ARMA models
part 3: mixed models (ARMA)

Beáta Stehlíková

ARMA models motivation I.

- We estimate ACF and PACF for the data and they do not like like neither AR nor MA process
- We would like to try to combine AR and MA terms

ARMA models - motivation II.

Consider a stationary and invertible process:

	AR(p)	MA(q)
ACF(τ)	nonzero	0 for $\tau > q$
PACF(τ)	0 for $\tau > p$	nonzero
AR(∞) representation	finite sum	infinite sum
MA(∞) repr. (Wold)	infinite sum	finite sum

- Neither of these models allows a possibility that both ACF and PACF are nonzero
- We would need a proces with infinite AR and MA representations
- This property holds for mixed ARMA models (mixed = both AR and MA terms)

VII.

Model ARMA(1,1)

ARMA(1,1) - definition

- Let u_t be a white noise, define

$$x_t = \delta + \alpha x_{t-1} + u_t - \beta u_{t-1},$$

where $\alpha \neq \beta$. Process x_t is then called an ARMA(1,1) process.

- Using lag operator L:

$$\begin{aligned} (x_t - \alpha x_{t-1}) &= \delta + (u_t - \beta u_{t-1}) \\ (1 - \alpha L) x_t &= \delta + (1 - \beta L) u_t \qquad (1) \end{aligned}$$

ARMA(1,1) - Wold repr. and stationarity

- From (1) we express the process x_t:

  (2) $\qquad x_t = (1-\alpha L)^{-1}\delta + (1-\alpha L)^{-1}(1-\beta L)u_t$
- We know that $(1-\alpha L)^{-1}$ exists, if $|\alpha| < 1$ and in this case we have:

$$(1-\alpha L)^{-1} = 1 + \alpha L + \alpha^2 L^2 + \dots$$

- Substitute into (2):

$$\begin{aligned} x_t &= \delta/(1-\alpha) + (1 + \alpha L + \alpha^2 L^2 + \dots)(1-\beta L)u_t \\ &= \delta/(1-\alpha) + u_t + (\alpha-\beta)u_{t-1} + \alpha(\alpha-\beta)u_{t-2} + \dots \end{aligned}$$

  so in Wold representation

$$\psi_0 = 1, \psi_1 = \alpha(\alpha-\beta), \psi_2 = \alpha^2(\alpha-\beta), \dots, \psi_k = \alpha^k(\alpha-\beta), \dots$$

- Stationarity condition: $|\alpha| < 1$ can we written as: root of $1-\alpha L$ has to be outside of the unit circle

ARMA(1,1) - on the condition $\alpha \neq \beta$

- We have Wold representation:

$$\begin{aligned} x_t &= \delta/(1-\alpha) + (1 + \alpha L + \alpha^2 L^2 + \ldots)(1 - \beta L) u_t \\ &= \delta/(1-\alpha) + u_t + (\alpha - \beta) u_{t-1} + \alpha(\alpha - \beta) u_{t-2} + \ldots \end{aligned}$$

- If $\alpha = \beta$, then

$$x_t = \delta/(1-\alpha) + u_t,$$

so the process is only constant + white noise

ARMA(1,1) - invertibility

- From (1) we express the white noise u_t, in order to obtain the process x_t written using its lagged values + the current value of the white noise:

$$
\begin{aligned}
-\delta + (1-\alpha L)x_t &= (1-\beta L)u_t \\
-(1-\beta L)^{-1}\delta + (1-\beta L)^{-1}(1-\alpha L)x_t &= u_t
\end{aligned}
$$

- We know that $(1-\beta L)^{-1}$ exists, if $|\beta| < 1$
- This invertibility condition can be written as: root of $1-\beta L$ has to be outside of the unit circle

ARMA(1,1) - summary

- We recall the process (1):

$$(1 - \alpha L)x_t = \delta + (1 - \beta L)u_t$$

- Stationarity condition:
    - ◇ root of $1 - \alpha L$ is outside of the unit circle
    - ◇ depends only on the AR part of the process
- Invertibility condition:
    - ◇ root of $1 - \beta L$ is outside of the unit circle
    - ◇ depends only on the MA part of the process

VIII.

