú na tom istom trhu, ich relatívna cena by sa mala pohybovať okolo určitej rovnovážnej hodnoty



In Chart 3.1 above, Panel A shows a data series that is non-stationary – it shows no tendency to return to a stable value over the period. Panel B, on the other hand, shows a stationary data series – the series rapidly reverts to a constant long-run value over the period. If two products are in the same market then we would expect the relative price of these two products to behave in a similar fashion to the data series in Panel B.

Stacionarita - dáta

Chart 3.2: Price of Scottish Salmon relative to the price of Norwegian salmon in the UK



The above chart shows that the price of Scottish salmon relative to the price of Norwegian salmon in the UK appears to vary randomly around a constant long-run value, which suggests that the relative price is stationary. The econometric test for stationarity confirms that the relative price of Scottish salmon is stationary, which is what we would expect to observe if Scottish and Norwegian gutted salmon compete in the same product market in the UK.

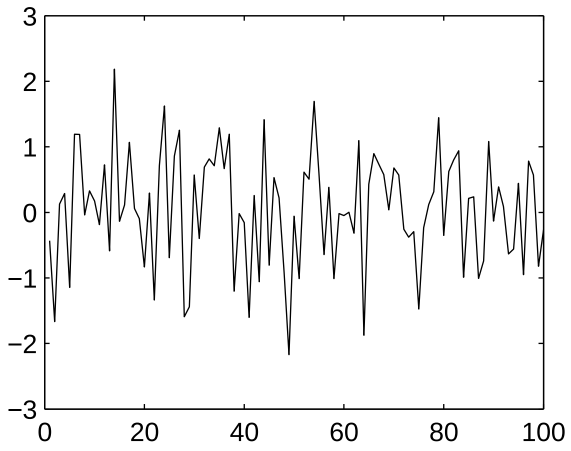
An Introduction to Quantitative Techniques in Competition Analysis, Lexecon, 2005

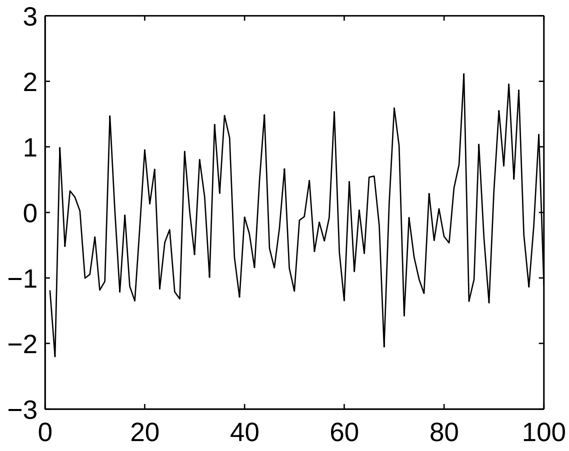
http://www.crai.com/ecp/publications/full_list.htm.

Biely šum

- Biely šum $\{u_t\}$ - dôležitý príklad stacionárneho procesu

$$
\begin{aligned}
E[u_t] &= 0 \qquad \forall t \\
Var[u_t] &= \sigma^2 \qquad \forall t \\
Cov[u_t, u_s] &= 0 \qquad \forall t \neq s
\end{aligned}
$$





- Základný náhodný proces, pomocou ktorého budeme definovať ďalšie, aj modely pre dáta

Stacionarita - príklad 1

- Nech u_t je biely šum, definujme

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

- Vypočítame:

$$E[x_t] = 0, \quad Var[x_t] = 2\sigma^2$$

$$Cov[x_t, x_s] = \begin{cases} \sigma^2 & \text{pre } |t-s| = 1 \\ 0 & \text{pre } |t-s| = 2, 3, \ldots \end{cases}$$

$\rightarrow$ proces je stacionárny

Stacionarita - príklad 2

- Nech u_t je biely šum; definujme

$$y_t = \begin{cases} u_1 & \text{pre } t = 1 \\ y_{t-1} + u_t & \text{pre } t = 2, 3, \dots \end{cases}$$

