

Modelovanie sezónnosti

Beáta Stehlíková
Časové rady, FMFI UK

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)

ACF procesu so sezónnymi členmi

Shumway, Robert H., and David S. Stoffer. Time series analysis and its applications: with R examples. Springer, 2011. Správime podrobný výpočet ACF z cvičenia 3.44

Example 3.44 A Mixed Seasonal Model

Consider an $\text{ARMA}(0, 1) \times (1, 0)_{12}$ model

$$x_t = \Phi x_{t-12} + w_t + \theta w_{t-1},$$

where $|\Phi| < 1$ and $|\theta| < 1$. Then, because x_{t-12} , w_t , and w_{t-1} are uncorrelated, and x_t is stationary, $\gamma(0) = \Phi^2 \gamma(0) + \sigma_w^2 + \theta^2 \sigma_w^2$, or

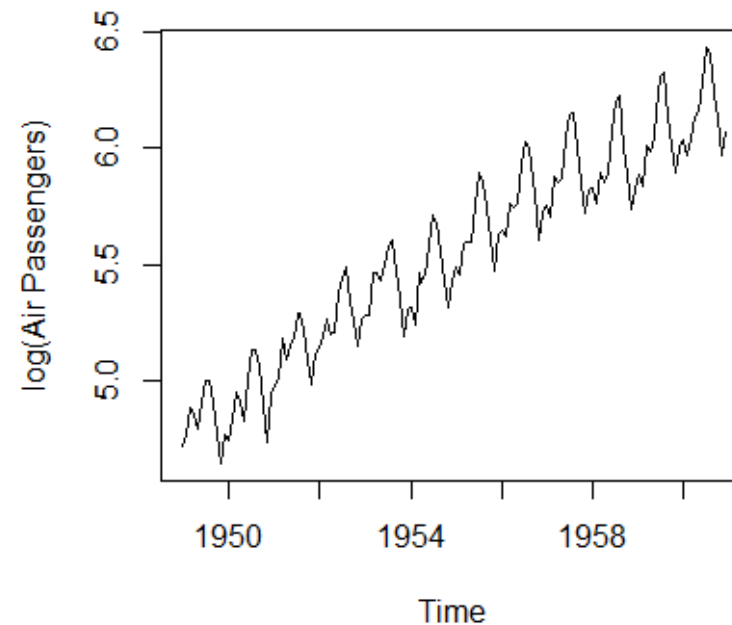
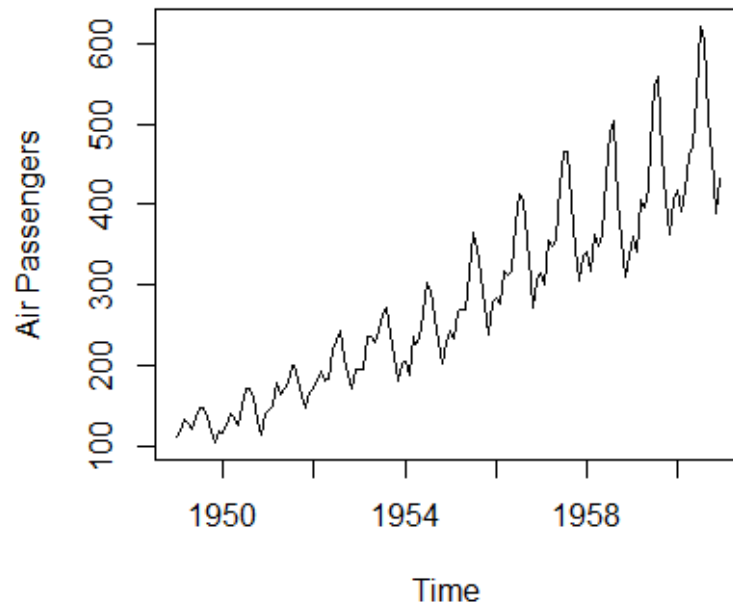
$$\gamma(0) = \frac{1 + \theta^2}{1 - \Phi^2} \sigma_w^2.$$

In addition, multiplying the model by x_{t-h} , $h > 0$, and taking expectations, we have $\gamma(1) = \Phi \gamma(11) + \theta \sigma_w^2$, and $\gamma(h) = \Phi \gamma(h - 12)$, for $h \geq 2$. Thus, the ACF for this model is

$$\begin{aligned} \rho(12h) &= \Phi^h \quad h = 1, 2, \dots \\ \rho(12h - 1) &= \rho(12h + 1) = \frac{\theta}{1 + \theta^2} \Phi^h \quad h = 0, 1, 2, \dots, \\ \rho(h) &= 0, \quad \text{otherwise.} \end{aligned}$$