Model ARMA(p,q)

ARMA(p,q) - definition

- Let u_t be a white noise, define

$$x_t = \delta + \alpha_1 x_{t-1} + \ldots + \alpha_p x_{t-p} + u_t - \beta_1 u_{t-1} - \ldots - \beta_q u_{t-q},$$

  this process is then called ARMA(p,q) process.

- Using lag operator L:

$$\begin{aligned} (1 - \alpha_1 L - \ldots \alpha_p L^p) x_t &= \delta + (1 - \beta_1 L - \ldots - \beta_q L^q) u_t \\ \alpha(L) x_t &= \delta + \beta(L) u_t \end{aligned} \tag{3}$$

  where we require that polynomials $\alpha(L)$, $\beta(L)$ do not have a common root (more about this later)

ARMA(p,q) - Wold repr., stationarity

- From (3) we express x_t:

$$\begin{aligned} \alpha(L)x_t &= \delta + \beta(L)u_t \\ x_t &= \alpha(L)^{-1}\delta + \alpha(L)^{-1}\beta(L)u_t \end{aligned}$$

- We need $\alpha(L)^{-1}\beta(L)$:

$$\begin{aligned} \alpha(L)^{-1}\beta(L) &= \psi_0 + \psi_1 L + \psi_2 L^2 + \ldots \\ \beta(L) &= \alpha(L)(\psi_0 + \psi_1 L + \psi_2 L^2 + \ldots) \\ (1 - \beta_1 L - \ldots - \beta_q L^q) &= (1 - \alpha_1 L - \ldots \alpha_p L^p) \times \\ &\quad \times(\psi_0 + \psi_1 L + \psi_2 L^2 + \ldots) \end{aligned}$$

Comparing coefficients at L^j

ARMA(p,q) - Wold repr., stationarity

- For coefficients ψ_j of the Wold representation we get:
  - ◇ difference equation

$$\psi_k - \alpha_1\psi_{k-1} - \ldots - \alpha_p\psi_{k-p} = 0$$

  - ◇ initial conditions
- To satisfy the convergence of $\sum \phi_j^2$ - roots of $\lambda^p - \alpha_1\lambda^{p-1} - \ldots \alpha_p = 0$ have to be inside, i.e., roots of $\alpha(L) = 0$ outside of the unit circle

ARMA(p,q) - invertibility

- From (3) we express u_t:

$$\begin{aligned}\alpha(L)x_t &= \delta + \beta(L)u_t \\ \beta(L)u_t &= -\delta + \alpha(L)x_t \\ u_t &= -\beta(L)^{-1}\delta + \beta(L)^{-1}\alpha(L)x_t\end{aligned}$$

- This can be done if $\beta(L)^{-1}$exists, which means that roots of $\beta(L) = 0$ sre outside of the unit circle

ARMA(p,q) - moments

- Expected value: μ:

$$x_t = \delta + \alpha_1 x_{t-1} + \ldots + \alpha_p x_{t-p} + u_t - \beta_1 u_{t-1} - \ldots - \beta_q u_{t-q}$$

$$\mu = \delta + \alpha_1 \mu + \ldots + \alpha_p \mu \Rightarrow \mu = \frac{\delta}{1 - \alpha_1 - \ldots - \alpha_p}$$

- Variance, autocovariances - - WLOG $\delta = 0$:

$$x_t = \alpha_1 x_{t-1} + \ldots + \alpha_p x_{t-p} + u_t - \beta_1 u_{t-1} - \ldots - \beta_q u_{t-q}$$

$$\Big/ \quad \times x_{t-s},\ E[.]$$

$$\begin{aligned}\gamma(s) &= \alpha_1 \gamma(s-1) + \ldots + \alpha_p \gamma(s-p) \\ &\quad + E[u_t x_{t-s}] - \beta_1 E[u_{t-1} x_{t-s}] - \ldots - \beta_q E[u_{t-q} x_{t-s}]\end{aligned}$$