- y_t sa dá zapísať ako $y_t = \sum_{i=1}^{t} u_i$
- Vypočítame:

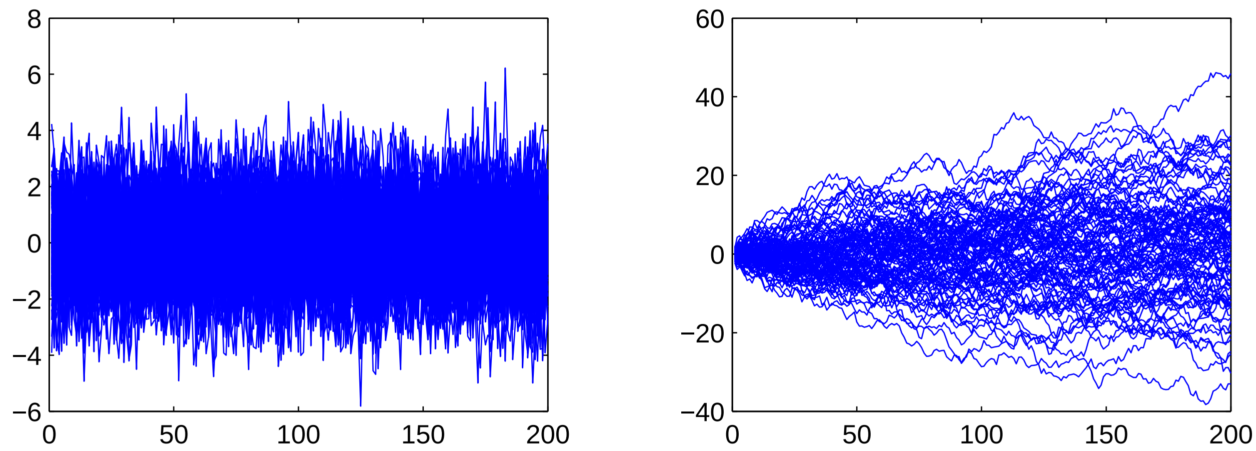
$$E[y_t] = 0, \;\; Var[y_t] = t\, \sigma^2$$

$$Cov[y_t, y_s] = \sigma^2 \, \min(t, s)$$

$\longrightarrow$ proces nie je stacionárny

Stacionarita - príklady 1, 2

- Porovnanie trajektórií procesov z predchádzajúcich dvoch príkladov:
  - ◇ vl’avo - stacionárny proces (pr. 1)
  - ◇ vpravo - nestacionárny proces (pr. 2), vidíme rastúcu disperziu



Stacionarita - príklad 3

- Nech $\{u_t\}_{t=-\infty}^{\infty}$ je biely šum, definujme

$$x_t = \mu + \sum_{j=0}^{\infty} \psi_j u_{t-j}, \tag{3}$$

pričom koeficienty ψ_j spĺňajú $\sum_{j=0}^{\infty} \psi_j^2 < \infty, \psi_0 = 1$

- Vypočítame:

$$E[x_t] = \mu, \;\; Var[x_t] = \sum_{j=0}^{\infty} \psi_j^2$$

$$Cov[x_t, x_{t+k}] = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{k+j}$$

