# 4.Estimation Validation and Forecasting

MESIO 2017-2018

M.Pilar Muñoz - J.A.Sanchez-Espigares

Dept. Statistics and Operations Research
UPC- BarcelonaTECH

# Outline

- Calculation of $\pi$-weights and $\psi$-weights
- Estimation of parameters
- Forecasting with ARIMA models

# Invertibility (1/2)

**AR and MA polynomials** are defined as:

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p, \phi_p \neq 0$$

and

$$\theta(z) = 1 + \theta_1 z - \cdots + \theta_q z^q, \theta_q \neq 0$$

Respectively, where z is a complex number

An **ARMA(p,q)** model is said to be **causal** if the time series $\{x_t; t = 0, 1, \cdots\}$ can be written as a one-sided linear process:

$$x_t = \sum_{j=0}^{\infty} \psi_j z_{t-j} = \psi(B) z_t$$

Where $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$ and $\sum_{j=0}^{\infty} |\psi_j| < \infty$, with $\psi_0 = 1$

# Invertibility (2/2)

**ARMA(p,q)** model is said to be **invertible** if the time series $\{x_t; t = 0, 1, \cdots\}$ can be written as

$$\pi(B)x_t = \sum_{j=0}^{\infty} \pi_j x_{t-j} = z_t$$

Where,
$\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j$ and $\sum_{j=0}^{\infty} |\pi_j| < \infty$, assuming $\pi_0 = 1$

**AR(2) with Complex Roots example**

Model: $x_t = 1.5x_{t-1} - 0.75x_{t-2} + z_t$ with $\sigma_z^2 = 1$

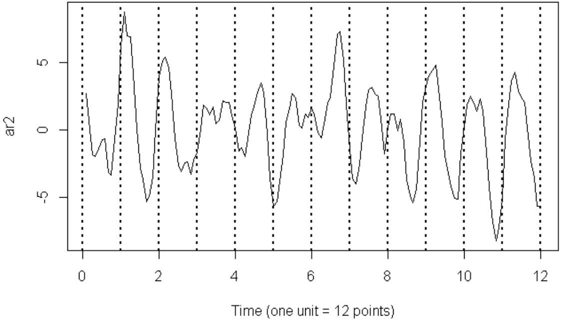
The **autoregressive polynomial** for this model is:

$$\phi(z) = 1 - 1.5z + 0.75z^2$$

The **roots** of $\phi(z)$ are $1 \neq i/\sqrt{3}$ and $\theta = tan^{-1}(\frac{1}{\sqrt{3}}) = 2\pi/12$ radians per unit time

# AR(2) model (2/4)

**Simulated AR(2) model**, n=144 with $\phi_1 = 1.5$ and $\phi_2 = -0.75$.




set.seed(90210)
ar2 = arima.sim(list(order=c(2,0,0), ar=c(1.5,-.75)), n = 144)
plot(1:144/12, ar2, type="l", xlab="Time (one unit = 12 points)")
abline(v=0:12, lty="dotted", lwd=2)

# AR(2) model (3/4)

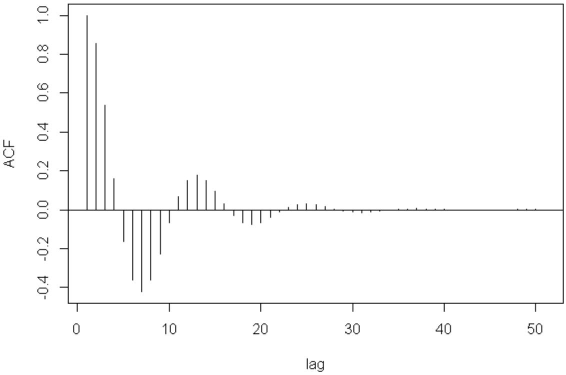
To **calculate the roots of the polynomial and solve** for arg in R:

```
# coefficients of the polynomial
z = c(1,-1.5,0.75)
# print one root: 1+0.57735i = 1 + i/sqrt(3)
(a = polyroot(z)[1])
# arg in cycles/pt
arg = Arg(a)/(2*pi)
# = 12, the pseudo period
1/arg
```