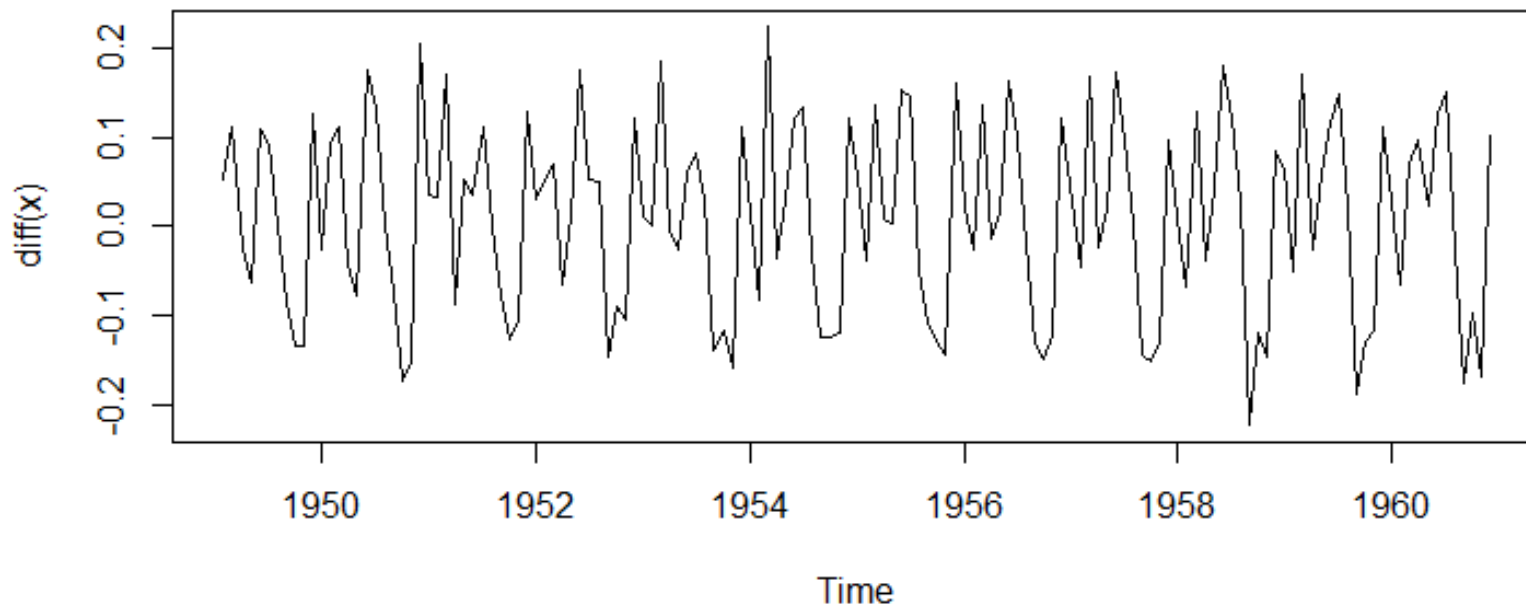
Príklad - dáta

- Počet cestujúcich aerolinkami od Boxa a Jenkinsa - príklad od zakladateľov ARIMA modelovania
- Mesačné dáta, január 1949 - december 1960
- Pracujeme s logaritmami, stabilizujú disperziu
- V R-ku: `data(AirPassengeres); x=log(AirPassengeres)`



