

Modelovanie sezónnosti

Beáta Stehlíková

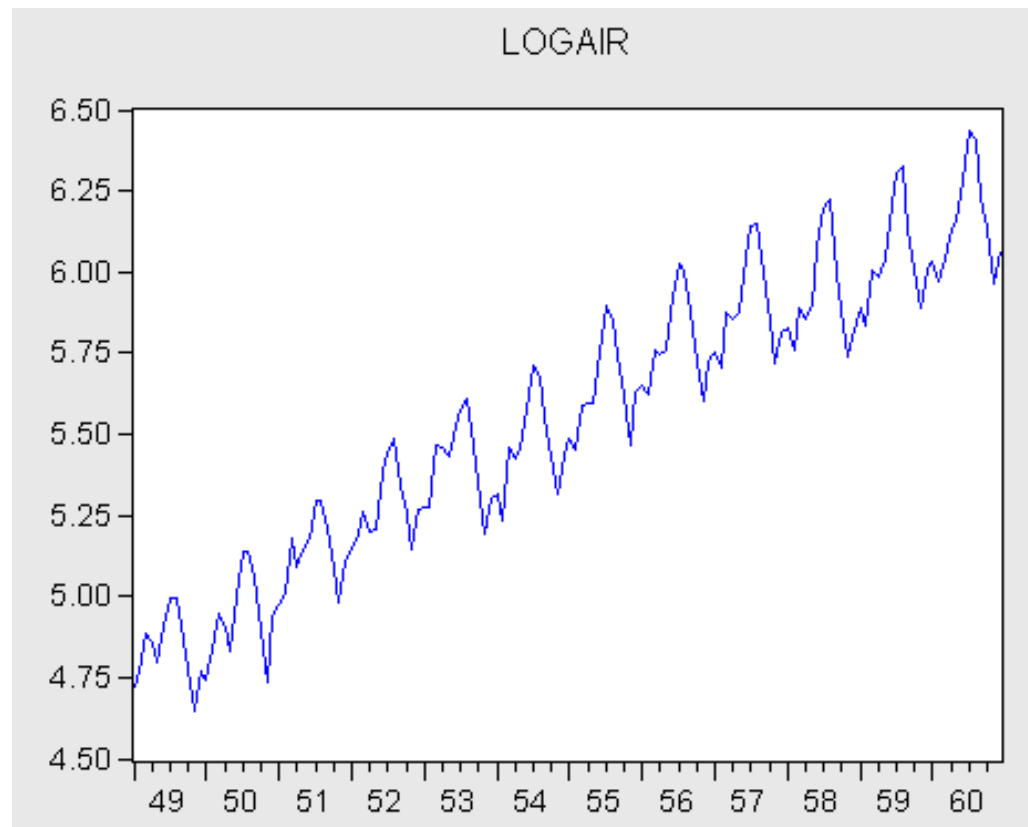
Časové rady, FMFI UK, 2012/2013

Modelovanie sezónnosti

- Mali sme už modely, ktoré mali sezónny charakter dát - AR(2) proces s komplexnými koreňmi.
- Nestačí to však na všetky sezónne časové rady
- Existujú modely špeciálne zamerané na modelovanie sezónnosti - **SARIMA modely** (sezónne ARIMA modely)