$\rightarrow$ proces je stacionárny

Woldova reprezentácia

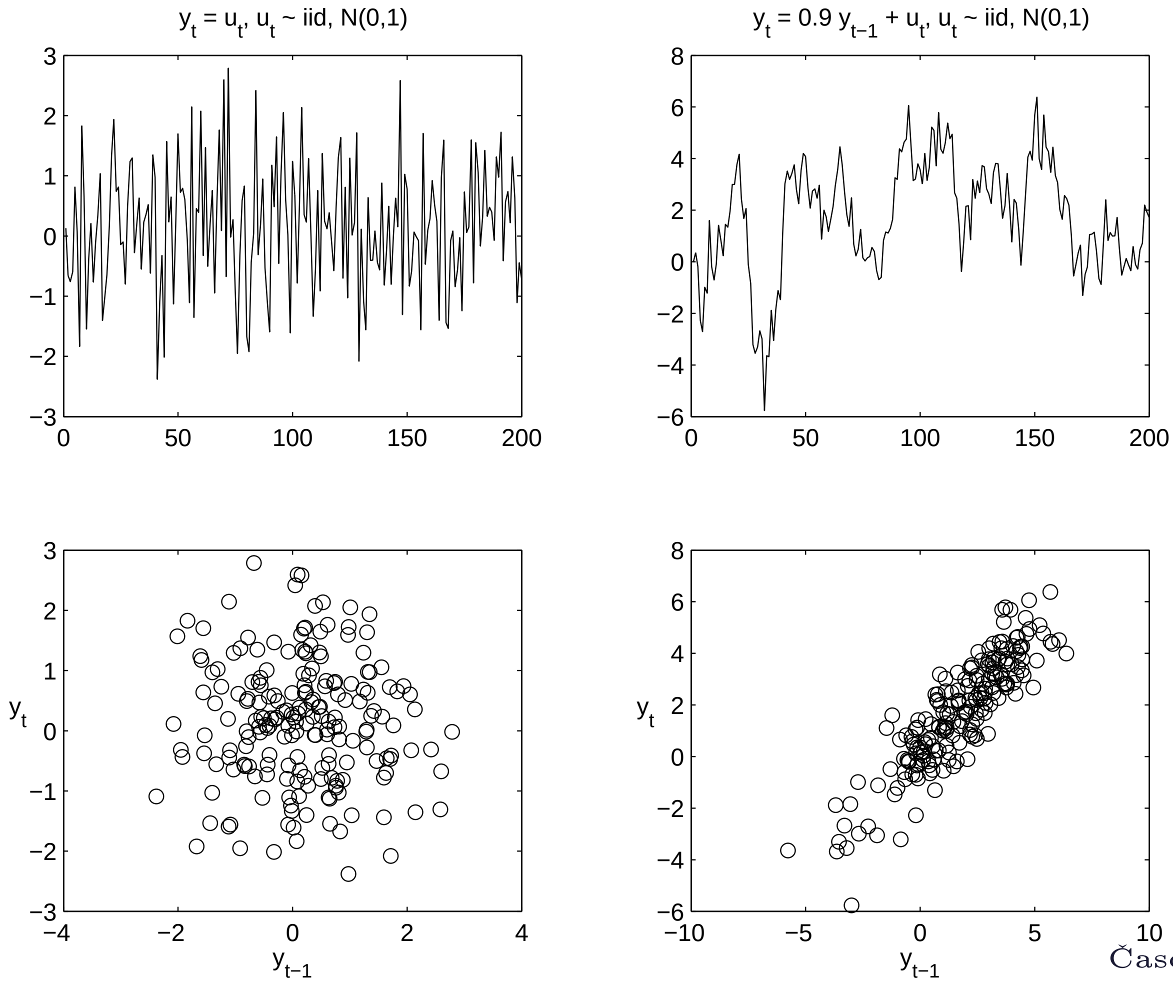
- Predchádzajúci príklad: proces x_t, ktorý má tvar (3), je stacionárny
- Dá sa dokázať':
  Každý stacionárny proces x_t sa dá zapísať' v tvare (3), t. j.

$$x_t = \mu + \sum_{j=0}^{\infty} \psi_j u_{t-j},$$

  pričom koeficienty ψ_j spĺňajú $\sum_{j=0}^{\infty} \psi_j^2 < \infty, \psi_0 = 1$ a u_t je biely šum
- Toto vyjadrenie stacionárneho procesu sa nazýva Woldova reprezentácia (Wold, 1938)

Autokorelačná funkcia - motivácia

- Vľavo: proces $y_t = u_t$
  Vpravo: proces $y_t = 0.9y_{t-1} + u_t$
- Realizácia procesu a ako závisí y_t od y_{t-1}



