Macroeconometria – Séries de tempo

FAUSTO JOSÉ ARAÚJO VIEIRA

Aula 2

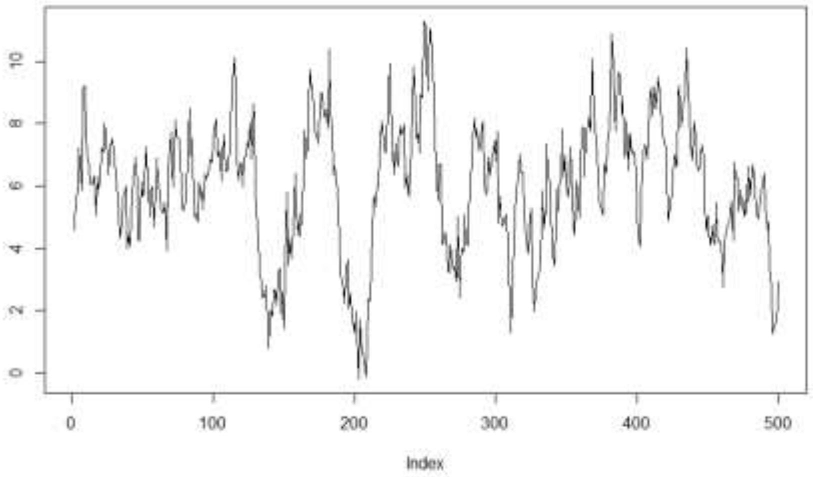
17 de abril a 22 de maio de 2018

RESUMO – SÉRIES DE TEMPO (AULA ANTERIOR)

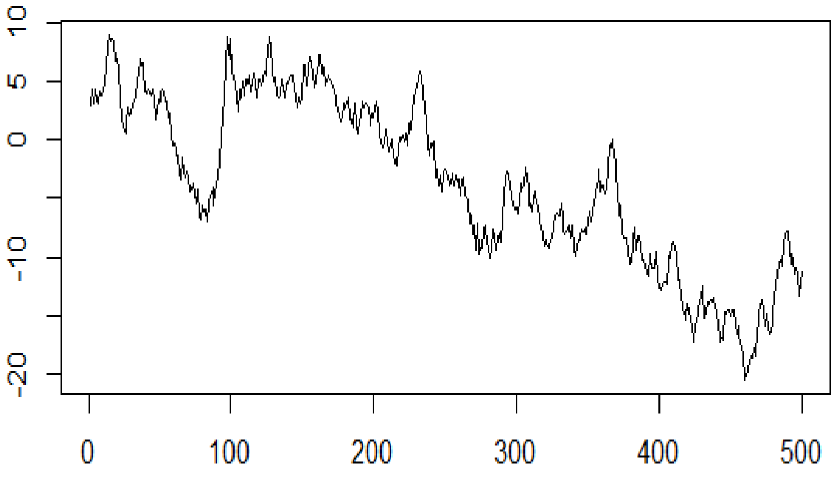
Estacionaridade x Tendência

Sem tendência

$$x_t = \mu + \beta\, x_{t-1} + u_t \text{ , onde } \beta < 1$$

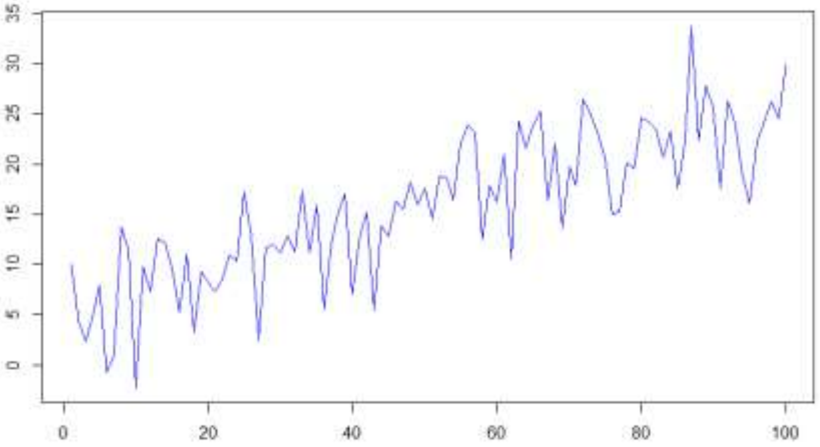


$$x_t = x_{t-1} + u_t$$

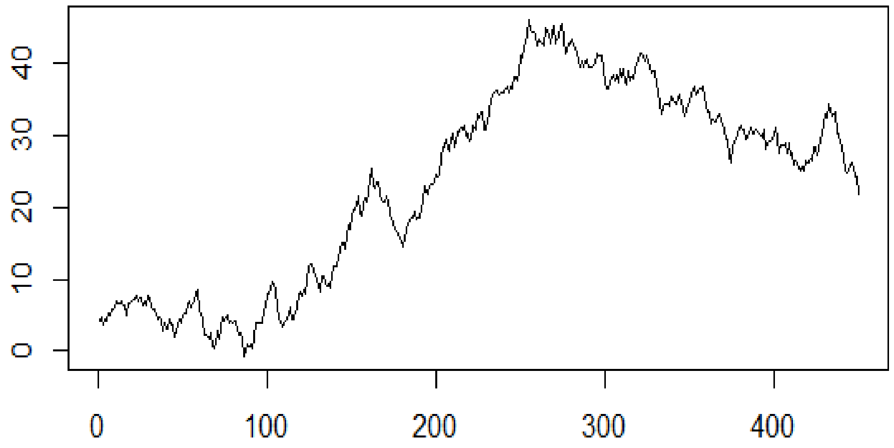


Com tendência

$$x_t = \mu + \delta t + u_t$$



$$x_t = \mu + x_{t-1} + u_t$$



Alguns conceitos básicos

1. Autocovariância;
2. Estacionariedade;
3. Ruído branco;
4. Ergodicidade;
5. Processo de médias móveis;
6. Processo autoregressivo;
7. Invertibilidade.

Autocovariância e autocorrelação

- Autocovariância
  - $\gamma_k = E(Y_t - \mu_t)(Y_{t-k} - \mu_{t-k}) = \mathrm{cov}(Y_t, Y_{t-k})$
- Autcorrelação
  - $\rho(k) = \dfrac{\gamma(k)}{\gamma(0)}$
- Matriz de convariância (dividido pelo desvio padrão – autocorrelação):

$$\Sigma_{XX} = \sigma^2 \begin{bmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(p-1) \\ \rho(1) & 1 & \rho(1) & \dots & \rho(p-2) \\ \rho(2) & \rho(1) & 1 & \ddots & \\ \vdots & & & \ddots & \rho(1) \\ \rho(p-1) & & & \rho(1) & 1 \end{bmatrix}$$

  - Pressupõe estacionariedade

Estacionaridade

- Propriedades estatísticas são invariantes no tempo
- Média finita e constante
  - $E(Y_t) = \mu \ \ para\ todo\ t$
- Variância finita
  - $E(Y_t)^2 < \infty$
- Autocovariância
  - $E(Y_t - \mu)\left(Y_{t-j} - \mu\right) = \gamma_j$ para todo t e j
- Ambos independem de t e são finitos – fracamente estacionário;
- Descrição de fracamente estacionário x estritamente estacionário
  - Estritamente estacionário: as funções de distribuição da série com mesmo tamanho de intervalo temporal terão propriedades estatísticas similares.
  - Para todo j, A função $\left(Y_{t_1}, Y_{t_2}, \cdots, Y_{t_n}\right)$ é semelhante a função de $\left(Y_{t_{1-j}}, Y_{t_{2-j}}, \cdots, Y_{t_{n-j}}\right)$

Ruído branco

- Média zero, variância constante e finta, não há autocorrelação

- $E(\varepsilon_t) = 0$

- $E(\varepsilon_t^2) = \sigma^2$

- $E(\varepsilon_t \varepsilon_k) = 0$ para todo t diferente de k

Ergodicidade

- O processo aleatório manterá as propriedades estatísticas da população;
- Ou seja, como as séries temporais econômicas são únicas, espera-se que a média temporal das observações seja uma boa aproximação da média populacional.
  - $E(\bar{y}^a) = \underset{T\to\infty}{\text{p lim}} \frac{1}{T}\sum_{i=0}^{T} y_t^a = E(y_t)$, onde "a" representa a amostra
- Em termos práticos, um processo é ergódico para a média desde que a autocovariância convirja suficientemente rápido para zero.
- Ruído branco é ergódico para todos os momentos.
- Ergodicidade não implica necessariamente em estacionariedade, nem o contrário.

Operador de defasagem e invertibilidade

- Seja z_t uma variável aleatória;
- $Lz_t = z_{t-1}$
- $L^j z_t = z_{t-j}$ para j pertencentes aos naturais
    - Exemplo:

$$z_t + \alpha_1 z_{t-1} + \alpha_2 z_{t-2} + \alpha_3 z_{t-3} = (1 + \alpha_1 L + \alpha_2 L^2 + \alpha_3 L^3) z_t$$

- Se | α |<1, então:
    - $(1 + \alpha\, L)^{-1} = 1 + \alpha\, L + \alpha^2 L^2 + \alpha^{\,3} L^3 \ldots$
    - Exemplo:

$$z_t + \alpha\, z_{t-1} = e_t$$

$$(1 + \alpha\, L) z_t = e_t$$

$$z_t = \frac{e_t}{(1+\alpha\, L)} = e_t (1 + \alpha L + \alpha^{\,2} L^2 + \alpha^3\, L^3 \ldots) = e_t + \alpha e_{t-1} + \alpha^{\,2} e_{t-2} + \alpha^{\,3}\, e_{t-3}\ \ldots$$

Estimar um modelo ARMA

- Para estimar uma série temporal, é necessário que o processo seja:
  - Estocástico;
  - Estacionário;
  - Ergótico.