# AR(2) model (4/4)

To **calculate** and display the **ACF** for this model:

ACF = ARMAacf(ar=c(1.5,-.75), ma=0, 50)
plot(ACF, type="h", xlab="lag")
abline(h=0)



**ARMA(p,q)** model: $\psi(B)x_t = \theta(B)w_t$

Any ARMA model can be converted to an **infinite order MA model**:

$$\psi\text{−weights}= \frac{\theta(B)w_t}{\psi(B)}$$

To solve for the $\psi$-**weights** in general, we must match the coefficients in $\phi(z)\psi(z) = \theta(z)$

$$(1-\phi_1 z-\phi_2 z^2-\cdots)(\psi_0+\psi_1 z+\psi_2 z^2+\cdots) = (1+\theta_1 z+\theta_2 z^2+\cdots)$$

The **first few values** are:

$$\psi_0 = 1$$
$$\psi_1 - \phi_1\psi_0 = \theta_1$$
$$\psi_2 - \phi_1\psi_1 - \phi_2\psi_0 = \theta_2$$
$$\psi_3 - \phi_1\psi_2 - \phi_2\psi_1 - \phi_3\psi_0 = \theta_3$$

The $\psi$-weights **satisfy the homogeneous difference equation** given by

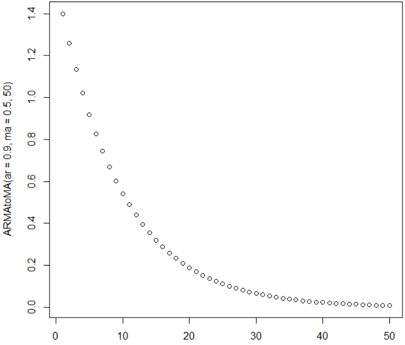
$$\psi_j - \sum_{k=1}^{p} \phi_k\psi_{j-k} = 0,\ j \geq max(p, q+1)$$

# $\psi$-weights for an ARMA model (3/3)

To solve for the $\psi$-**weights in general**, we must match the coefficients in $\Phi(z)\Psi(z) = \theta(z)$

First 50 $\psi$-weights **with R**:

```
ARMAtoMA(ar=.9,ma=.5, 50)
plot(ARMAtoMA(ar=.9,ma=.5, 50)
```



Let $X = \{x_t\}_{t=0}^{N}$ a time series with conditional gaussian distribution

Likelihood of the model:

$$L(\phi, \theta, \sigma_z^2; X) = f_1(x_1) \sum_{i=2}^{N} f(x_i | x_{i-1}, ..., x_1)$$

For example, for an AR(1) model:

$$x_i = \phi x_{i-1} + z_t \quad z_t \sim N(0, \sigma_z^2)$$

$$E(x_i | x_{i-1}, ..., x_1) = \phi x_{i-1}$$

$$V(x_i | x_{i-1}, ..., x_1) = \sigma_z^2$$

$$L(\phi, \sigma_z^2; X) = f_1(x_1) \sum_{i=2}^{N} f(\frac{x_i - \phi x_{i-1}}{2\sigma_z^2})$$

**Maximum likelihood** estimation

$$(\hat{\phi}, \hat{\theta}, \hat{\sigma_z^2}) = argmaxL(\phi, \theta, \sigma_z^2; X)$$

Maximum likelihood estimators are **consistent** and **asymptotically eficient and gaussian**

$$\hat{\Theta}_{ML} \approx N(\Theta, I_{\Theta}^{-1})$$

where $I_{\Theta} = E(-\frac{\partial^2 logL(\Theta; X)}{\partial \Theta^2})$

# Forecasting ARMA(1,1) model (1/2)

Given data $x_1, \cdots, x_n$ for forecasting purpose, write the model as

$$x_{n+1} = \phi x_n + z_{n+1} + \theta z_n \tag{1}$$

- **One-step-ahead truncated forecast**

$$\tilde{x}^n_{n+1} = \phi x_n + 0 + \theta \tilde{z}^n_{n+m-1} \tag{2}$$

- **Errors forward in time**

$$\tilde{w}^n_t = x_t - \phi x_{t-1} + \theta \tilde{z}^n_{t-1},\ t = 1, \cdots, n$$