Autokorelačná funkcia

- Autokorelačná funkcia (ACF) stacionárneho procesu:

$$\rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)}$$

t. j.

$$\rho(\tau) = Cor(x_t, x_{t+\tau}) = \frac{Cov(x_t, x_{t+\tau})}{\sqrt{Var(x_t)\,Var(x_{t+\tau})}}$$

- Platí:

$$\rho(0) = 1, \quad \rho(\tau) = \rho(-\tau)$$

$\rightarrow$ stačí nám počítať $\rho(\tau)$ pre $\tau = 1, 2, \ldots$

Autokorelačná funkcia

- Ergodický proces $\rightarrow$ stredná hodnota, disperzia a kovariancie sa dajú konzistentne odhadnúť z dát $x_1, \ldots, x_T$:

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} x_t, \quad \hat{\gamma}(0) = \frac{1}{T}\sum_{t=1}^{T}(x_t - \hat{\mu})^2$$

$$\hat{\gamma}(\tau) = \frac{1}{T}\sum_{t=1}^{T-\tau}(x_t - \hat{\mu})(x_{t+\tau} - \hat{\mu})$$

$\rightarrow$ konzistetný odhad autokorelačnej funkcie:

$$\hat{\rho}(\tau) = \frac{\hat{\gamma}(\tau)}{\hat{\gamma}(0)}$$

- je asymptoticky nevychýlený

Autokorelačná funkcia - testy

- Odhad ACF v prípade bieleho šumu:
  - asymptoticky nevychýlený
  - variancia $\approx 1/T$
  - $\Rightarrow$ približný 95 % interval spoľahlivosti: $\pm 2/\sqrt{T}$, často sa zobrazuje spolu s odhadnutými autokoreláciami
- V prípade stochastického procesu, pre ktorý $\rho(\tau) = 0$ pre $\tau > k$, pre tieto τ platí

$$Var[\hat{\rho}(\tau)] \approx \frac{1}{T}\left(1 + 2\sum_{j=1}^{k}\hat{\rho}(j)^2\right)$$

Autokorelačná funkcia - testy

- Testovanie, či daný časový rad je biely šum:
  1. interval spoľahlivosti $\pm 2/\sqrt{T}$ pre každú autokoreláciu samostatne
  2. testovanie nulovosti $\rho(1), \ldots, \rho(m)$ súčasne:
     - Box & Pierce, 1970: za platnosti H_0 asymptoticky platí