RESUMO DA AULA

Resumo

1. Processo de Médias Móveis;
2. Processo Autoregressivo;
3. Invertibilidade;
4. Função de Autocorrelação (FAC) e Autocorrelação Parcial (FACP);
5. Estimação dos modelos;
6. Critério de informação;
7. Projeção.

PROCESSO AUTORREGRESSIVO E MÉDIAS MÓVEIS

Médias Móveis – MA(1)

- Considere o seguinte processo estocástico:

  $y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}$ onde ε_t é um ruído branco ~N(0, σ^2);

- Média móvel mais geral do que normalmente escutamos:
  - Exemplo: média móvel de 3 meses da produção industrial.
  - Pesos não são necessariamente idênticos.
- Processo é estacionário?
  - Média:
  - $E(y_t) = \mu + E(\varepsilon_t) + \theta E(\varepsilon_{t-1}) = \mu$

Médias Móveis – MA(1)

- Função

$y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}$

- Variância

$Var(y_t) = (y_t - \mu)^2$
$= (\varepsilon_t + \theta\varepsilon_{t-1})^2$
$= (\varepsilon_t^2 + 2\theta\varepsilon_t\varepsilon_{t-1} + \theta^2\varepsilon_{t-1}^2)$
$= \sigma^2 + 0 + \theta^2\sigma^2$

$= \sigma^2(1 + \theta^2)$

- Covariância

$= E[(y_t - \mu)(y_{t-1} - \mu)]$
$= E[(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-1} + \theta\varepsilon_{t-2})]$
$= E[(\varepsilon_t\varepsilon_{t-1} + \theta\varepsilon_{t-1}^2 + \theta\varepsilon_t\varepsilon_{t-2})]$
$= \theta\sigma^2$

Médias Móveis – MA(1)

$$y_t = \mu + \varepsilon_t + \theta \varepsilon_{t-1}$$

- Processo é estacionário?

  Sim, média constante, variância finita e covariância independe de t

- Processo é ergótico?

  Sim, a autocovariância converge rapidamente para zero

- IMPORTANTE!!!
  - Autocovariância
  - $\rho_1 = {}^{\gamma_1}/_{\gamma_0}$
    $$= \frac{\theta\sigma^2}{\sigma^2(1+\theta^2)} = \frac{\theta}{(1+\theta^2)}$$
  - ρ_j=0 para todo j>1

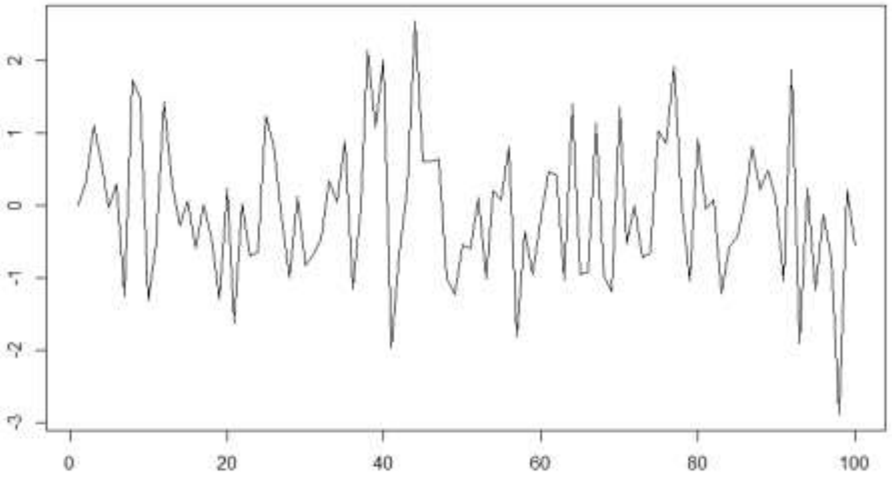
Médias Móveis – MA(1)

- Criar um script no R
- Digitar as seguintes linhas:
  - y<- e <- rnorm(100,0,1)
  - y[0]=0
  - y[1]=0
  - for (t in 2:100) y[t] <- e[t]+0.1*e[t-1]
  - plot(y,type= "l")
- Colocar a primeira linha como comentário "#" para usar diferentes coeficientes, mas com a mesma amostra aleatória.
- Salvar os valores: y_ma1_0_3=y

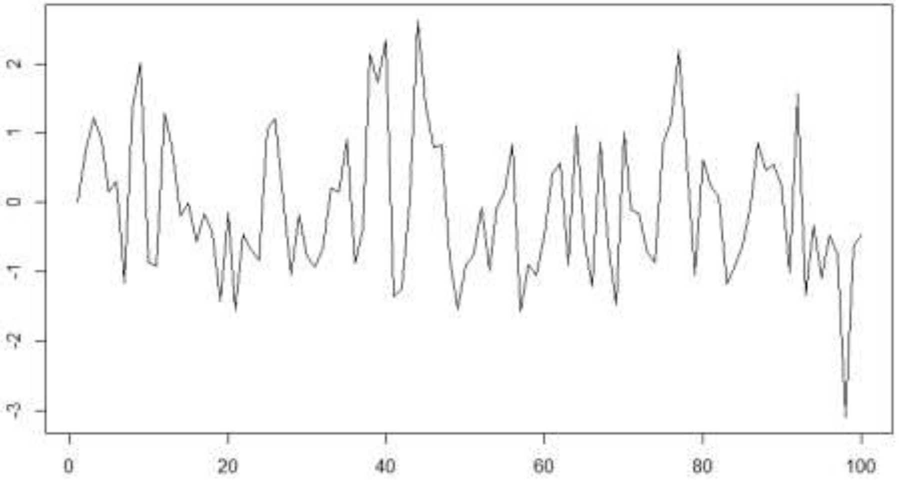
  y_ma1_0_9=y

Médias Móveis – MA(1)

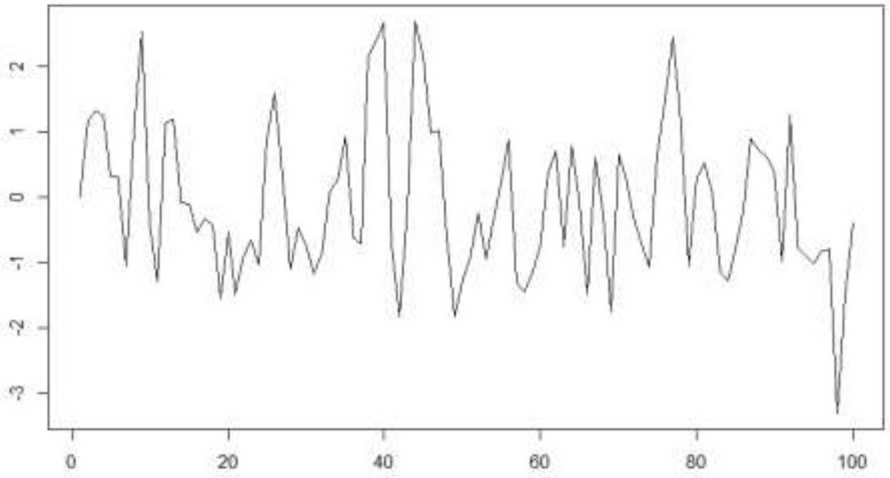
$$y_t = \mu + \varepsilon_t$$



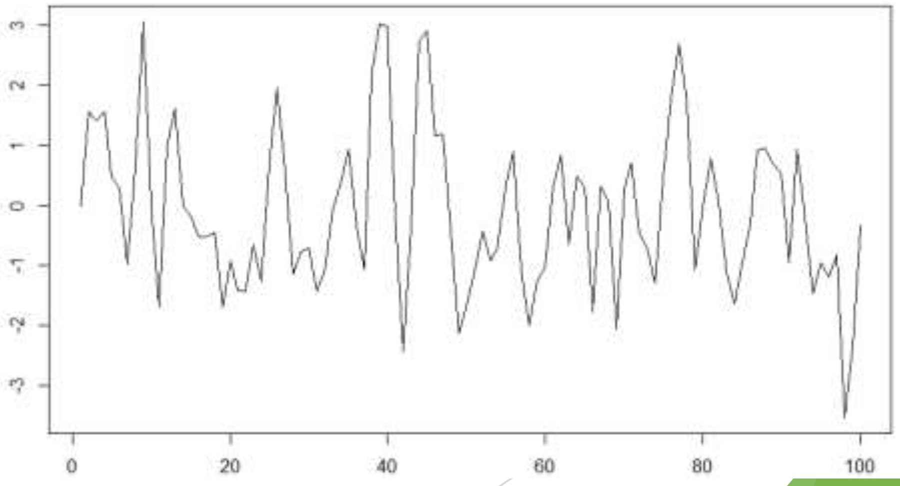
$$y_t = \mu + \varepsilon_t + 0.3\varepsilon_{t-1}$$



$$y_t = \mu + \varepsilon_t + 0.6\varepsilon_{t-1}$$



$$y_t = \mu + \varepsilon_t + 0.9\varepsilon_{t-1}$$



Médias Móveis – MA(1)