ARMA(p,q) - moments

- For $s > q$ are all expected values

$$E[u_t x_{t-s}], E[u_{t-1} x_{t-s}], \ldots, E[u_{t-p} x_{t-s}]$$

zero $\Rightarrow$ for $s > q \wedge s > p$ (because we need at least p initial values) we have a difference equation for autocovariances:

$$\gamma(s) = \alpha_1 \gamma(s-1) + \ldots + \alpha_p \gamma(s-p) \tag{4}$$

- ACF - dividing (4) by variance $\gamma(0)$ - we get a difference equation for the autocorrelations $\rho(s)$, $s > \max(p, q)$:

$$\rho(s) = \alpha_1 \rho(s-1) + \ldots + \alpha_p \rho(s-p) \tag{5}$$

- the same as for the process without MA terms; they, however, enter the initial condition

Example: ARMA(1,1)

- Expected value μ:

$$\begin{aligned} x_t &= \delta + \alpha x_{t-1} + u_t - \beta u_{t-1} \qquad \Big/ \quad E[.] \\ \mu &= \delta + \alpha\mu + 0 \Rightarrow \mu = \frac{\delta}{1-\alpha} \end{aligned}$$

- Variance, autocovariances - WLOG $\delta = 0$:

$$\begin{aligned} x_t &= \alpha x_{t-1} + u_t - \beta u_{t-1} \qquad \Big/ \quad \times x_{t-s},\ E[.] \\ E[x_t x_{t-s}] &= \alpha E[x_{t-1}x_{t-s}] + E[u_t x_{t-s}] - \beta E[u_{t-1}x_{t-s}] \\ \text{(6)} \quad \gamma(s) &= \alpha\gamma(s-1) + E[u_t x_{t-s}] - \beta E[u_{t-1}x_{t-s}] \end{aligned}$$

Expected value $E[u_t x_{t-s}]$ is nonzero only for $s = 0$, $E[u_{t-1}x_{t-s}]$ is nonzero only for $s = 0$ and $s = 1$

Example: ARMA(1,1)

- Concrete values $E[u_t x_{t-s}]$ and $E[u_{t-1} x_{t-s}]$ are obtained from the Wold representation

$$x_{t-s} = u_{t-s} + (\alpha - \beta) u_{t-s-a} + \alpha(\alpha - \beta) u_{t-s-2} + \dots$$

We get:

$$E[u_t x_{t-s}] = \begin{cases} \sigma^2 & \text{for } \tau = 0 \\ 0 & \text{for } \tau = 1, 2, 3, \dots \end{cases}$$

$$E[u_{t-1} x_{t-s}] = \begin{cases} (\alpha - \beta)\sigma^2 & \text{for } \tau = 0 \\ \sigma^2 & \text{for } \tau = 1 \\ 0 & \text{for } \tau = 2, 3, \dots \end{cases}$$

and substitute into (6).

Example: ARMA(1,1)

- Finally, from (6) we get for $s = 0$, $s = 1$:

$$s = 0 \quad \Rightarrow \quad \gamma(0) = \alpha\gamma(1) + \sigma^2 - \beta(\alpha - \beta)\sigma^2$$
$$s = 1 \quad \Rightarrow \quad \gamma(1) = \alpha\gamma(0) - \beta\sigma^2$$

$\rightarrow$ system of 2 equations with 2 unknowns, the solution is

$$\gamma(0) \quad = \quad \frac{1 + \beta^2 - 2\alpha\beta}{1 - \alpha^2}\sigma^2, \gamma(1) = \frac{(\alpha - \beta)(1 - \alpha\beta)}{1 - \alpha^2}\sigma^2$$

(7)

- For $s = 2, 3, \ldots$ we get a recurrent relation for the next $\gamma(s)$:

$$\gamma(s) = \alpha\gamma(s - 1)$$

Example: ARMA(1,1)

- For $s = 2, 3, \ldots$ we have

$$\begin{aligned} \gamma(s) &= \alpha\gamma(s-1) \quad \Big/ \quad \frac{1}{\gamma(0)} \\ \rho(s) &= \alpha\rho(s-1) \end{aligned}$$

$\rightarrow$ the same difference equation for the ACF as for the process without the MA part

- but swith a different initial condition - from(7) we have

$$\rho(1) = \frac{\gamma(1)}{\gamma(0)} = \frac{(\alpha - \beta)(1 - \alpha\beta)}{(1 + \beta^2 - 2\alpha\beta)}$$