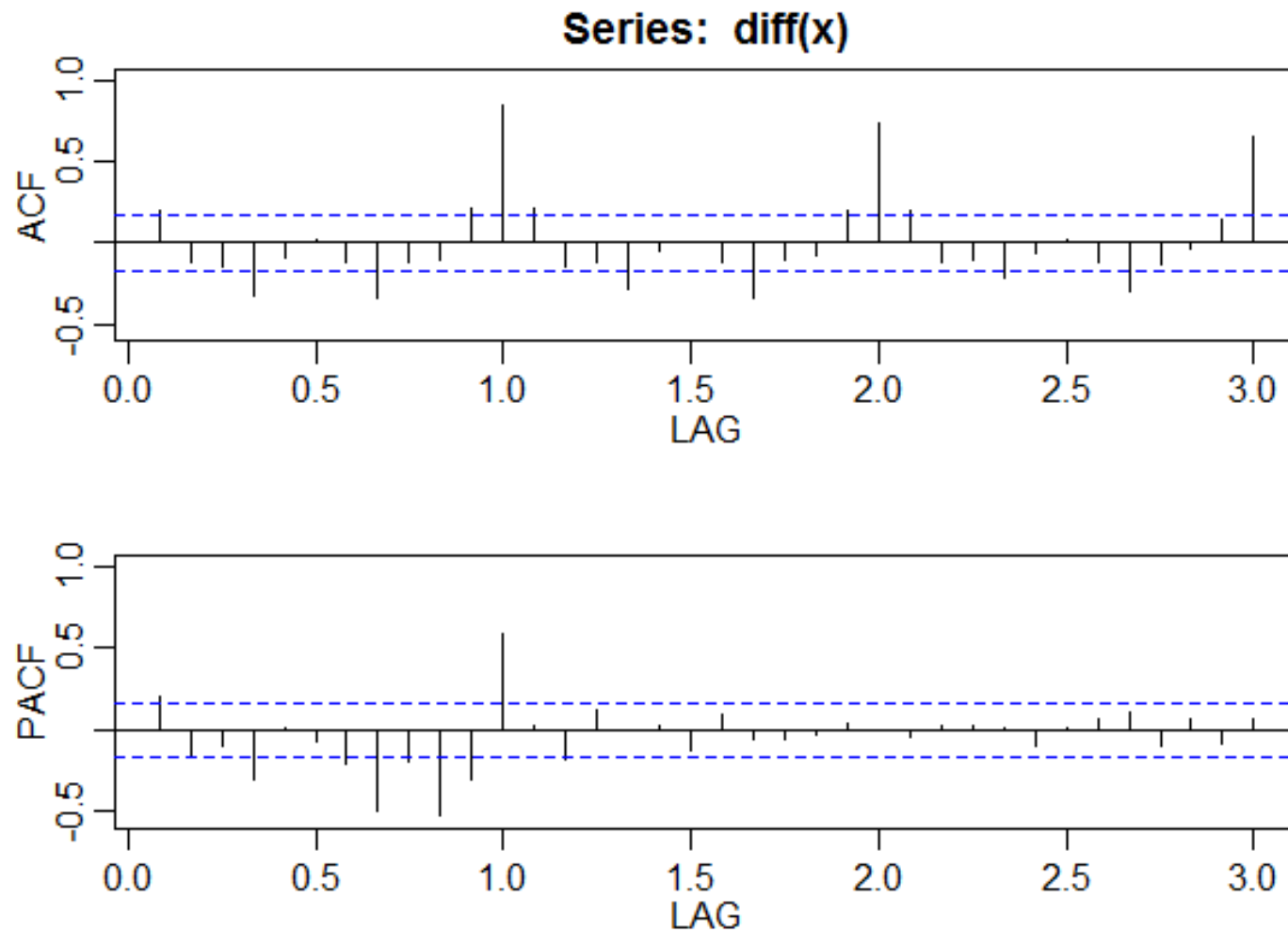
Príklad - diferencovanie

- Spravíme diferencie - v R-ku `diff(x)` - je v nich ročná sezónnosť:



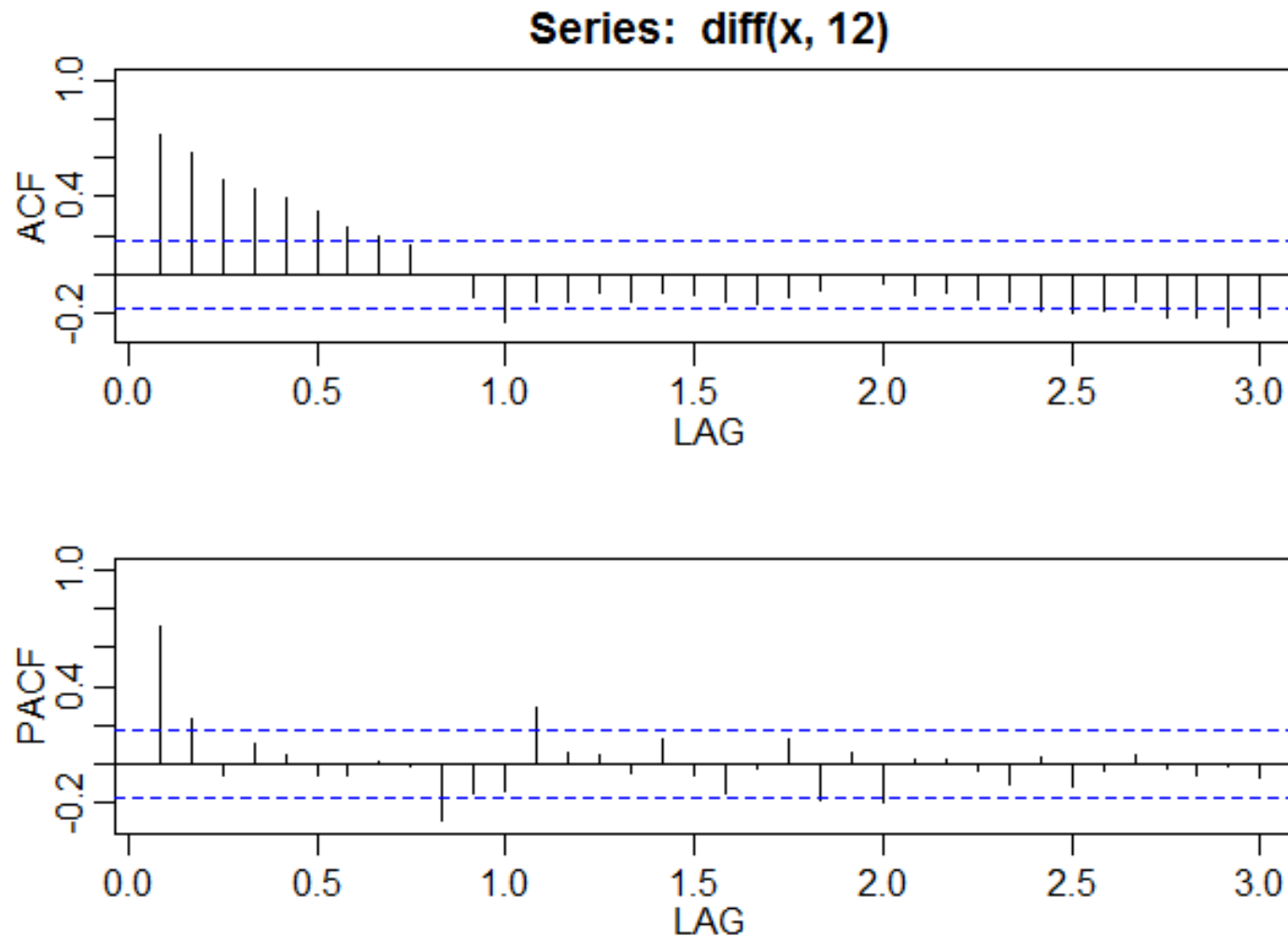
Príklad - diferencovanie

- ACF a PACF pre tieto diferencie - v R-ku `acf2(diff(x))`:



Príklad - diferencovanie

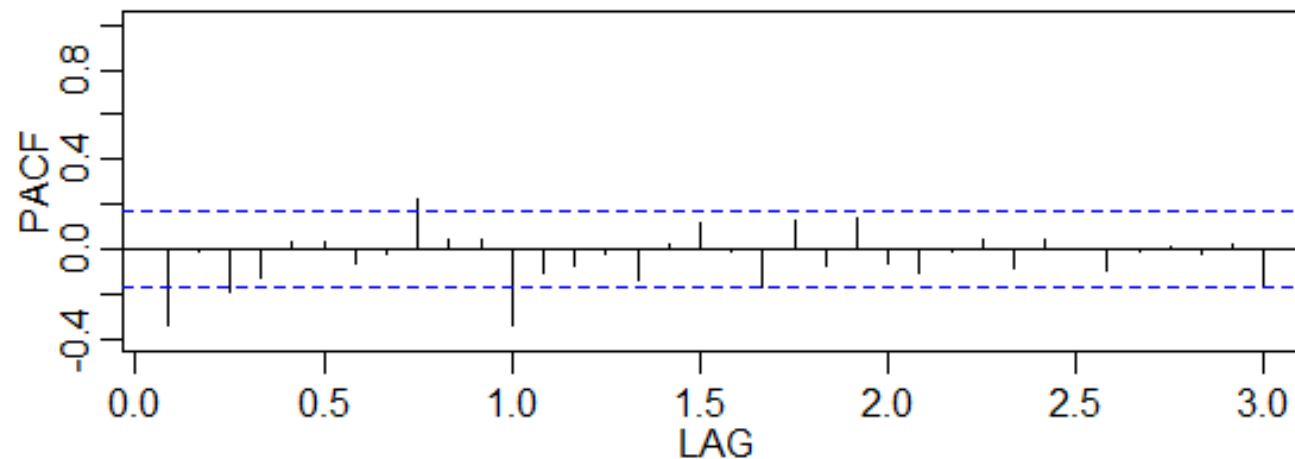
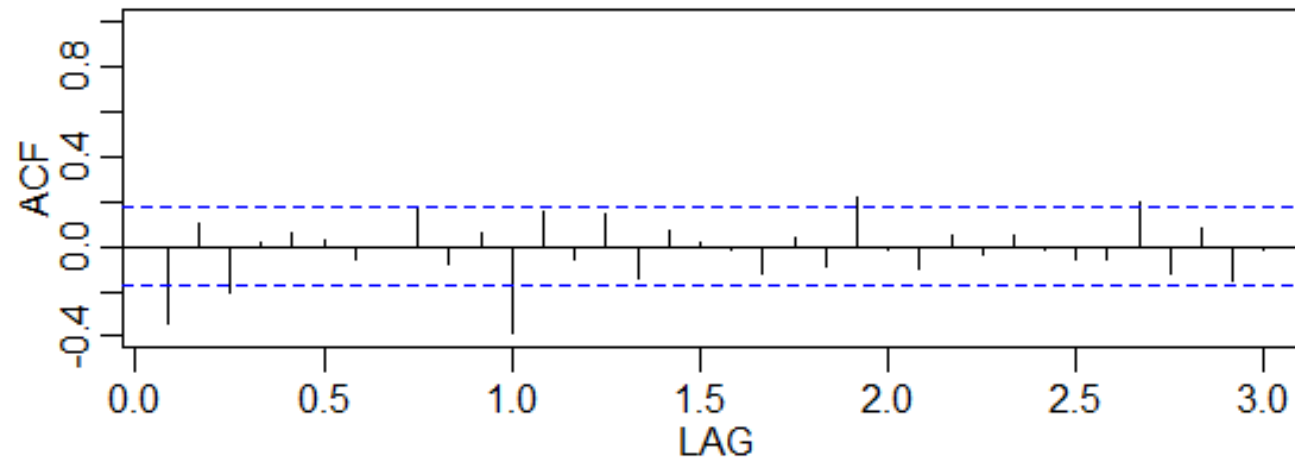
- Môžeme spraviť ročné diferencie $x_t - x_{t-12}$ - v R-ku `diff(x,12)`:



Príklad - diferencovanie

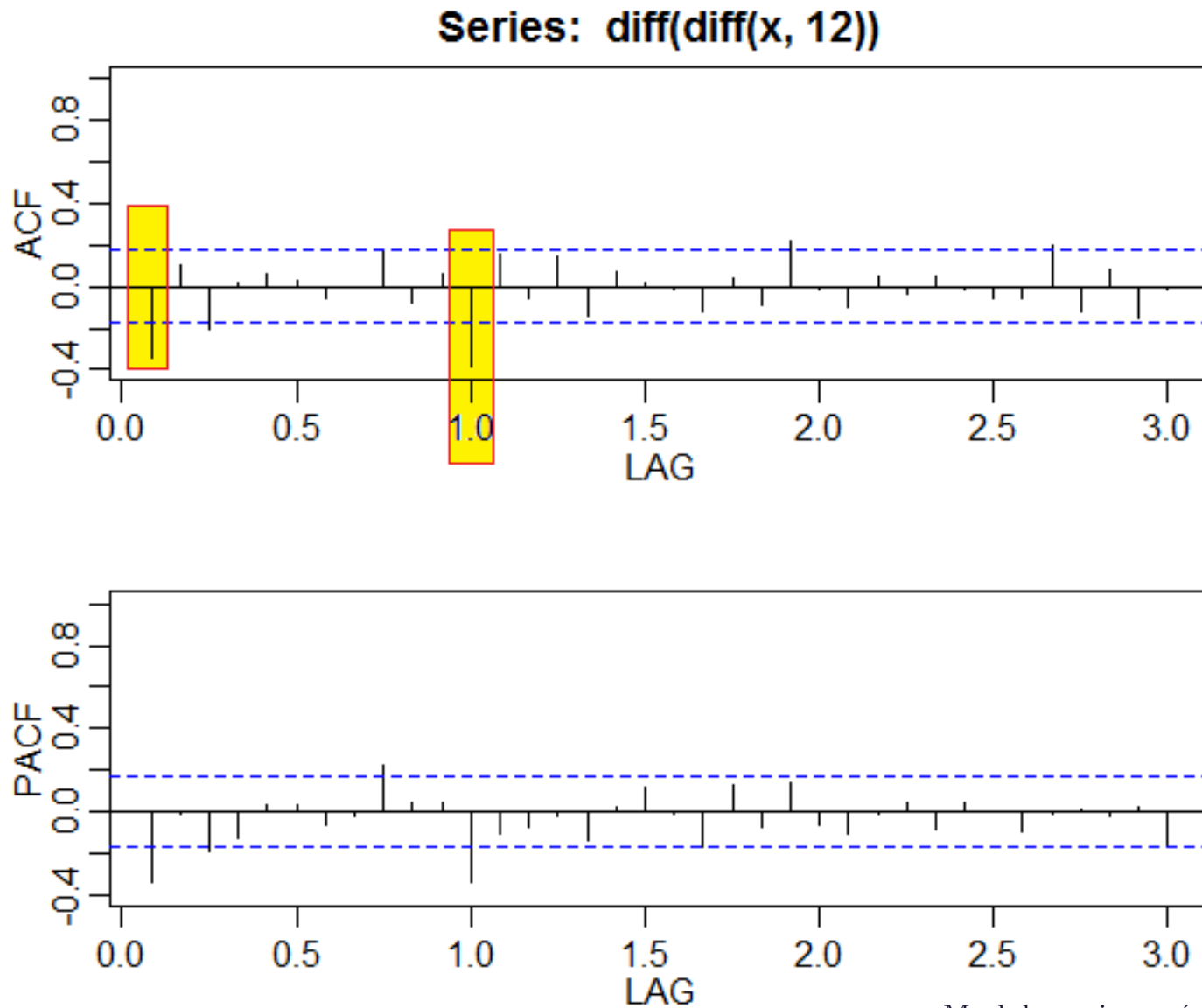
- Aj klasické, aj ročné diferencie - klasické kvôli trendu, sezónne kvôli sezónnosti - v R-ku teda `diff(diff(x,12))`:

Series: `diff(diff(x, 12))`



Príklad - diferencovanie

- Čo s tým:



Príklad - sezónne AR a MA členy

- Podľa ACF by sme mohli skísit' úst' až po **ma(12)**
- Box a Jenkins:
 - ◇ nie vetky **ma(1), ma(2), ..., ma(12)**
 - ◇ ani nie iba **ma(1)** a **ma(12)**
 - ◇ ale vynásobit' polynómy rádu 1 a rádu 12:

$$(1 - \beta L)(1 - \theta L^{12})u_t$$

- dostaneme **13 ma** členov, a potrebujeme na to iba **2 koeficienty**

- Analogicky + dajú sa kombinovat':
 - ◇ sezónne **ma** členy vyššieho rádu: $1 - \theta_1 L^{12} - \theta_2 L^{24}$
 - ◇ sezónny **ar** člen s klasickým: $(1 - \alpha L)(1 - \theta_1 L^{12})x_t$
 - ◇ sezónny **ma** člen s klasickým: $(1 - \alpha L)(1 - \theta_1 L^{12})u_t$

SARIMA modely - terminológia

- Pripomeňme si **ARIMA** (p, d, q) :
 - ◇ p - počet AR členov
 - ◇ d - koľkokrát dáta diferencujeme
 - ◇ q - počet MA členov
- **SARIMA** $(p, d, q) \times (P, D, Q)_s$ má navyše:
 - ◇ P - počet sezónnych AR členov
 - ◇ D - koľkokrát dáta sezónne diferencujeme
 - ◇ Q - počet sezónnych MA členov
 - ◇ s - perióda dát
- Po diferencovaní (klasickom a/alebo sezónnom) musíme mať časový rad, v ktorom nie je trend (sezónnosť tam byť môže) ani jednotkový koreň