- Séries ficaram mais persistentes com a elevação do coeficiente do MA(1)
- Variância também aumenta:
  - A escala dos gráficos aumenta com o coeficiente do MA(1)
  - = $\sigma^2(1+\theta^2)$
- Autocorrelação aumenta com o valor maior de θ

Médias Móveis – MA(q)

- Definição do modelo geral

$$y_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}$$, onde $\theta_1, \theta_2, \ldots, \theta_q \ (\theta_q \neq 0)$

- Usando o operador de defasagem

$$y_t = \theta(B)\epsilon_t$$, onde $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$

- Momentos do MA(q)
  - Média:

$$\begin{aligned} E[y_t] &= E[\epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \ldots + \theta_q \epsilon_{t-q}] \\ &= E[\epsilon_t] + \theta_1 E[\epsilon_{t-1}] + \theta_2 E[\epsilon_{t-2}] + \ldots + \theta_q E[\epsilon_{t-q}] = 0 \end{aligned}$$

Médias Móveis – MA(q)

- Momentos do MA(q)
  - Variância:

$$Var[y_t] = Var[\epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \ldots + \theta_q\epsilon_{t-q}]$$

$$= Var[\epsilon_t] + \theta_1^2 Var[\epsilon_{t-1}] + \theta_2^2 Var[\epsilon_{t-2}] + \ldots + \theta_q^2 Var[\epsilon_{t-q}]$$

$$= (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)\sigma^2$$ - Covariância é nula (Ruído Branco)

  - Covariância:

$$\gamma_j = E\left(\sum_{i=0}^{q} \theta_i\epsilon_{t-1} \sum_{i=0}^{q} \theta_i\epsilon_{t-i-j}\right) = E(\theta_j\epsilon_{t-j}^2 + \theta_{j+1}\,\theta_1\epsilon_{t-j-1}^2 + \cdots + \theta_q\theta_{q-j}\epsilon_{t-q}^2)$$

$$\gamma_j = E(\theta_j + \theta_{j+1}\,\theta_1 + \theta_{j+2}\,\theta_2 + \cdots + \theta_q\theta_{q-j})\sigma^2$$ para j=1,2,...,q

$$\gamma_j = 0$$ para j>q

Médias Móveis – MA(q)

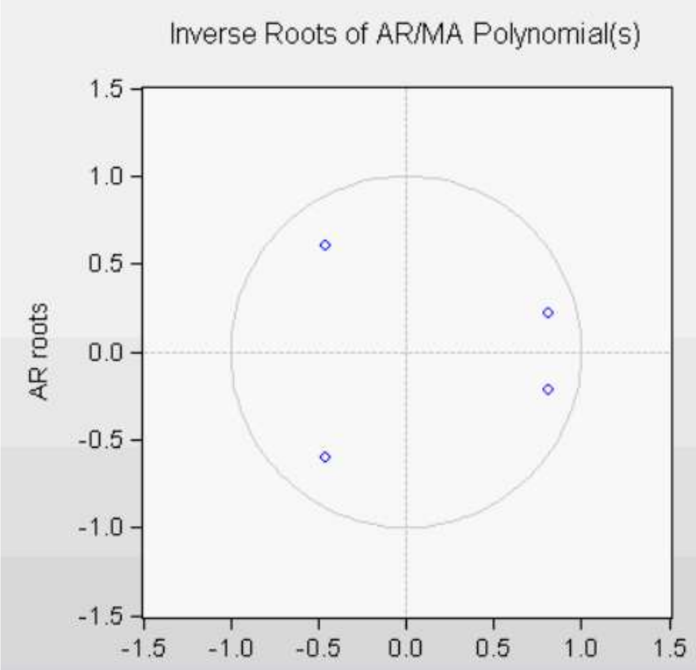
$$y_t = \theta_0 + \epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} - \theta_3\epsilon_{t-3} - \cdots - \theta_q\epsilon_{t-q}$$

- MA(q) é estacionário?
  - (Hamilton, 1994): As raízes de $\theta(L)$ =0 tem que estar fora do círculo unitário.

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \theta_3 L^3 - \cdots - \theta_q L^q$$

  - (Enders, 2009): As raízes de $\theta(L)$ =0 tem que estar dentro do círculo unitário.

$$\theta(L) = \lambda^2 - \theta_1\lambda - \theta_2\lambda^2 - \theta_3\lambda^3 - \cdots - \theta_q\lambda^q$$



Exemplos de MA(2)

- MA(2)

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2)\sigma^2$$

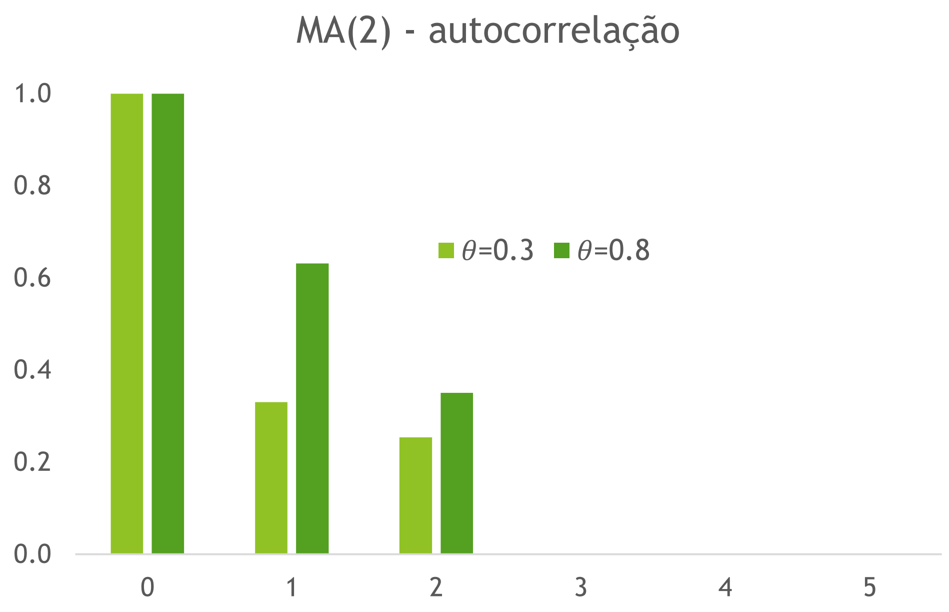
$$\gamma_1 = (\theta_1 + \theta_1\theta_2)\sigma^2$$

$$\gamma_2 = (\theta_2)\sigma^2$$

$$\gamma_j = 0, para\ j > 2$$

- Autocorrelação é dividir por γ_0
- **Importante!!!**
  - A defasagem do MA é determinada pelo autocorrelação

$$\rho_j = \begin{cases} 1, & j = 0 \\ \frac{\theta_1+\theta_1\theta_2}{1+\theta_1^2+\theta_2^2}, & j = 1 \\ \frac{\theta_2}{1+\theta_1^2+\theta_2^2}, & j = 2 \\ 0, & j > 2. \end{cases}$$



Autorregressivo – AR(1)

- Considere o seguinte processo estocástico:

$y_t = c + \varphi y_{t-1} + \varepsilon_t$ onde ε_t é um ruído branco ~N(0, σ^2) e | φ|<1;

- Usando o operador de defasagem e a invertibilidade:

$$y_t = c + \varphi y_{t-1} + \varepsilon_t \rightarrow (1 - \varphi \text{L})\, y_t = c + \varepsilon_t$$

$$y_t = \frac{c}{1-\varphi} + \sum_{j=0}^{\infty} \varphi^j \varepsilon_{t-j} = \mu + \vartheta(L)\varepsilon_t, desde\ que\ |\varphi| < 1$$

- Prova (sem constante):

$$y_t = \phi y_{t-1} + u_t$$
$$y_{t-1} = \phi y_{t-2} + u_{t-1}$$
$$y_t = \phi(\phi y_{t-2} + u_{t-1}) + u_t$$

$\Longrightarrow$

$$y_t = \phi^2 y_{t-2} + \phi u_{t-1} + u_t$$
$$y_t = \phi^m y_{t-m} + \sum_{j=0}^{m} \phi^j u_{t-j},\ |\phi|<1$$
$$y_t = \sum_{j=0}^{\infty} \phi^j u_{t-j}$$

Autorregressivo – AR(1)

$$y_t = c + \varphi y_{t-1} + \varepsilon_t$$

- É estacionário?
- Média:

$$E(y_t) = \mu + \sum_{j=0}^{\infty} \varphi^j E(\varepsilon_{t-j}) = \mu$$

- Variânica:

$$Var(y_t) = E(y_t - \mu)^2 = E\left(\sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}\right)^2 = \sum_{j=0}^{\infty} \phi^{2j} E\left(\varepsilon_{t-j}^2\right) = \frac{\sigma^2}{1-\phi^2}$$

Autorregressivo AR(1)