- depends also on the MA part

PACF

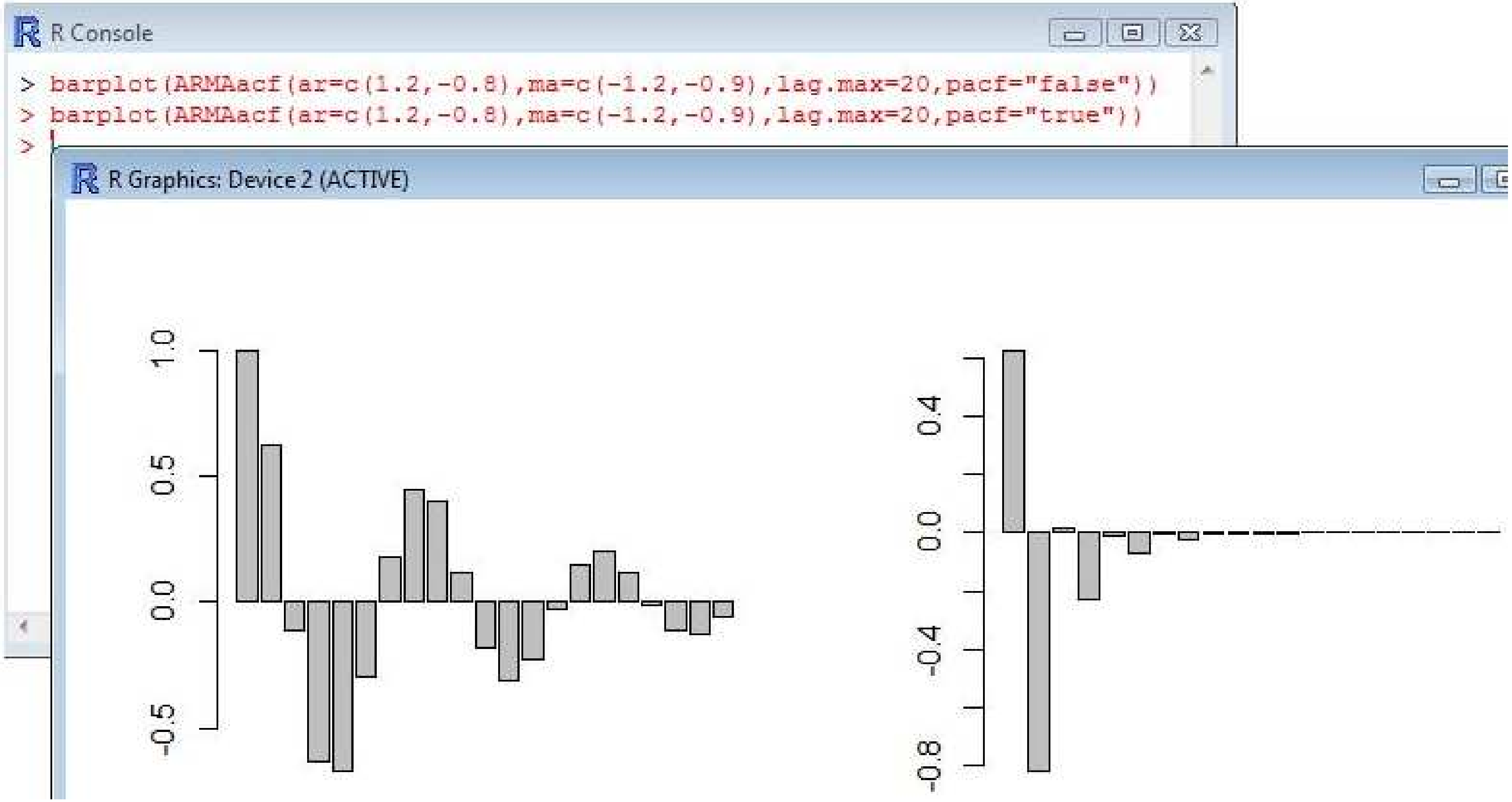
- PACF in the same way as before:

$$\Phi_{kk} = \frac{\det \begin{pmatrix} 1 & \rho(1) & \dots & \rho(1) \\ \rho(1) & 1 & \dots & \rho(2) \\ & \dots & \dots & \\ \rho(k-1) & \rho(k-2) & \dots & \rho(k) \end{pmatrix}}{\det \begin{pmatrix} 1 & \rho(1) & \dots & \rho(k-1) \\ \rho(1) & 1 & \dots & \rho(k-2) \\ & \dots & \dots & \\ \rho(k-1) & \rho(k-2) & \dots & 1 \end{pmatrix}} \tag{8}$$