$$Q = T \sum_{j=1}^{m} \rho(j)^2 \sim \chi_m^2$$

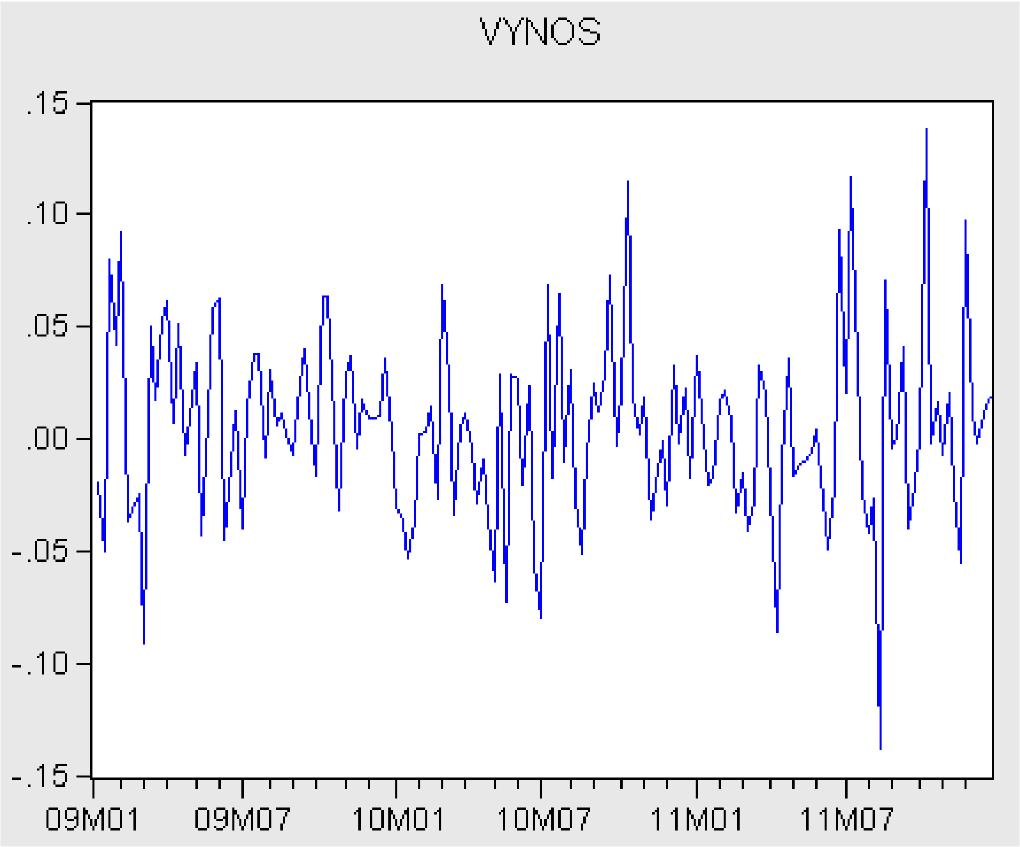
     - Ljung & Box, 1978: modifikácia s lepšími vlastnosťami pri menšom počte dát:

$$Q = T(T+2) \sum_{j=1}^{m} \frac{\rho(j)^2}{T-j} \sim \chi_m^2$$

     - počet stupňov voľnosti sa zmení, ak ide o rezíduá z modelu

Autokorelačná funkcia - príklad

- S_t - cena akcie GOOG, $vynos_t = \log(S_t/S_{t-\Delta t})$
- Týždenné výnosy, 2009 - 2011



Autokorelačná funkcia - príklad

- ACF výnosov v EViews:

Correlogram of VYNOS

Date: 08/21/12 Time: 13:11
Sample: 1/01/2009 1/04/2012
Included observations: 156

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.045	0.045	0.3279	0.567
		2	-0.001	-0.003	0.3280	0.849
		3	-0.032	-0.032	0.4890	0.921
		4	-0.162	-0.160	4.7606	0.313
		5	-0.091	-0.080	6.1146	0.295
		6	-0.056	-0.053	6.6222	0.357
		7	0.127	0.126	9.3109	0.231
		8	0.037	-0.000	9.5416	0.299
		9	0.021	-0.009	9.6151	0.383
		10	0.014	-0.007	9.6476	0.472
		11	-0.034	-0.004	9.8465	0.544
		12	0.015	0.042	9.8828	0.626
		13	-0.010	0.007	9.9018	0.702
		14	0.075	0.067	10.891	0.695
		15	0.065	0.055	11.619	0.708
		16	-0.020	-0.020	11.689	0.765
		17	0.066	0.074	12.454	0.772
		18	-0.080	-0.058	13.611	0.754

autokorelácie, číselné hodnoty tu
intervaly spoľahlivosti

Q štatistika
a P hodnota

Operátor posunu

- Operátor posunu (lag operator) L - užitočný pri práci s časovými radmi
- Vráti hodnotu procesu posunutú o jedno obdobie dozadu:

$$Lx_t = x_{t-1}$$

- Vlastnosti:
  - mocniny: $L^2 x_t = L(Lx_t) = x_{t-2}$
  - $L^0 = 1$ je identita: $(1 - L)x_t = x_t - x_{t-1}$
  - počítanie s mocninami: $L^2(L^3) = L^5$
  - násobenie: $(1 - 0.5L)(1 - 0.3L) = 1 - 0.8L + 0.15L^2$
  - ak c je konštanta, tak napr. $(1 - 0.1L + 2L^2)c = (1 - 0.1 + 2)c = 2.9c$