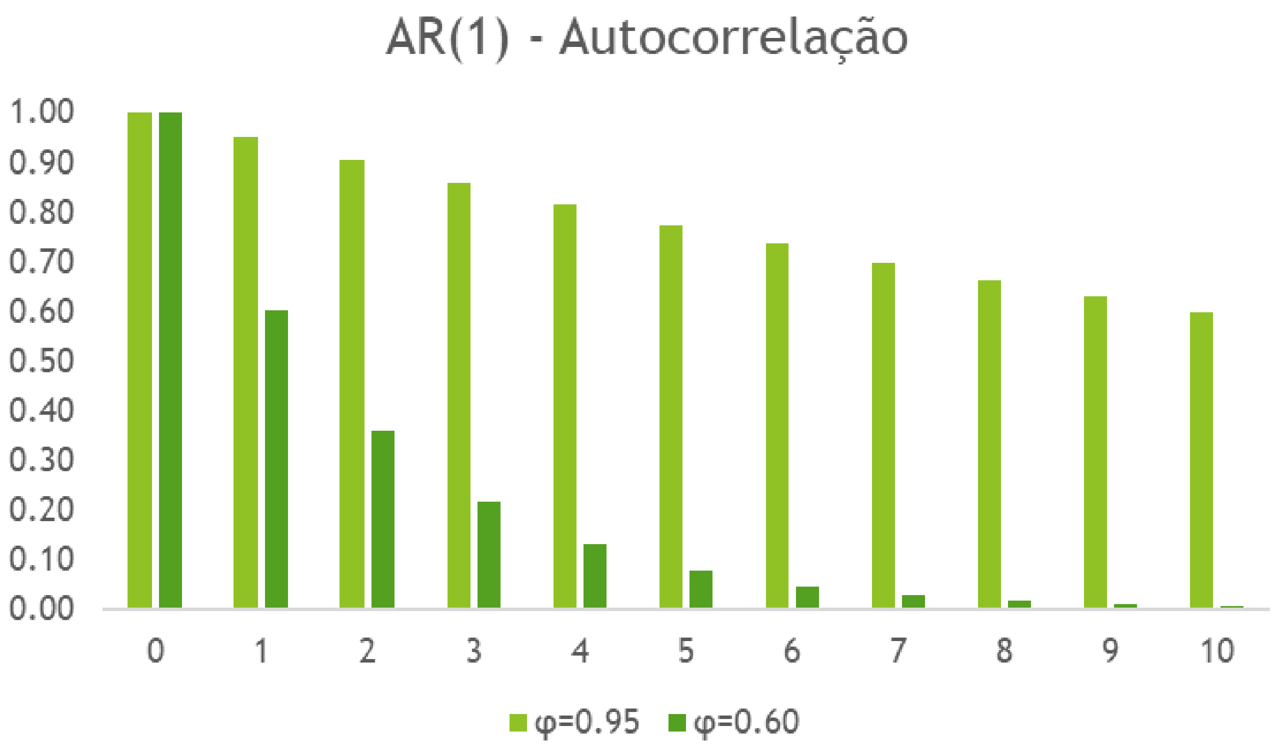
- Covariância:

$$
\begin{aligned}
E\left[(y_t-\mu)(y_{t-j}-\mu)\right] &= E\left[\left(\sum_{s=0}^{\infty}\phi^s\varepsilon_{t-s}\right)\left(\sum_{s=0}^{\infty}\phi^s\varepsilon_{t-s-j}\right)\right]= \\
&= \sigma^2\left(\phi^j+\phi^{j+2}+\phi^{j+4}+\cdots\right)=\left(\frac{\phi^j}{1-\phi^2}\right)\sigma^2
\end{aligned}
$$

- Autocorrelação:

$$
\rho_j=\frac{\left(\frac{\phi^j}{1-\phi^2}\right)\sigma^2}{\frac{\sigma^2}{1-\phi^2}}=\phi^j,\quad j=1,2,\ldots
$$

Autocorrelação - AR(1)



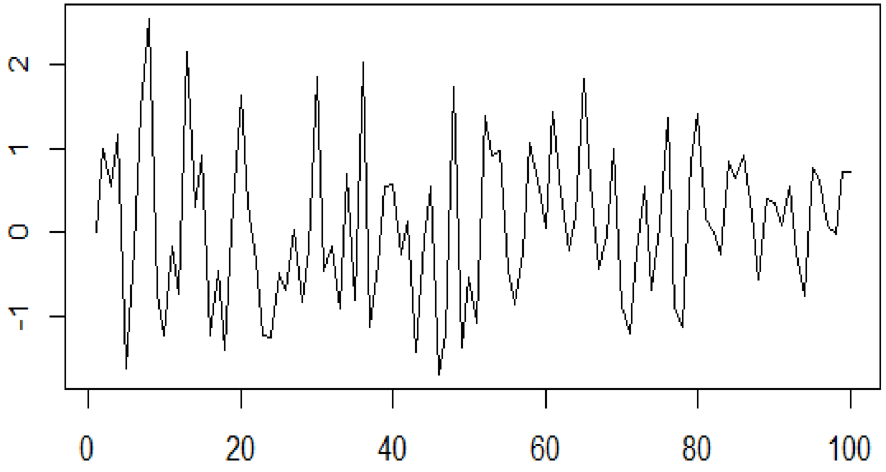
Autorregressivo AR(1)

- Criar um script no R
- Digitar as seguintes linhas:
  - y<- e <- rnorm(100,0,1)
  - y[0]=0
  - y[1]=0
  - for (t in 2:100) y[t] <- e[t]+0.1*y[t-1]
  - plot(y,type= "l")
- Colocar a primeira linha como comentário “#” para usar diferentes coeficientes, mas com a mesma amostra aleatória.
- Salvar os valores: y_ar1_0_4=y

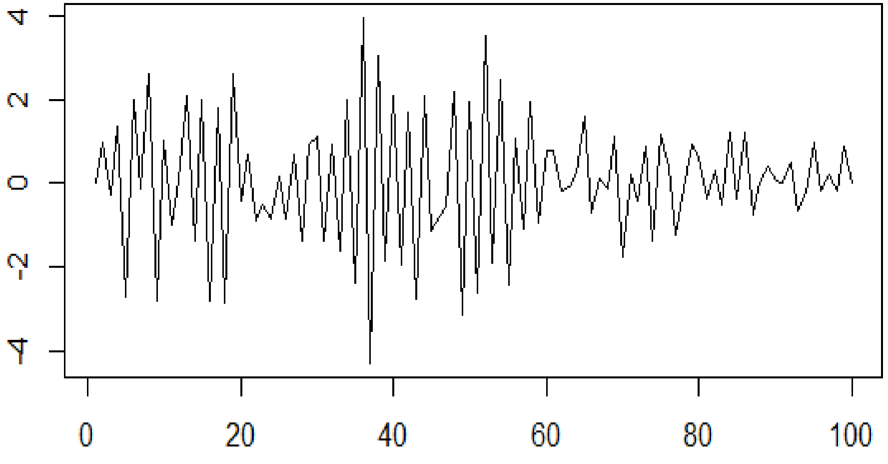
  y_ar1_0_8=y

Autorregressivo AR(1)

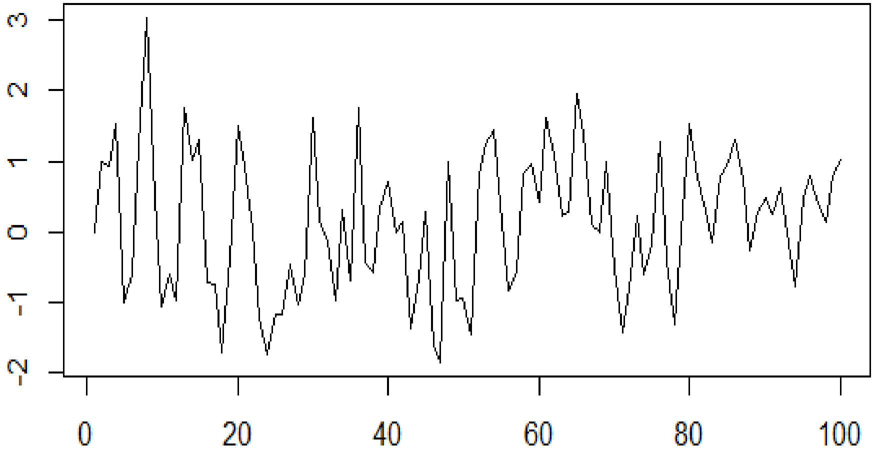
$$y_t = \mu + \varepsilon_t$$



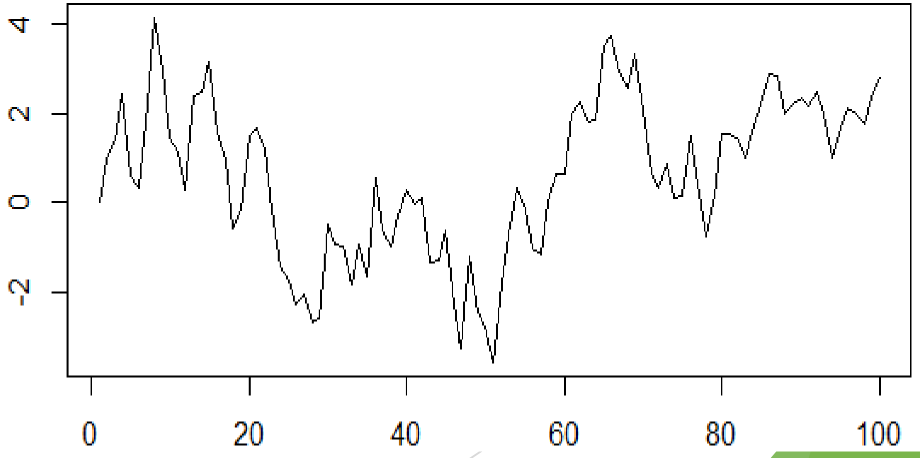
$$y_t = \mu + \varepsilon_t - 0.8y_{t-1}$$