- Example: for ARMA(1,1) process we substitute $\rho(k) = \alpha^{k-1}\rho(1)$, $\rho(1) = \frac{(\alpha-\beta)(1-\alpha\beta)}{(1+\beta^2-2\alpha\beta)}$

Example - ACF and PACF

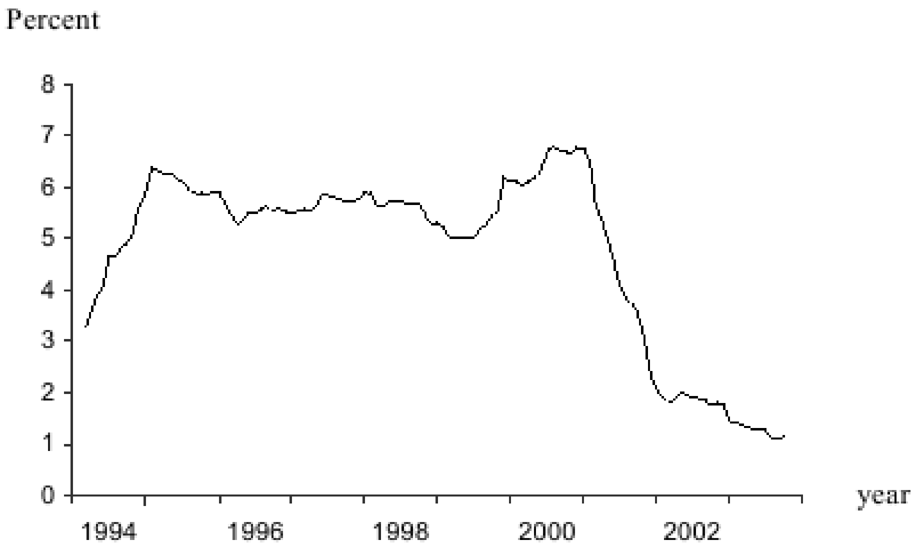
It does NOT holdthat $ACF(k) = 0$ for $k > q$ and $PACF(k) = 0$ for $k > p$ - for example:



Example - real data

[Kirchgässner, Wolters], example 2.15

- USA, March 1994 - August 2003
- USR_t = 3-month interest rate



a) New York three months money market rate, 1994 – 2003

Example - real data

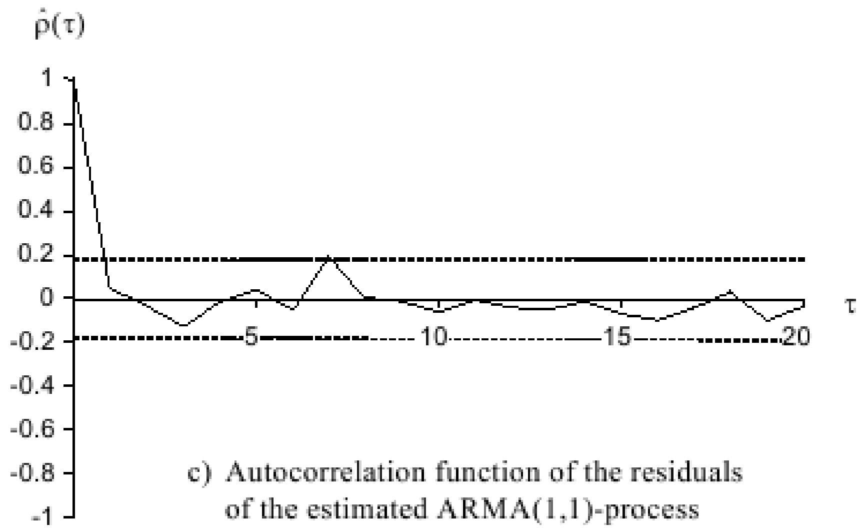
Estimated model for the differences of USR:

The following ARMA(1,1) model has been estimated for this time series:

$$\Delta USR_t = \underset{(-0.73)}{-0.006} + \underset{(10.91)}{0.831}\, \Delta USR_{t-1} + \hat{u}_t - \underset{(-3.57)}{0.457}\, \hat{u}_{t-1},.$$

$$\bar{R}^2 = 0.351, \quad SE = 0.166, \quad Q(10) = 7.897 \ (p = 0.639).$$

The AR(1) as well as the MA(1) terms are different from zero at the 0.1 percent significance level. The autocorrelogram of the estimated residuals, which is also given in *Figure 2.10*, as well as the Box-Ljung Q statistic, which is calculated for this model with 12 autocorrelation coefficients (i.e. with 10 degrees of freedom), do not provide any evidence of a higher order process.



c) Autocorrelation function of the residuals of the estimated ARMA(1,1)-process with confidence intervals

Example - real data

Questions:

- Is the model stationary? Is it invertible?
- *"The autocorrelogram of the estimated residuals... not provide any evidence of a higher order process"* - explain
- *"...the Box-Ljung Q statistic, which is calculated for this model with 12 autocorrelation coefficients (i.e. with 10 degrees of freedom)..."*
  - ◇ what is the null hypothesis?
  - ◇ explain the degrees of freedom
  - ◇ what is the outcome?

ARMA(p,q) - common AR and MA roots

- Recall the definition of the ARMA(p,q) process:

$$\begin{aligned}(1-\alpha_1 L-\ldots-\alpha_p L^p)x_t &= \delta+(1-\beta_1 L-\ldots-\beta_q L^q)u_t\\ \alpha(L)x_t &= \delta+\beta(L)u_t\end{aligned}$$

  where we require that $\alpha(L)$, $\beta(L)$ do not have common roots

- Why there cannot be common roots of $\alpha(L)$, $\beta(L)$?
- Generalization of the property that for ARMA(1,1) we need $\alpha \neq \beta$, otherwise we have trivial process "constant + white noise"

ARMA(p,q) - common AR and MA roots

- Consider "ARMA(2,2)" process

$$(1-\alpha_1 L-\alpha_2 L^2)x_t=\delta+(1-\beta_1 L-\beta_2 L^2)u_t,$$

where
$$\begin{aligned}1-\alpha_1 L-\alpha_2 L^2 &= (1-\gamma L)(1-\gamma_1 L)\\ 1-\beta_1 L-\beta_2 L^2 &= (1-\gamma L)(1-\gamma_2 L)\end{aligned}$$

i.e., AR and MA have a common root γ

- Then:

$$\begin{aligned}(1-\gamma L)(1-\gamma_1 L)x_t &= \delta+(1-\gamma L)(1-\gamma_2 L)u_t\\ (1-\gamma_1 L)x_t &= (1-\gamma L)^{-1}\delta+(1-\gamma_2 L)u_t\end{aligned}$$

so it is ARMA(1,1), and not ARMA(2,2) model

- From a practical point of view - if we have close AR and MA roots, instead of ARMA(p,q) we should try ARMA(p-1,q-1) model