# Forecasting ARMA(1,1) model (2/2)

- The aproximate **forecast variance**

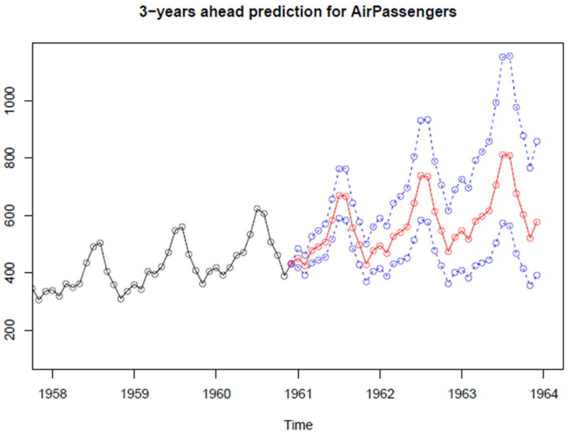
$$P_{n+m}^n = \sigma_z^2[1+(\phi+\theta)^2\sum_{j=1}^{m-1}\phi^{2(j-1)} = \sigma_z^2[1+\frac{(\phi+\theta)^2(1-\phi^{2(m-1)})}{(1-\phi^2)}]] \tag{3}$$

- $(1-\alpha)$ **prediction interval**

$$x_{n+m}^n \pm c_{\alpha/2}\sqrt{P_{n+m}^n} \tag{4}$$

Where $c_{\alpha/2}$ is chosen to get the desired degree of confidence. *i.e. if the process is Gaussian, choosing* $c_{\alpha/2}$ *will yield an aproximate* 95% *prediction interval for* $x_{n+m}$

**ARIMA(0,1,1)(0,1,1)**$_s$; period $s = 12$; $n.ahead = 36$



**Remember:**
A times series $\{x_t; t = 0, \pm 1, \pm 2, \cdots\}$ is an **ARMA(p,q)** model if it is **stationary** and

$$x_t = \phi_1 x_{t-1} + \cdots + \phi_2 x_{t-p} + z_t + \theta_1 z_{t-1} + \cdots + \theta_q z_{t-q} \quad (5)$$

with $\phi_p \neq 0$; $\theta_q \neq 0$ and $\sigma_z^2 > 0$

**ARMA(p,q)** model can be written in concise form as

$$\Phi(B) x_t = \theta(B) z_t \quad (6)$$

**Objective:**
To find the best forecast for $x_{t+h}$ at time t using the **h-step-ahead minimum mean square error predictor**, denoted by $\tilde{X}_{t+h|t}$

**Questions to help the construction process:**

- Which is the **forecasting function**?
  **i.e.**, Which are the **values for** $\tilde{X}_{t+h|t}$ as a function of **h**?

- Which is the **memory of the model**? How are the forecast expressed as a (linear) **function of the previous** (known) observations?

- Which is the **forecasting confidence interval**?

**Three equivalent ARMA(p,q) equations:**

- **Previous observations** and **random noises**

  $$x_{t+h} = \phi_1 x_{t+h-1} + \cdots + \phi_p x_{t+h-p} + z_{t+h} + \theta_1 z_{t+h-1} + \cdots + \theta_q z_{t+h-q}$$