$$y_t = \mu + \varepsilon_t + 0.4y_{t-1}$$

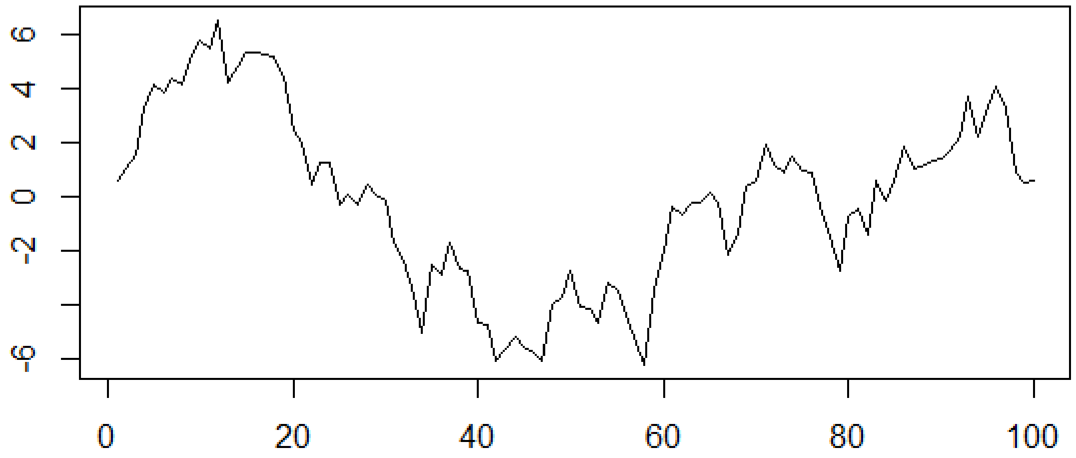


$$y_t = \mu + \varepsilon_t + 0.8y_{t-1}$$



Autorregressivo AR(1)

- Outra forma de criar um AR(1)
  - x1<-arima.sim(model=list(ar=0.9),n=100)
  - plot(x1)



Autorregressivo AR(1)

- Séries ficaram mais persistentes com a elevação do coeficiente do AR(1)
  - Coeficiente negativo têm maior influência das mudanças.
  - Dos dois gráfico à direita, tem-se a impressão de maior volatilidade do gráfico superior, no entanto, a volatilidade das séries é a mesma.
- Variância também aumenta:
  - A escala dos gráficos aumenta com o coeficiente do AR(1) – valor absoluto
- Autocorrelação aumenta com o valor maior de φ
  - O decaimento é lento quando o valor $|\varphi|$ é próximo de 1
  - Todo AR tem um decaimento infinito, pois é pode ser representado por um MA(∞)

Momentos do AR(2)

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$$

- Média

$$\begin{aligned} E(y_t) &= c + \phi_1 E(y_{t-1}) + \phi_2 E(y_{t-2}) + E(\varepsilon_t) \Longrightarrow \\ E(y_t) &\equiv \mu = \frac{c}{1 - \phi_1 - \phi_2}. \end{aligned}$$

- Autocovariância:

Reescrevendo a equação (variável passando pela origem):

$$y_t - \mu = \phi_1 (y_{t-1} - \mu) + \phi_2 (y_{t-2} - \mu) + \varepsilon_t$$

Multiplicando ambos os lados por $(y_{t-j} - \mu)$ e tomando a esperança:

$$E(y_t - \mu)(y_{t-j} - \mu) = \phi_1 E(y_{t-1} - \mu)(y_{t-j} - \mu) +$$
$$+\phi_2 E(y_{t-2} - \mu)(y_{t-j} - \mu) + E[\varepsilon_t (y_{t-j} - \mu)].$$

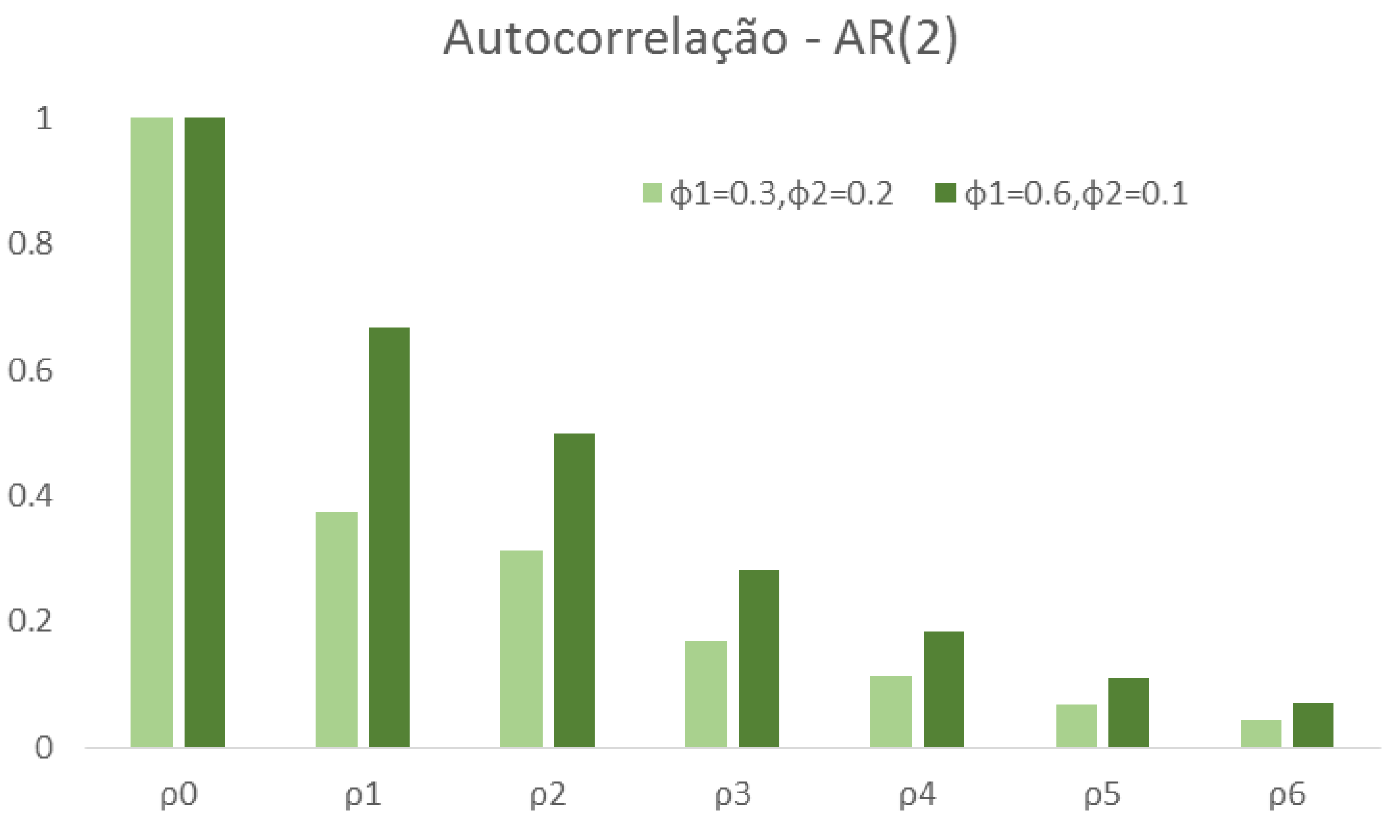
Então:

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2}, \; j = 1, 2, \ldots$$

Dividindo por γ_0

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2}, \; j = 1, 2, \ldots$$

Exemplo de AR(2) - Autocorrelação

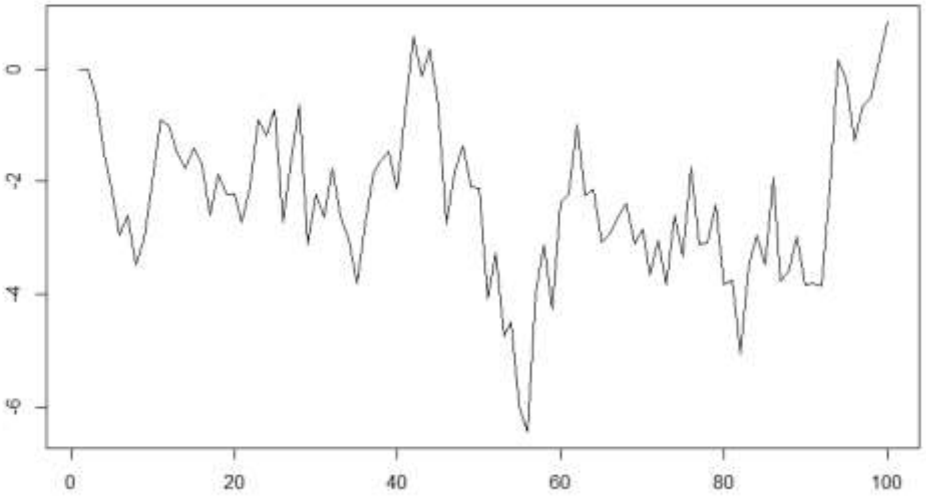


Autorregressivo AR(2)