Príklad - model v R-ku

- Čo chceme pre naše dáta: **SARIMA** $(0, 1, 1) \times (0, 1, 1)_s$, pričom $s = 12$
- Časový rad `diff(diff(x,12))` nemá trend ani jednotkový koreň
- V R-ku: `sarima(x,0,1,1,0,1,1,12)`
- Dostaneme:

```
> sarima(x,0,1,1,0,1,1,12,details="FALSE")
$fit
Series: xdata
ARIMA(0,1,1)(0,1,1)[12]

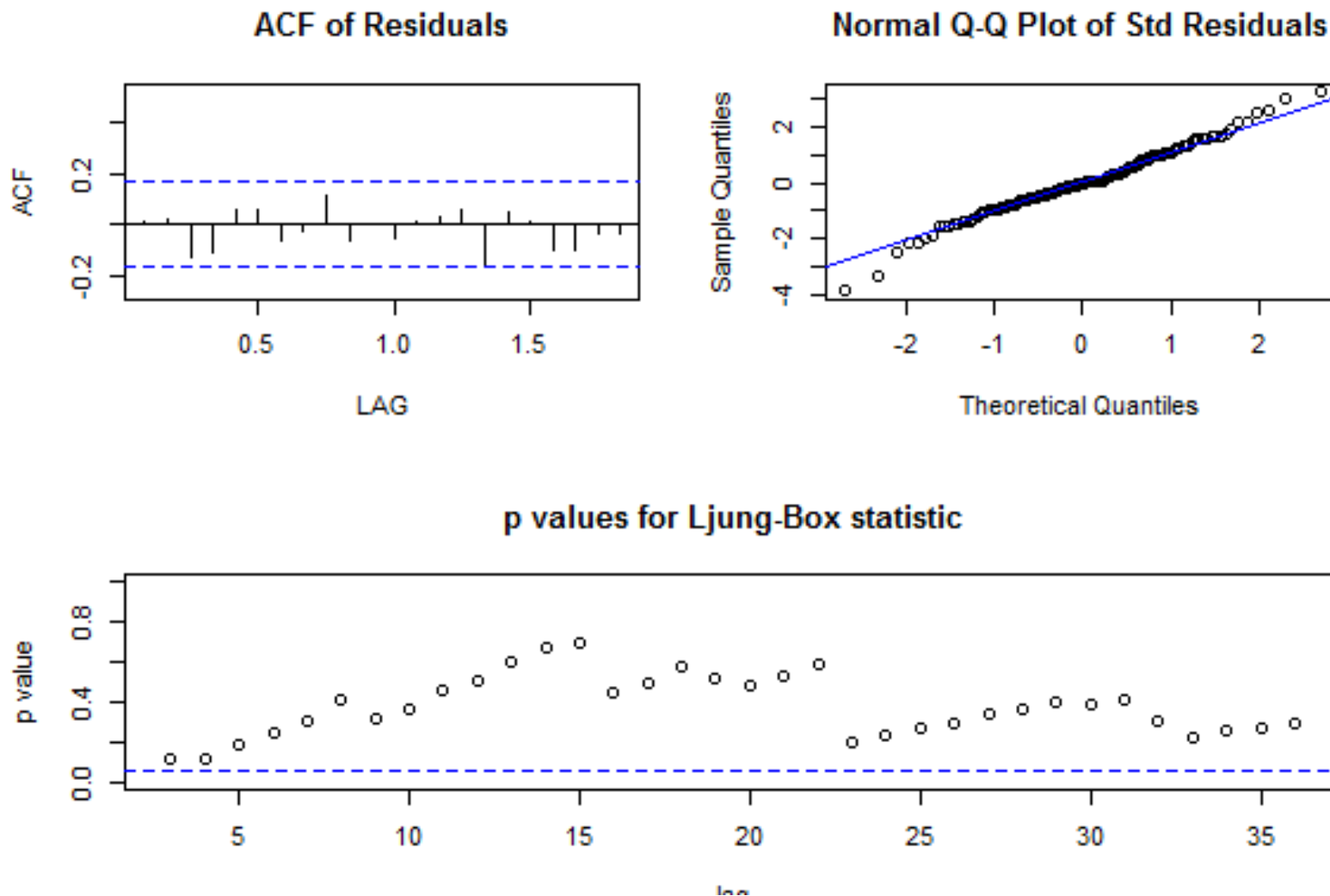
Coefficients:
            ma1      sma1
            -0.4018  -0.5569
s.e.         0.0896   0.0731

sigma^2 estimated as 0.001348:  log likelihood=244.7
AIC=-483.4   AICc=-483.21   BIC=-474.77

$AIC
[1] -5.58133
```