- **AR** process of **infinite order**

  $$x_{t+h} = -\pi_1 x_{t+h-1} - \pi_2 x_{t+h-2} - \cdots + z_{t+h}$$

- **MA** process of **infinite order**

  $$x_{t+h} = z_{t+h} + \Psi_1 z_{t+h-1} + \Psi_2 z_{t+h-2} + \cdots$$

# Forecasting and Confidence Interval (5/5)

**1-step-ahead-forecast**

- **Previous observations** and **random noises**

  $x_{t+h} = \phi_1 x_{t+h-1} + \cdots + \phi_p x t + h - p + z_{t+h} + \theta_1 z_{t+h-1} + \cdots + \theta_q z_{t+h-q}$

- **AR** process of **infinite order**

  $x_{t+h} = -\pi_1 x_{t+h-1} - \pi_2 x_{t+h-2} - \cdots + z_{t+h}$

- **MA** process of **infinite order**

  $x_{t+h} = z_{t+h} + \Psi_1 z_{t+h-1} + \Psi_2 z_{t+h-2} + \cdots$

**Given the model**

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} \qquad (7)$$

Suppose that **values of past observations are known**, then
$Z_t = X_t - \tilde{X}_{t|t-1}$; $Z_{t-1} = X_{t-1} - \tilde{X}_{t-1|t-2}$

Substituting

$$\begin{aligned}
\tilde{X}_{t+1|t} &= \phi_1 X_t + \phi_2 X_{t-1} + E(Z_{t+1}) + \theta_1 Z_t + \theta_2 Z_{t-1} \\
\tilde{X}_{t+2|t} &= \phi_1 \tilde{X}_{t+1|t} + \phi_2 X_t + E(Z_{t+2}) + \theta_1 E(Z_{t+1}) + \theta_2 Z_t \\
\tilde{X}_{t+3|t} &= \phi_1 \tilde{X}_{t+2|t} + \phi_2 \tilde{X}_{t+1|t} + E(Z_{t+3}) + \theta_1 E(Z_{t+2}) + \theta_2 Z_{t+1} \\
&\dots \\
\tilde{X}_{t+h-3|t} &= \phi_1 \tilde{X}_{t+h-1|t} + \phi_2 \tilde{X}_{t+h-2|t} \ , \ h > 2
\end{aligned}$$

**In fact,**

Given a general **ARMA(p,q)** model, for $h > q$:

**H-step-ahead predictor** is determined by the difference equations for the autocorrelations (which only depend on the autocorrelation characteristic polynomial)

# Non Stationary ARIMA process Forecast. ARIMA(0,1,1) (1/3)

**Model:**

$$X_t = X_{t-1} + Z_t + \theta Z_{t-1} \tag{8}$$

**also defined by**

$$(1 - B)X_t = (1 - \theta B)Z_t \tag{9}$$

$X_t$ **can be expressed as**

$$\pi(B) = \frac{1 - B}{1 + \theta B} = 1-(\theta+1)B+\theta(\theta+1)B^2-\theta^2(\theta+1)B^3+\cdots \tag{10}$$

# Non Stationary ARIMA process Forecast. ARIMA(0,1,1) (2/3)

**Model:**

$$\tilde{X}_{t+1|t} = (\theta+1)X_t - \theta(\theta+1)X_{t-1} + \theta^2(\theta+1)X_{t-2} - \cdots \quad (11)$$

**moving** $\theta$

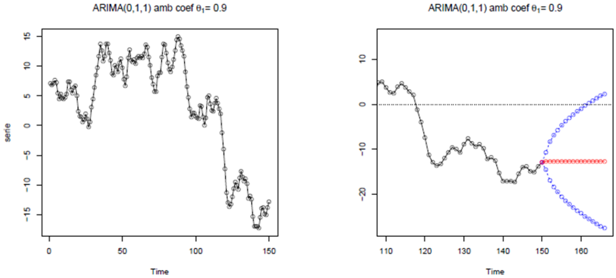
$$\tilde{X}_{t+1|t} = (\theta+1)X_t - \theta\tilde{X}_{t|t-1} \quad (12)$$

**Using the parameter** $\lambda = 1+\theta$. Clearly $\theta = \lambda - 1 = -(1-\lambda)$

$$\tilde{X}_{t+1|t} = \lambda X_t - (1-\lambda)\lambda X_{t-1} + (1-\lambda)^2\lambda X_{t-2} - \cdots = \lambda X_t + (1-\lambda)\tilde{X}_{t|t-1} \quad (13)$$

# Non Stationary ARIMA process Forecast. ARIMA(0,1,1) (3/3)

Hence, the **forecast is the a linear combination**of the last observation and the forecast obtained in the previous step.