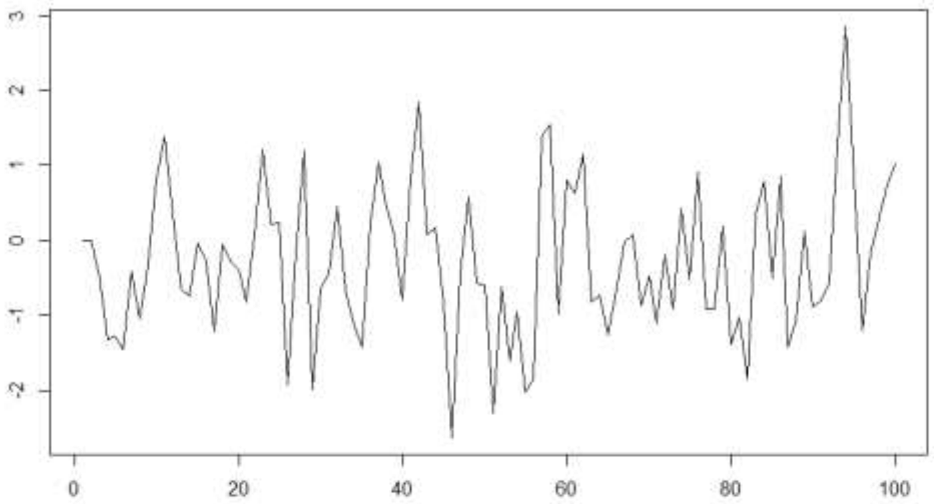
- Criar um script no R
- Digitar as seguintes linhas:
  - y<- e <- rnorm(100,0,1)
  - y[0]=0
  - y[1]=0
  - for (t in 3:100) y[t] <- e[t]+0.7*y[t-1] +0.2*y[t-2]
  - plot(y,type= "l")
- Colocar a primeira linha como comentário “#” para usar diferentes coeficientes, mas com a mesma amostra aleatória.
- Salvar os valores: y_ar2_0_7_0_2=y

Autorregressivo AR(2)

$$y_t = \varepsilon_t + 0.7y_{t-1} + 0.2y_{t-2}$$



$$y_t = \varepsilon_t + 0.3y_{t-1} - 0.1y_{t-2}$$



Autorregressivo – AR(p)

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t = c + \sum_{j=1}^{p} \phi_j y_{t-j} + \varepsilon_t$$

- Raízes polinomiais:

$\left(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p\right)$, estacionário se as raízes estiverem fora do círculo unitário

- Média

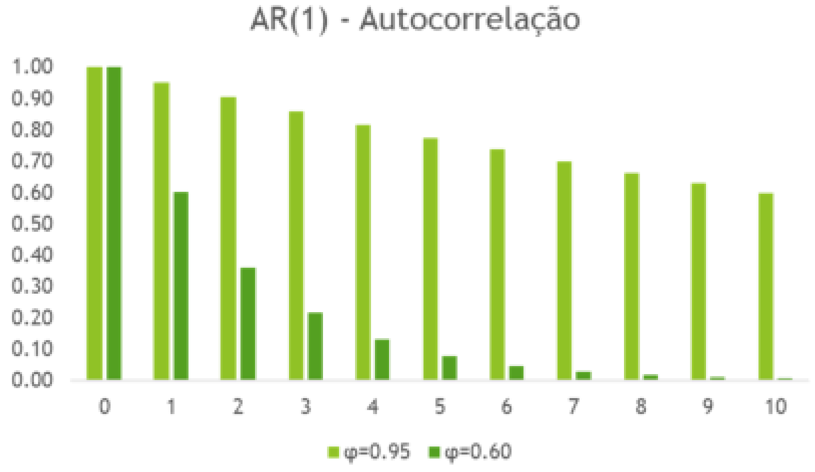
$$y_t = \mu + \psi(L)\,\varepsilon_t, \qquad \mu = \frac{c}{1-(\phi_1+\phi_2+\cdots+\phi_p)}$$

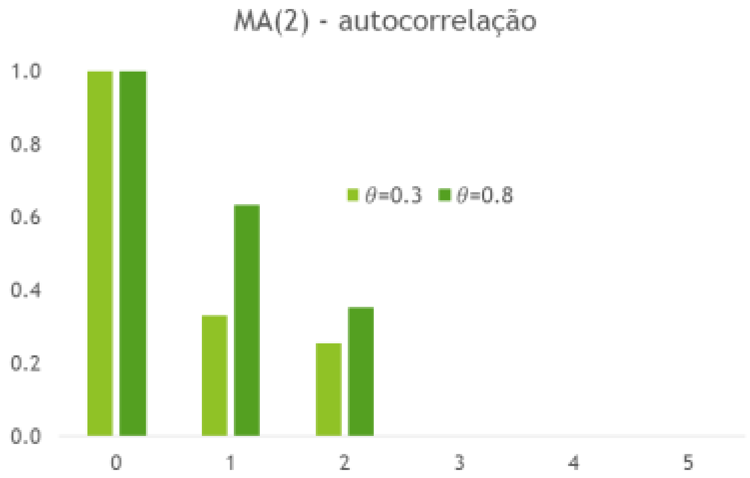
- Autocorrelação:

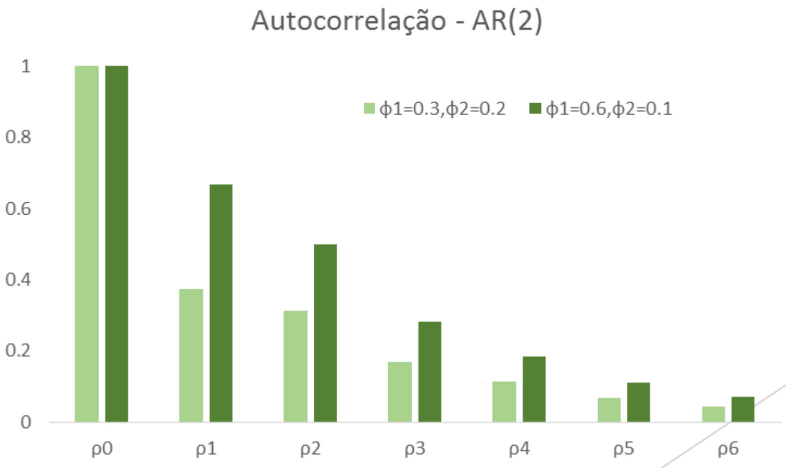
$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \cdots + \phi_p \rho_{j-p}, \; j = 1, 2, \ldots, p$$

Autocorrelação – Não se esqueça!

- MA – ela interrompe na última defasagem do MA;
- AR – declinante com as defasagens;
- Maior o coeficiente, mais lento o declínio da autocorrelação;
- Variância aumento com o coeficiente.







Processo ARMA(p,q)

$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=0}^{q} \theta_j \varepsilon_{t-j}$$

- Invertibilidade e definição

$$y_t = \mu + \sum_{i=1}^{p} \phi_i \left(y_{t-i} - \mu \right) + \sum_{j=0}^{q} \theta_j \varepsilon_{t-j} = \mu + \psi \left(L \right) \varepsilon_t, \text{ onde } \psi\left(L\right) = \frac{1+\sum_{j=1}^{q} \theta_j L^j}{1-\sum_{i=1}^{p} \phi_i L^i}$$

- Média incondicional:

$$E\left(y_t\right) \equiv \mu = c + \sum_{i=1}^{p} \phi_i E\left(y_{t-i}\right) = \frac{c}{1-\sum_{i=1}^{p} \phi_i}$$

- Autocovariâcia:

$$\gamma_h = \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2} + \cdots + \phi_p \gamma_{h-p}, \quad h = q+1, q+2, \ldots$$

Covariância é declinante, como o AR(P)

ARMA(1,1) - Autocovariância

- Autocovariância declinante:

$$
\begin{aligned}
\gamma_0 &= \frac{1+\theta_1^2+2\phi_1\theta_1}{1-\phi_1^2}\sigma^2; \\
\gamma_1 &= \frac{(1+\phi_1\theta_1)(\phi_1+\theta_1)}{1-\phi_1^2}\sigma^2; \\
\gamma_2 &= \phi_1\gamma_1; \\
&\vdots \\
\gamma_h &= \phi_1^{h-1}\gamma_1.
\end{aligned}
$$

ARIMA(p,d,q)

- A grande diferença entre o ARIMA e o ARMA é se a variável analisada é estacionária ou integrada de grau 1 ou superior.
  - Veremos essa discussão na aula 4, quando falarmos sobre séries não estacionárias;

- Se uma variável aleatória X_t segue um modelo ARIMA(p,d,q), então $\Delta^d X_t$ segue um modelo ARMA(p,q);

- O “I” está relacionado com a quantidade de vezes necessárias para tornar a variável estacionária.