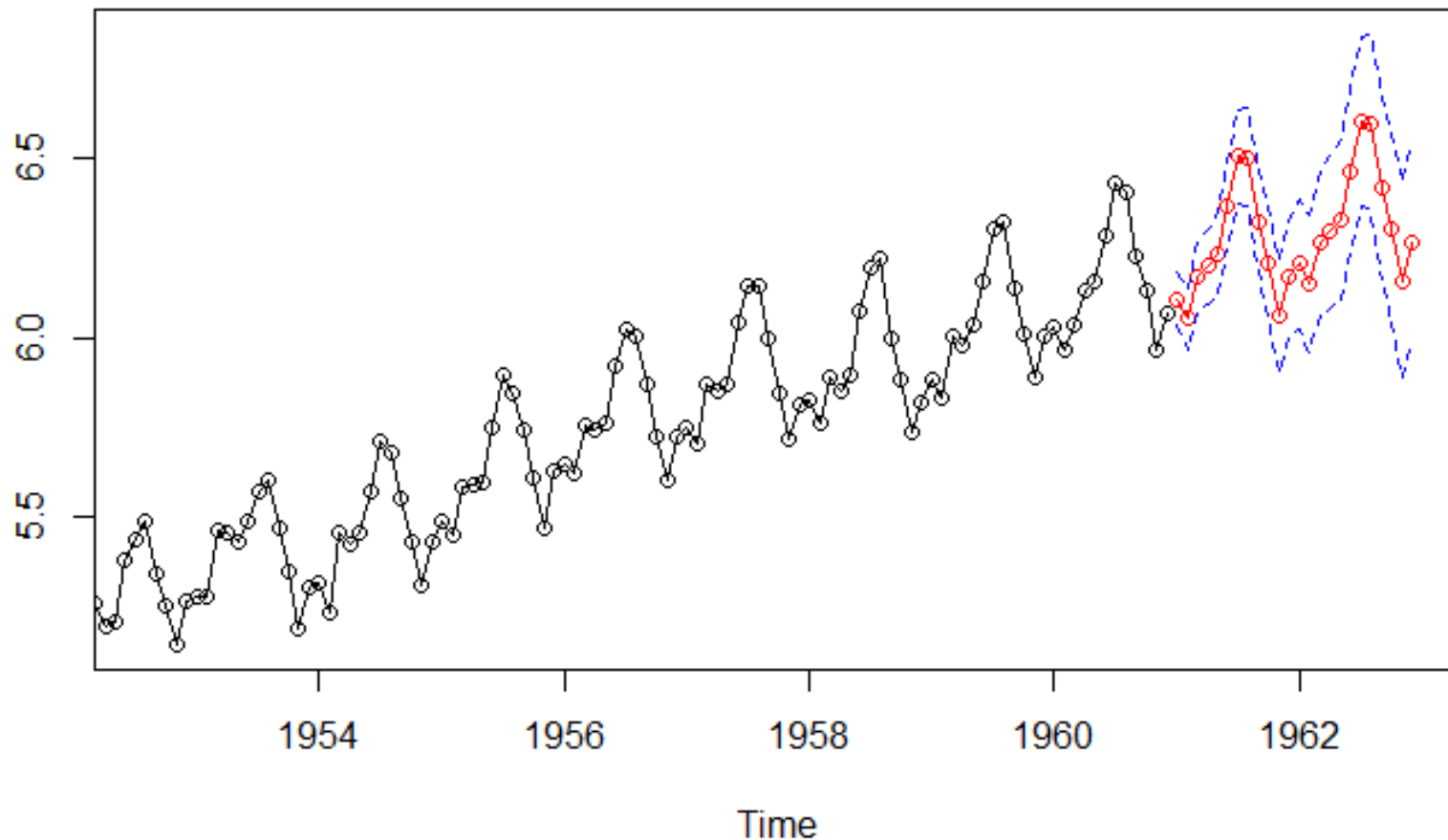
Príklad - model v R-ku

- Rezíduá sú dobré:



Príklad - predikcie v R-ku

- Máme model **SARIMA** $(0, 1, 1) \times (0, 1, 1)_s$
- Spravíme predikciu na nasledujúcich 24 mesiacov:
`sarima.for(x,24,0,1,1,0,1,1,12)`



Cvičenia

Nájdite vhodný SARIMA model (dáta na stránke):

- `spanielsko.txt` - počet turistov v Španielsku; mesačné dáta od januára 1970 do marca 1989
- `suveniry.txt` - tržby v obchode so suvenírmi v lodenici na pláži v Austrálii; mesačné dáta od januára 1987 do decembra 1993