# Variance of the forecasting error

**Variance of the 1-step-ahead forecasting error**

$$Var(e_t(1)) = E(X_{t+1} - \tilde{X}_{t+1|t}) = E(Z_{t+1}) = \sigma_Z^2 \qquad (14)$$

**H step forecast for h 2 is**

$$Var(e_t(h)) = \sigma_Z^2(1 + \psi_1^2 + \psi_2^2 + \cdots \psi_{h-1}^2) \qquad (15)$$

# Seasonal Processes Forecasting. ARIMA(0,1,1)(0,1,1)$_{12}$ (1/4)

**Model:**

$$(1-B)(1-B^{12})X_t = (1+\theta B)(1+\theta_{12}B^{12})Z_t \tag{16}$$

**also defined by**

$$(X_t - X_{t-12}) - (X_{t-1} - X_{t-13}) = Z_t \theta_1 Z_{t-1} + \theta_1 2 Z_{t-12} + \theta_1 \theta_{12} Z_{t-13} \tag{17}$$

**Hence,**

$$X_{t+1} = X_{t-11} + (X_t - X_{t-12} + Z_{t+1} + \theta_1 Z_{t-11} + \theta_1 \theta_{12} Z_{t-12}) \tag{18}$$

# Seasonal Processes Forecasting. ARIMA(0,1,1)(0,1,1)$_{12}$ (2/4)

**The forecasting function:**

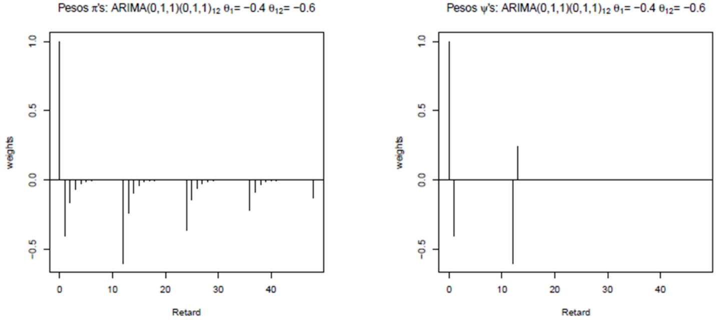
$$\tilde{X}_{t+1|t} = X_{t-11} + (X - X_{t-12}) + \theta_1 Z_t + \theta_1 2 Z_{t-11} + \theta_1 \theta_{12} Z_{t-12}$$
$$\tilde{X}_{t+2|t} = X_{t-10} + (\tilde{X}_{t+1|t} - X_{t-11}) + \theta_{12} Z_{t-10} + \theta_1 \theta_{12} Z_{t-11}$$
$$\dots$$
$$\tilde{X}_{t+h|t} = X_{t+h-12} + (\tilde{X}_{t+h-1|t} - X_{t+h-13}) \ , \ h > 13$$

# Seasonal Processes Forecasting. ARIMA(0,1,1)(0,1,1)$_{12}$ (3/4)

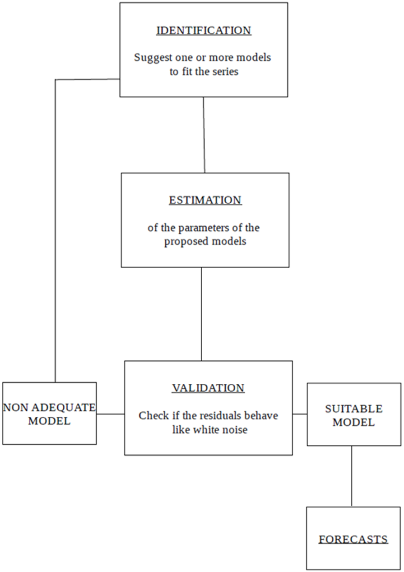
$$\pi(B) = \pi_0 + \pi_1 B + \pi_2 B^2 + \cdots = \frac{1-B}{1+\theta_1 B}\frac{1-B^{12}}{1+\theta_{12}B^{12}} \qquad (19)$$