Função de Autocorrelação - FAC

- É o gráfico de autocorrelação em relação a defasagem
- Permite identificar a ordem q de um processo MA
  - Lembre-se, a FAC do AR tem decaimento exponencial em relação às defasagens.
- Em grandes amostras, a variância das estimativas de um RB pode ser aproximada por ser dado por T^{-1}

Modelo	FAC
$MA\ (q)$	$\rho_j = \frac{\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \cdots + \theta_q\theta_{q-j}}{\sum_{j=0}^{q}\theta_j^2}, j = 1, 2, \ldots, q$
$AR\ (1)$	$\rho_j = \phi^j, j = 1, 2, \ldots.$
$AR\ (p)$	$\rho_j = \phi_1\rho_{j-1} + \phi_2\rho_{j-2} + \cdots + \phi_p\rho_{j-p}, j = 1, 2, \ldots$
$ARMA\ (1, 1)$	$\begin{cases} \rho_1 = \frac{(1+\phi_1\theta_1)(\phi_1+\theta_1)}{1+\theta_1^2+2\phi_1\theta_1} \\ \rho_j = \phi_1\rho_{j-1} = \phi_1^{j-1}\rho_1, j > 1. \end{cases}$

Estimando a FAC

- Obtenha a média amostral da série:

$$\overline{y} = \frac{\sum_{t=1}^{T} y_t}{T}$$

- Calcule a autocorrelação amostral:

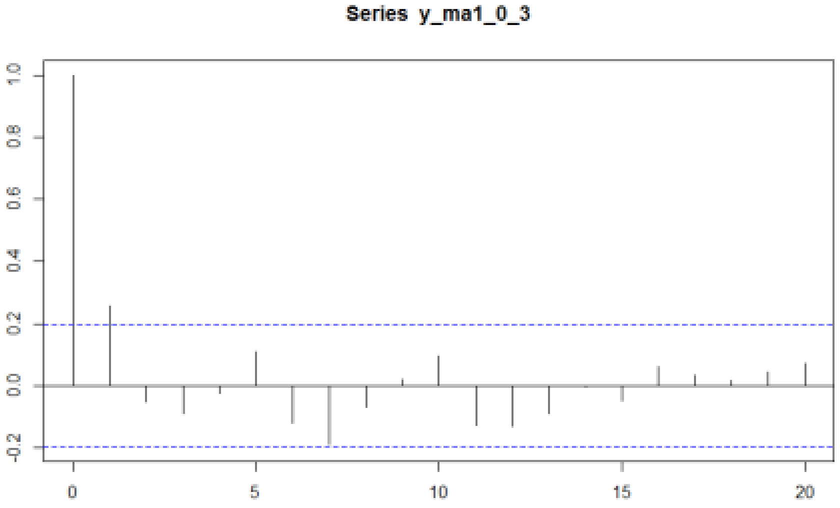
$$\widehat{\rho}_j : \widehat{\rho}_j = \frac{\frac{\sum_{t=j+1}^{T} (y_t - \overline{y})(y_{t-j} - \overline{y})}{T}}{\frac{\sum_{t=1}^{T} (y_t - \overline{y})^2}{T}}, \quad j = 1, 2, \ldots$$

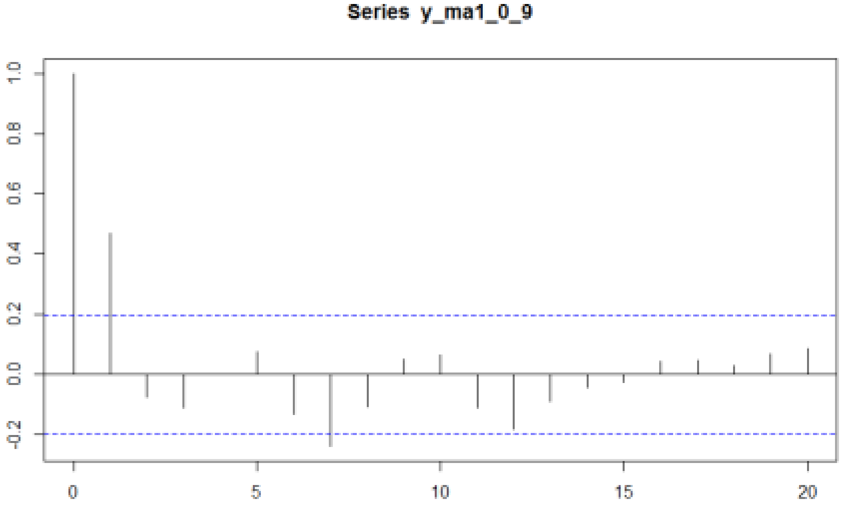
- Trace o gráfico de $\widehat{\rho}_j$ contra j.

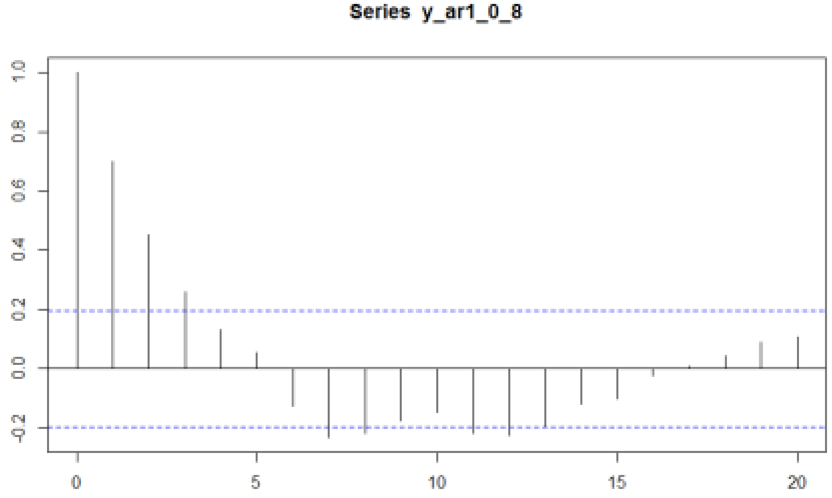
FAC dos modelos simulados

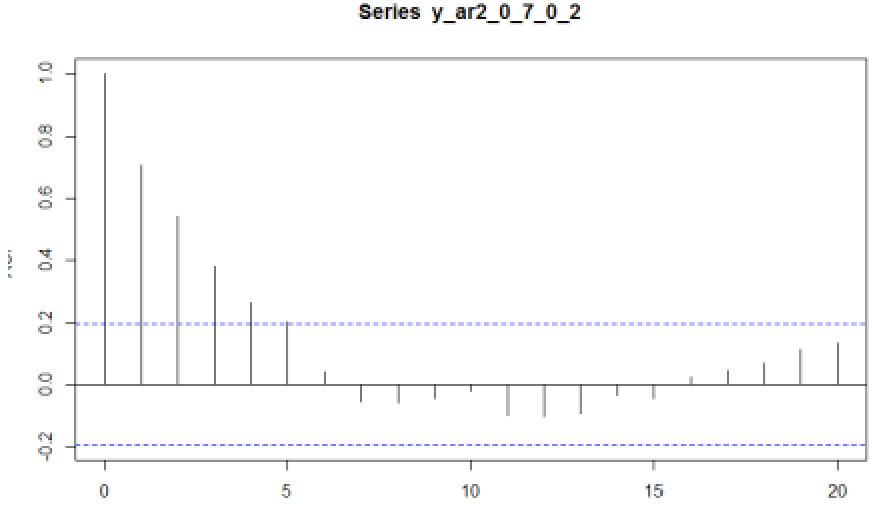
- Instalando o “Linear and Nonlinear Mixed Effects Model”
  - install.packages("nlme")
- Deixando-o ativo:
  - require(nlme)
- Estimando as FAC
  - acf(y_ar1_0_8)
  - acf(y_ma1_0_3)
  - acf(y_ma1_0_9)
  - acf(y_ar2_0_7_0_2)

Resultado do FAC – modelos simulados









Função de Autocorrelação Parcial - FACP

- A FACP elimina as correlações implícitas entre a variável e suas defasagens, tornando possível estimar o coeficiente da defasagem.
- Como se faz? Elimina-se as correlações implícitas entre duas variáveis.
- Equação: $y_t = \phi_{j,1} y_{t-1} + \phi_{j,2} y_{t-2} + \cdots + \phi_{j,j} y_{t-j} + e_t, \quad j = 1, 2, \ldots,$
- Procedimento consiste em regredir y_t contra y_{t-1} e obter $\hat{\Phi}_{1,1}$, depois estima-se y_t contra y_{t-1} e y_{t-2}, obtendo $\hat{\Phi}_{2,1}$ e $\hat{\Phi}_{2,2}$. A FACP apresenta $\hat{\Phi}_{1,1}$ e $\hat{\Phi}_{2,2}$, descartando $\hat{\Phi}_{2,1}$.

	y_{t-1}	y_{t-2}	y_{t-3}	…	y_{t-p}
1	$\hat{\Phi}_{1,1}$	0	0	…	0
2	$\hat{\Phi}_{2,1}$	$\hat{\Phi}_{2,2}$	0	…	0
3	$\hat{\Phi}_{3,1}$	$\hat{\Phi}_{3,2}$	$\hat{\Phi}_{3,3}$	…	0
…					
P	$\hat{\Phi}_{p,1}$	$\hat{\Phi}_{p,2}$	$\hat{\Phi}_{p,3}$	…	$\hat{\Phi}_{p,p}$

- A FACP é formada pela diagonal principal!!!