Príklad - dáta

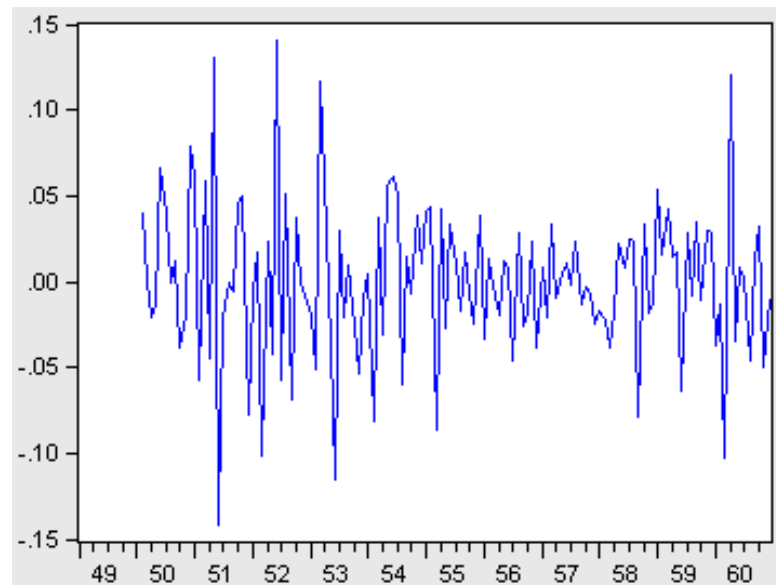
- Počet cestujúcich aerolinkami od Boxa a Jenkinsa - príklad od zakladateľov ARIMA modelovania
- Mesačné dáta, január 1949 - december 1960
- Pracujeme s logaritmami.



Príklad - diferencovanie

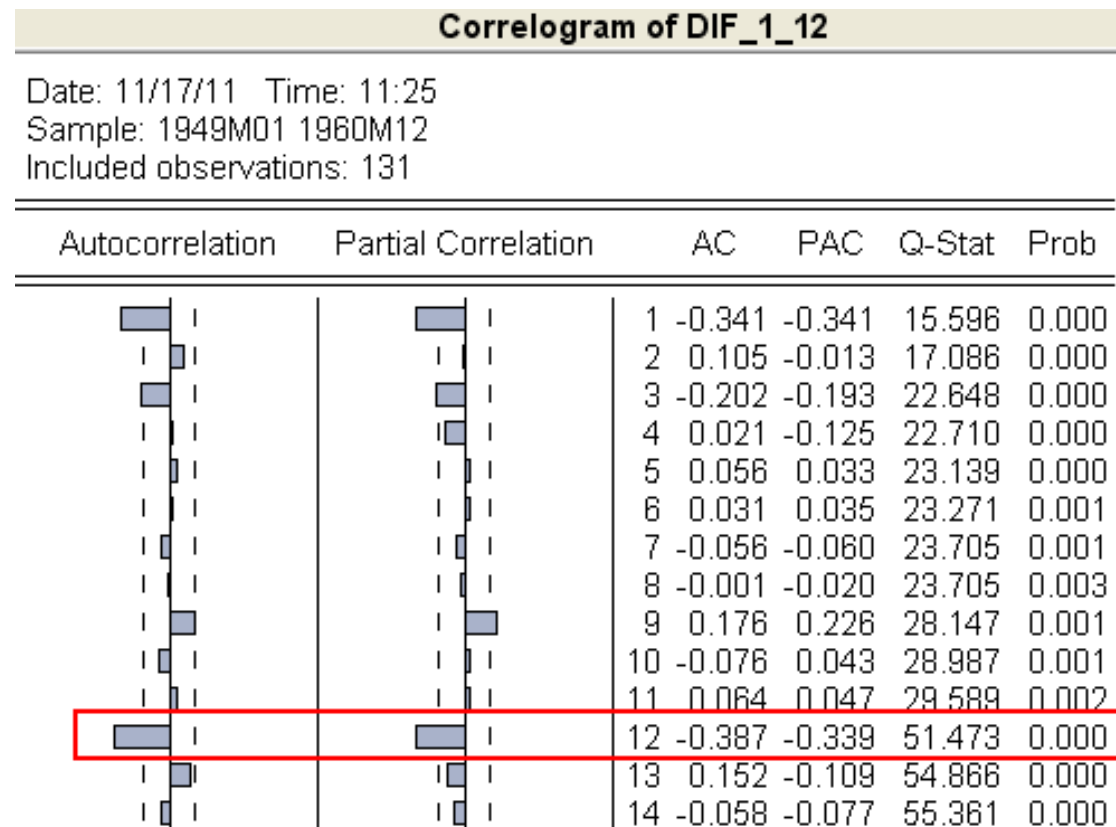
- Ročná sezónnosť, rastúci trend \rightarrow v takomto prípade sa robia
 - ◇ klasické diferencie $x_t - x_{t-1}$ - kvôli trendu
 - ◇ sezónne diferencie $x_t - x_{t-12}$ - kvôli ročnej sezónnosti
- Modelujeme teda premennú:

$$(1 - L)(1 - L^{12})x_t$$



Príklad - korelogram

- Výrazná hodnota ACF a PACF pri lagu 12:



Príklad - zostavenie modelu

- Mal by tam byť AR/MA člen rádu 12
- SARIMA modely:
 - ◇ nebudeme do modelu dávať samostatne x_{t-12} alebo všetky $x_{t-1}, x_{t-2}, \dots, x_{t-12}$
 - ◇ ale vynásobíme napríklad s AR(1) členom $(1 + \alpha L)$ - dostaneme $(1 + \alpha L)(1 + \theta L^{12})x_t$
 - ◇ vyššieho rádu, napríklad: $1 + \theta_1 L^{12} + \theta_2 L^{24}$
- Náš príklad:
 - ◇ Box & Jenkins navrhujú zobrať MA(1) člen (výrazná ACF pri lagu 1) a prenásobiť $(1 - \theta L^{12})$
 - ◇ dostaneme MA časť $(1 - \beta L)(1 - \theta L^{12})u_t$
 - ◇ takýto model sa dá často použiť na modelovanie sezónnych časových radov (samozrejme, nie je univerzálny), nazýva sa **airline model**

Príklad - odhadnutý model

- Výstup:

Dependent Variable: D(LOGAIR,1,12)
 Method: Least Squares
 Date: 11/17/11 Time: 11:28
 Sample (adjusted): 1950M02 1960M12
 Included observations: 131 after adjustments
 Convergence achieved after 8 iterations
 MA Backcast: 1949M01 1950M01

	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000448	0.000861	-0.520609	0.6035
MA(1)	-0.404629	0.080477	-5.027909	0.0000
SMA(12)	-0.640750	0.069172	-9.263115	0.0000
R-squared	0.372548	Mean dependent var		0.000291
Adjusted R-squared	0.362744	S.D. dependent var		0.045848
S.E. of regression	0.036600	Akaike info criterion		-3.754907
Sum squared resid	0.171463	Schwarz criterion		-3.689063
Log likelihood	248.9464	Hannan-Quinn criter.		-3.728152
F-statistic	37.99986	Durbin-Watson stat		1.932469
Prob(F-statistic)	0.000000			
Inverted MA Roots	.96	.83-.48i	.83+.48i	.48+.83i
	.48-.83i	.40	.00+.96i	-.00-.96i
	-.48+.83i	-.48-.83i	-.83-.48i	-.83+.48i
	-.96			

Príklad - odhadnutý model

- Stacionarita - po roznásobení je to MA model, ten je stacionárny
- Invertovateľnosť:

Inverse Roots of AR/MA Polynomial(s)
Specification: D(LOGAIR,1,12) C MA(1) SMA(12)
Date: 11/17/11 Time: 11:29
Sample: 1949M01 1960M12
Included observations: 131

MA Root(s)	Modulus	Cycle
-0.963587	0.963587	
-8.25e-17 ± 0.963587i	0.963587	4.000000
0.481793 ± 0.834490i	0.963587	6.000000
0.834490 ± 0.481793i	0.963587	12.000000
0.963587	0.963587	
-0.834490 ± 0.481793i	0.963587	2.400000
-0.481793 ± 0.834490i	0.963587	3.000000
0.404629	0.404629	

No root lies outside the unit circle.
ARMA model is invertible.