$$\pi_1(B) = \frac{1-B}{1+\theta_1 B} = 1-(1-\theta_1)+\theta_1(1+\theta_1)B^2-\theta_1^2(1+\theta_1)B^3+\cdots \qquad (20)$$

# Seasonal Processes Forecasting. ARIMA(0,1,1)(0,1,1)$_1$2 (4/4)



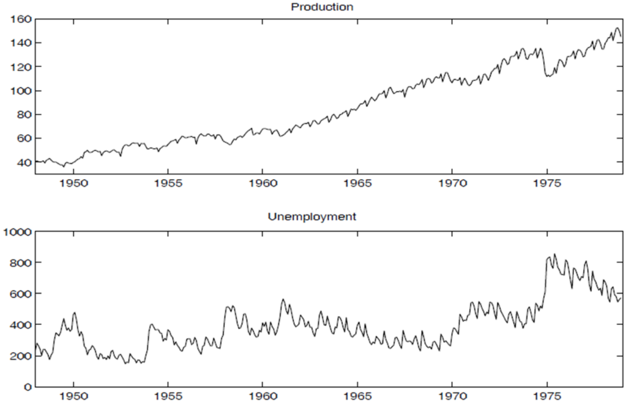

Weights $Z$ and $\Psi$ for the model ARIMA(0,1,1)(0,1,1)$_{12}$ for the parameters above


IDENTIFICATION
Suggest one or more models to fit the series
ESTIMATION
of the parameters of the proposed models
NON ADEQUATE MODEL
VALIDATION
Check if the residuals behave like white noise
SUITABLE MODEL
FORECASTS

# Outline

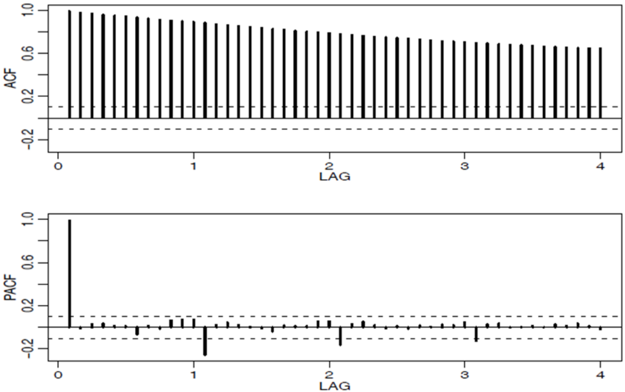
- The Federal Reserve Board Production Index example

# The Federal Reserve Board Production Index example (1/7)



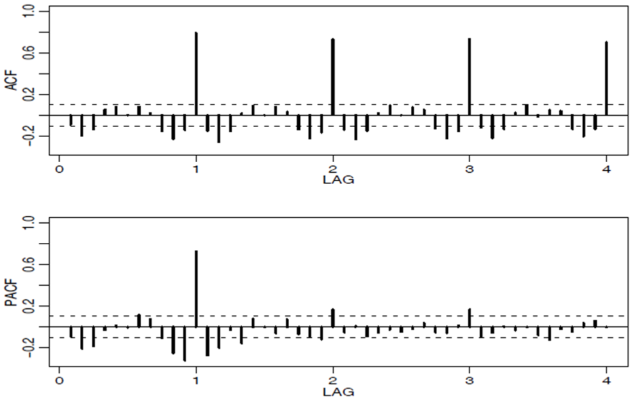

Values of the **Monthly Federal Reserve Board Production Index and Unemployment** (1948-1978, n = 372 months).