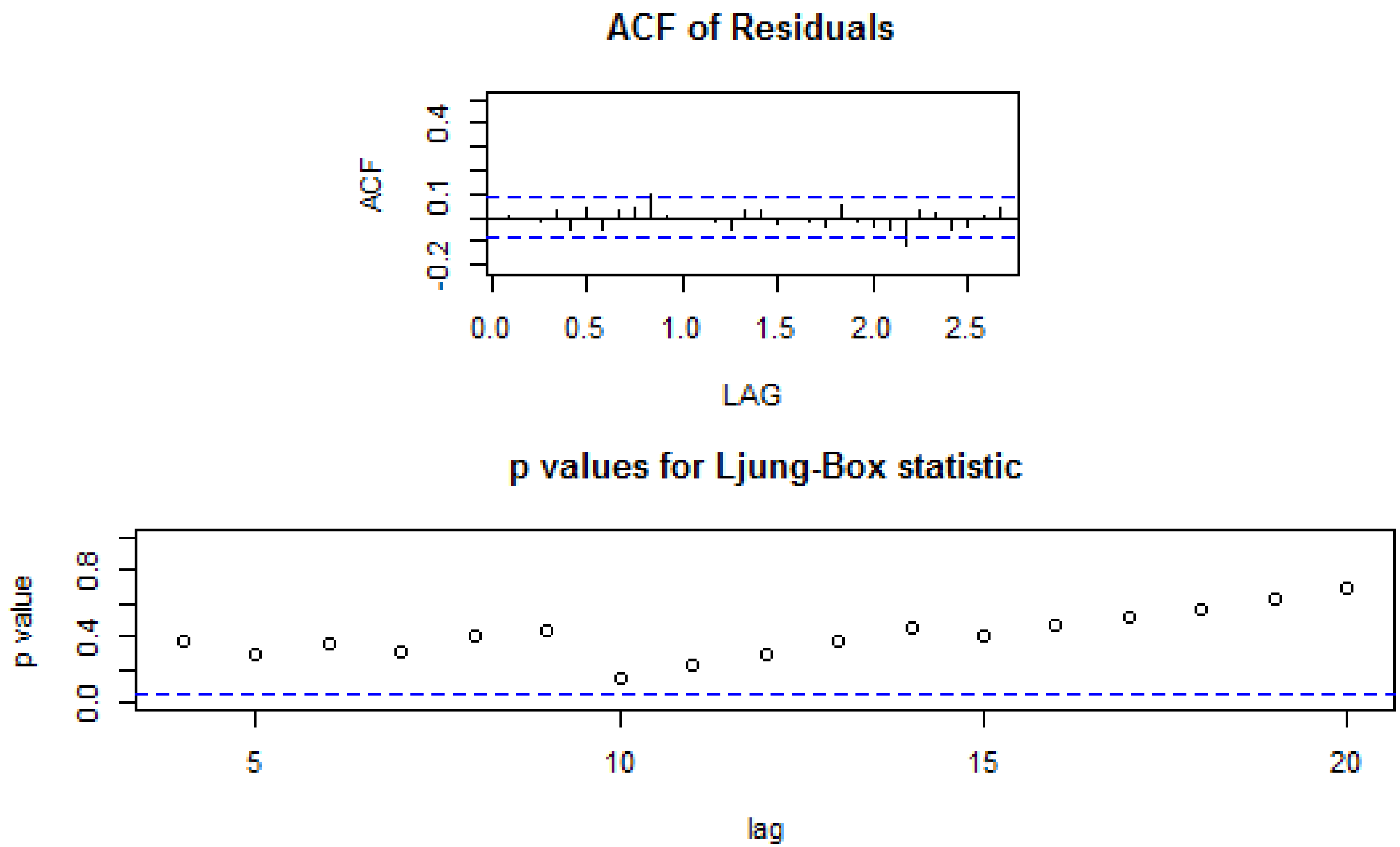
ARMA(p,q) - example

- EXAMPLE: ARMA(1,2) model for the differenced of log prices of cocoa (fromt the previous chapter on MA models):

```
> p=read.table("pcocoa.txt")
> p=ts(p,frequency=12,start=c(1960,1))
> sarima(log(p),1,1,2,details=FALSE)
$fit

Call:
stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,
    Q), period = S), xreg = constant, optim.control = list(trace = trc, REPORT =$
    reltol = tol))

Coefficients:
         ar1      ma1      ma2  constant
      0.8708  -0.5174  -0.3030    0.0025
s.e.  0.3563   0.3622   0.1401    0.0038

sigma^2 estimated as 0.003897:  log likelihood = 693.62,  aic = -1377.24
```

ARMA(p,q) - example

- Residuals:



ARMA(p,q) - example

- EXERCISE:
  Compute the roots of AR and MA parts
- We get: AR root is close to one of the MA roots
- So we should tryARMA(0,1) = MA(1) model instead of ARMA(1,2), and it was indeed a good model for the data