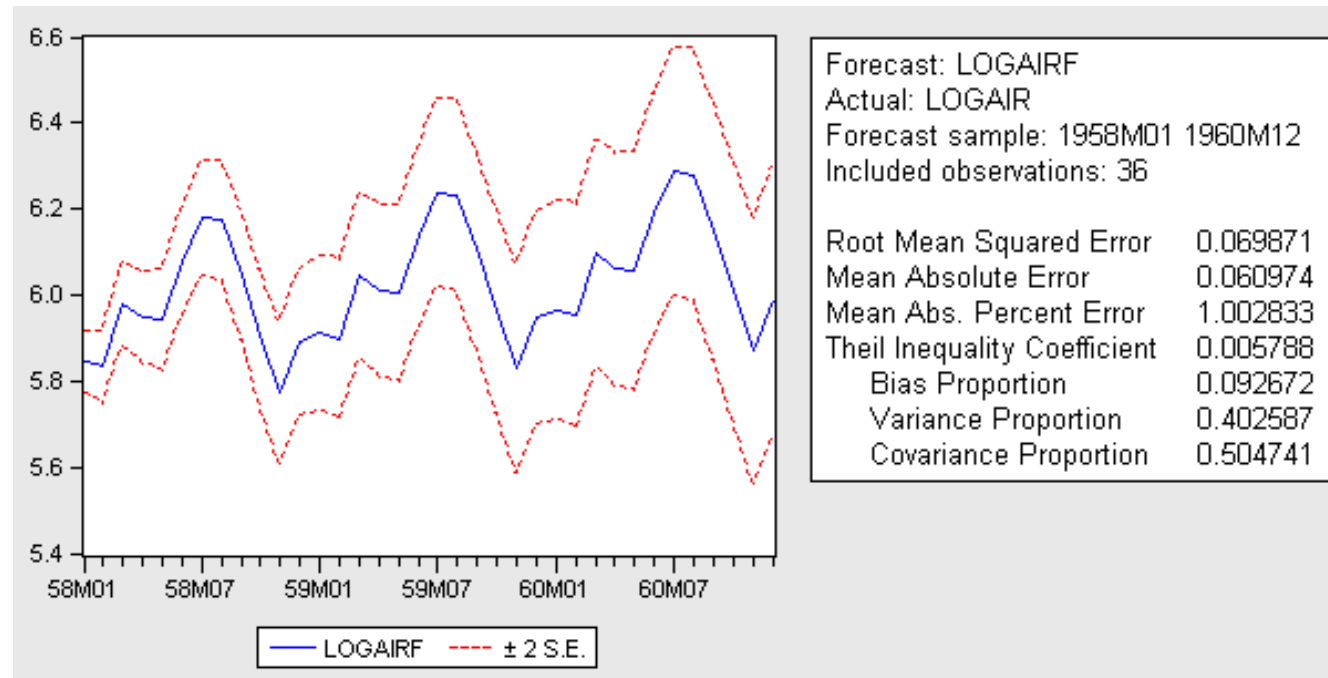
Príklad - odhadnutý model

- Korelogram rezíduí:

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.029	0.029	0.1092	
		2	0.017	0.016	0.1493	
█	█	3	-0.129	-0.130	2.3981	0.121
█	█	4	-0.156	-0.152	5.7495	0.056
		5	0.039	0.052	5.9595	0.114
		6	0.055	0.046	6.3875	0.172
█	█	7	-0.081	-0.129	7.2990	0.199
█	█	8	-0.048	-0.064	7.6305	0.266
		9	0.092	0.138	8.8490	0.264
█	█	10	-0.081	-0.101	9.7827	0.281
		11	0.030	-0.034	9.9159	0.357
		12	0.012	0.047	9.9368	0.446
		13	0.039	0.076	10.159	0.516
		14	0.046	-0.011	10.479	0.574
		15	0.041	0.029	10.738	0.633
█	█	16	-0.163	-0.121	14.775	0.394
		17	0.015	0.040	14.811	0.465
		18	-0.008	-0.007	14.820	0.538
█	█	19	-0.111	-0.143	16.751	0.471
█	█	20	-0.107	-0.164	18.541	0.421
		21	-0.037	0.021	18.758	0.472
		22	-0.025	-0.036	18.859	0.531
		23	0.220	0.127	26.696	0.181
		24	0.048	-0.012	27.069	0.209
		25	0.006	0.020	27.075	0.252