# The Federal Reserve Board Production Index example (2/7)



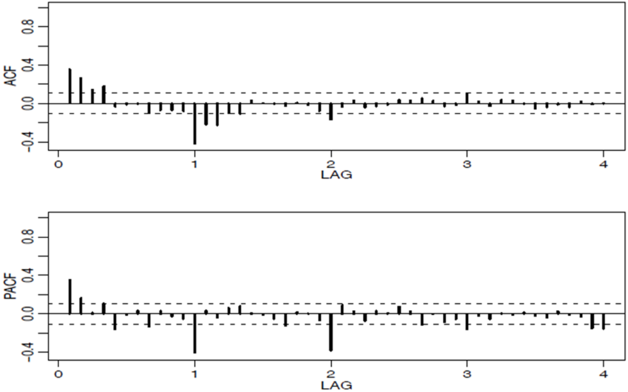

**ACF** and **PACF** of the **production series**

# The Federal Reserve Board Production Index example (3/7)




**ACF** and **PACF** of **differenced production**, $(1 - B)x_t$

# The Federal Reserve Board Production Index example (4/7)




**ACF** and **PACF** of the **first differenced and then seasonally differenced production**, $(1-B)(1-B^{12})x_t$

# The Federal Reserve Board Production Index example (5/7)

Three **models suggested** by these observations:

1. **ARIMA(2,1,0) x (0,1,1)**$_{12}$:

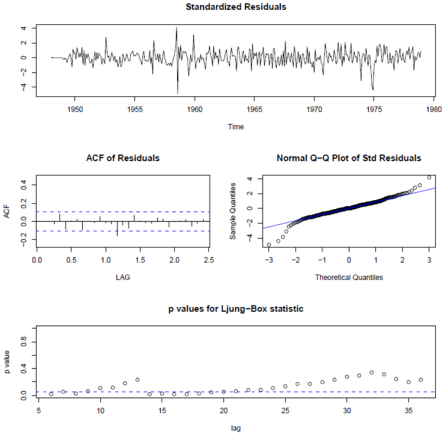
   AIC$=1.372$, BIC$=0.404$
2. **ARIMA(2,1,0) x (0,1,3)**$_{12}$:

   AIC$=1.299$, BIC$=0.351$
3. **ARIMA(2,1,0) x (2,1,1)**$_{12}$:

   AIC$=1.326$, BIC$=0.379$

**Model that best fits**: ARIMA(2,1,0)x(0,1,3)$_1$2

# The Federal Reserve Board Production Index example (6/7)




**Diagnostics for the ARIMA(2,1,0) x (0,1,3)$_{12}$** on the Production Index.

**Residual Analysis**

- Homogeneity of variance
- Normality
- Independence

**Homogeneity of variance. Tools**

- Residuals plot
- Square root of absolute values of the residuals with smooth fit
- ACF/PACF of square of residuals

**Homogeneity of variance. Problems**

- Outlier observations: Outlier detection and treatment
- Volatility: Models for the variance (GARCH)

**Normality. Tools**

- Quantile-Quantile plot
- Histogram with theoretical density overlapped
- Shapiro-Wilks test

**Normality. Problems**

- Outlier observations: Outlier detection and treatment
- Asymmetry or bi-modality: Transformation, outliers or change residuals distribution
- Heavy tails (excess of kurtosi): Volatility models or t-distributions for the residuals

**Independence. Tools**

ACF/PACF of residuals

LJung-Box test

Durbin-Watson test

**Independence. Problems**

Significant lags: Re-identify or add parameters to the model

**Information Criterion**

Balance between godness-of-fit and simplicity of the model

Component for the godness-of-fit: $-2\log L(\phi,\theta,\sigma^2|X)$

Component for the simplicity: $K * p$

$$AIC = -2\log L(\phi,\theta,\sigma^2|X) + 2p$$
$$BIC = -2\log L(\phi,\theta,\sigma^2|X) + log(N)p$$