Príklad - predikcie

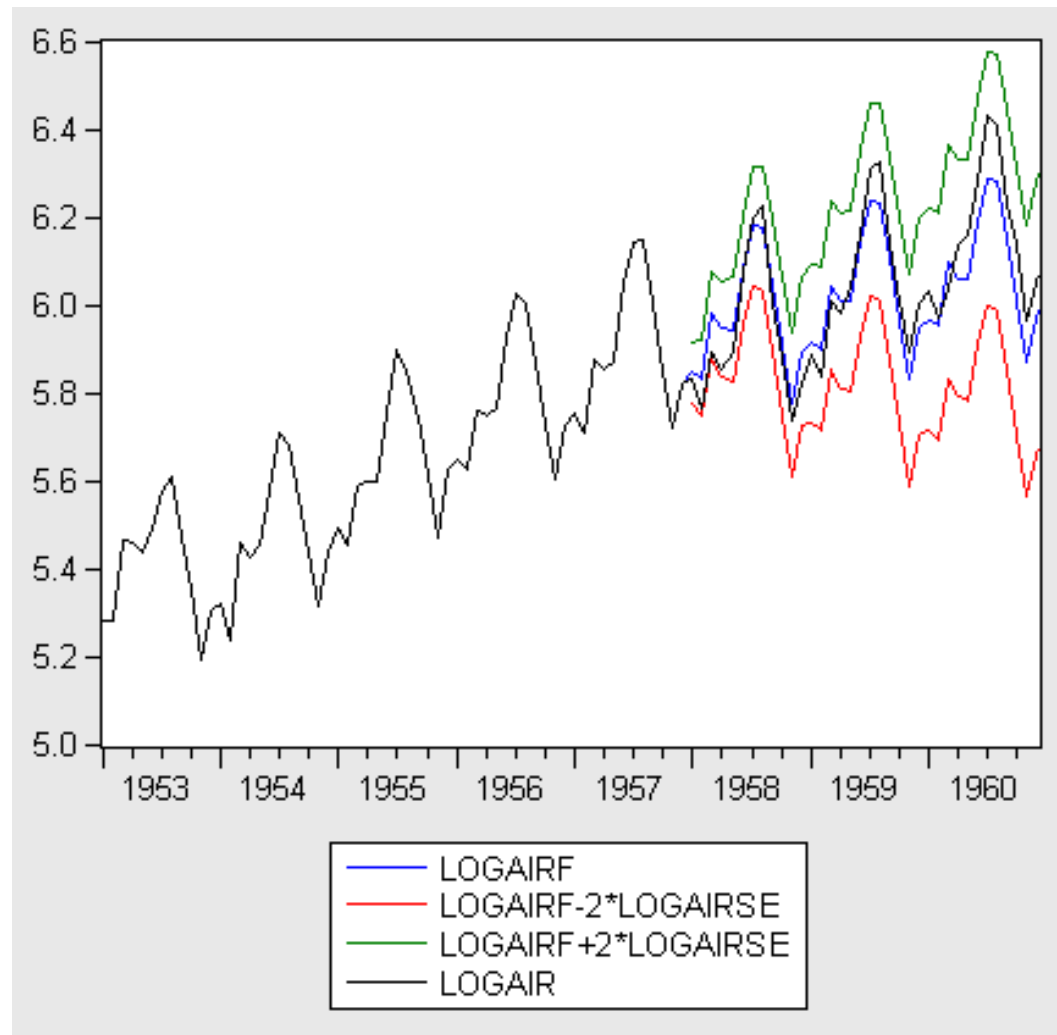
- Vynecháme posledné tri roky a spravíme pre ne predikcie:



- majú rovnaký sezónny charakter ako pôvodné dáta

Príklad - predikcie

- Porovnanie predikcií so skutočnými hodnotami:



SARIMA modely - terminológia

- Pripomeňme si **ARIMA** (p, d, q) :
 - ◇ p - počet AR členov
 - ◇ d - koľkokrát dáta diferencujeme
 - ◇ q - počet MA členov
- **SARIMA** $(p, d, q) \times (P, D, Q)_s$:
 - ◇ P - počet sezónnych AR členov
 - ◇ D - koľkokrát dáta sezónne diferencujeme
 - ◇ Q - počet sezónnych MA členov
 - ◇ s - perióda dát
- "Airline model" je SARIMA $(0, 1, 1) \times (0, 1, 1)_s$, v našom prípade $s = 12$