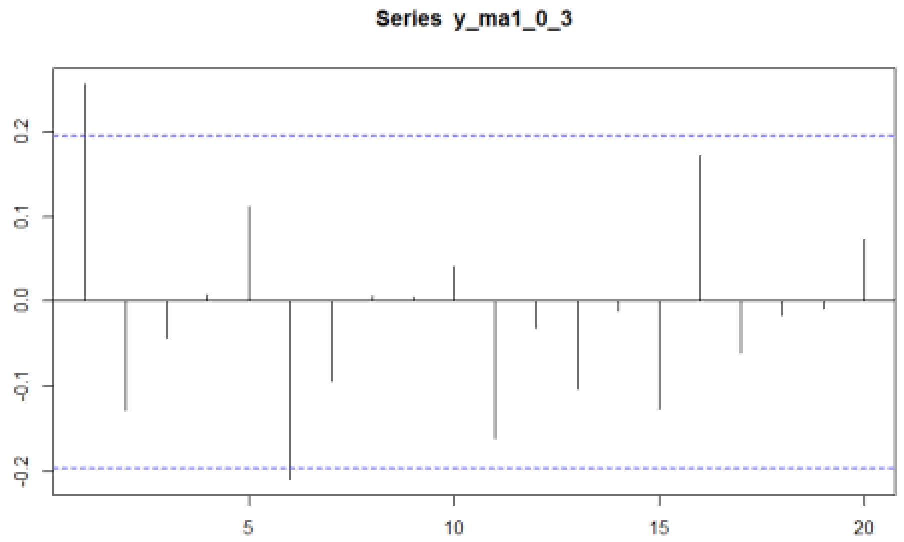
Função de Autocorrelação Parcial - FACP

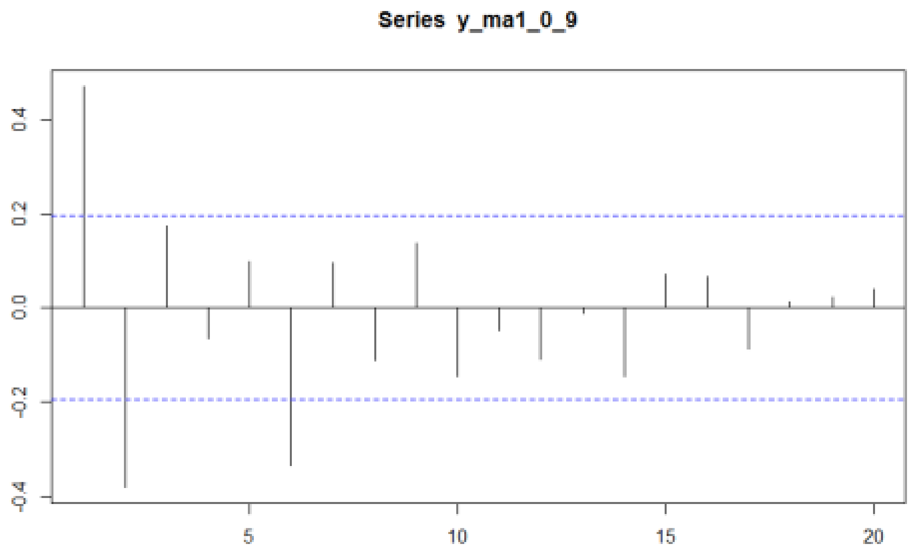
- A FACP determina a defasagem “p” do AR(p);
- Ou seja, espera-se que os coeficientes $\widehat{\Phi}_{j,j}$ para j<=p são diferentes de zero, já para os coeficientes j>p, são nulos.
- Devido a condição de invertibilidade, o MA(q) pode se tornar um processo AR infinito. Assim, o coeficiente $\widehat{\Phi}_{j,j}$ terá um decaimento exponencial.
- Qual é o total de parâmetros j para ser estimado?
  - Enders (2009) sugere calcular a FACP até j=T/4, em que T é o tamanho da amostra.

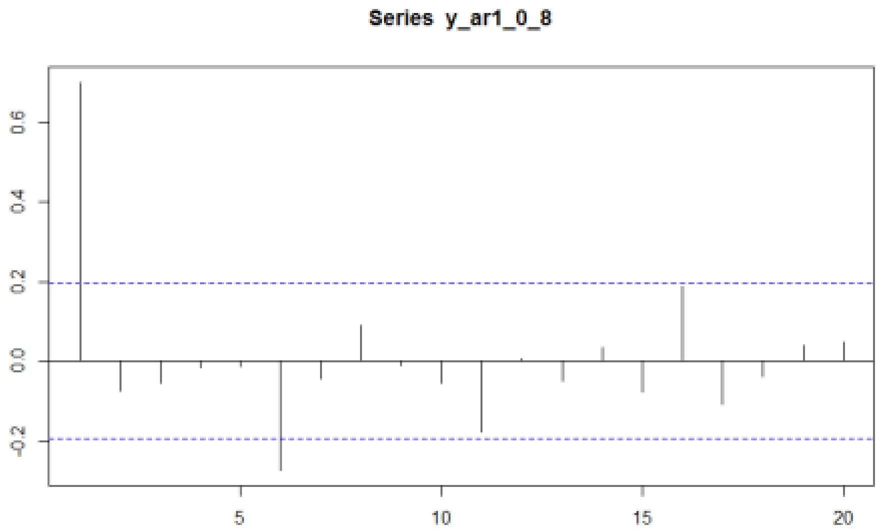
FAC dos modelos simulados

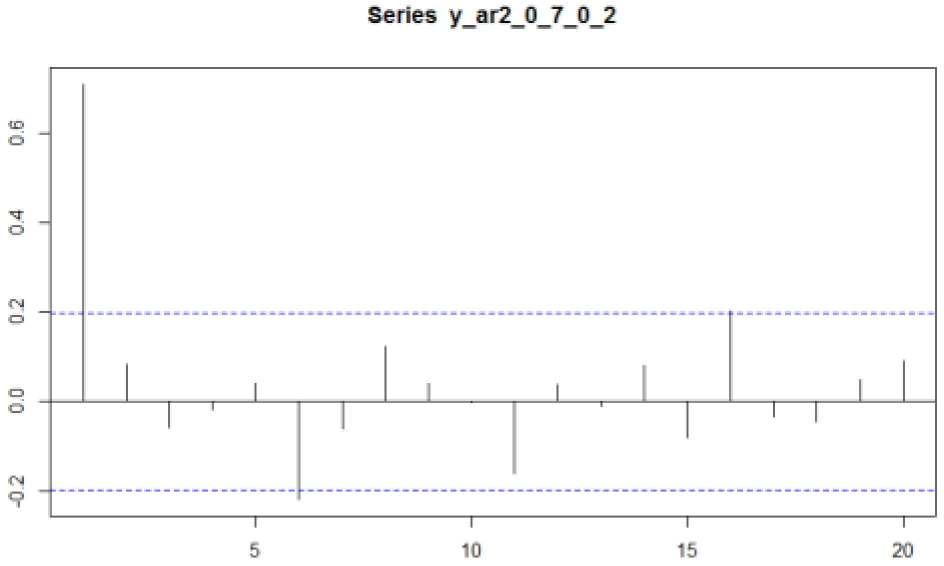
- Estimando as FACP
  - pacf(y_ar1_0_8)
  - pacf(y_ma1_0_3)
  - pacf(y_ma1_0_9)
  - pacf(y_ar2_0_7_0_2)

Resultado do FACP – modelos simulados









Função de Autocorrelação Parcial - FACP

- A FAC define a defasagem do MA(q). A FACP define defasagem do AR(p).
- No primeiro caso, a FAC decai com o aumento de defasagens, e a função de autocorrelação parcial é truncada a partir da defasagem p.
- No segundo caso, a função de autocorrelação é truncada na defasagem q, e a função de autocorrelação parcial decai.
- No caso de uma ARMA (p, q), ambas as funções decaem a partir da defasagem de truncagem.

Modelo	FAC	FACP
$AR\ (p)$	Decai	Truncada na defasagem p
$MA\ (q)$	Truncada na defasagem q	Decai
$ARMA\ (p, q)$	Decai se $j > q$	Decai se $j > p$

Teste de Ljung-Box

- Hipóteses:

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_n = 0 \quad \text{para } n = 1, 2, \ldots, J$$

$$H_1 : \exists\ \rho_j \neq 0 \text{ para algum } j = 1, 2, \ldots, n$$

- Estatística:

$$Q = T\,(T + 2) \sum_{j=1}^{n} \frac{\widehat{\rho}_j^{\,2}}{T - j} \xrightarrow{d} \chi_n^2$$

- Se uma das autocorrelações for diferente de zero, há evidência de um modelo ARMA(p,q).
- Se após estimado o modelo, os resíduos não sugerirem alguma autocorrelação, o modelo estará bem estimado.
  - Resíduo é um Ruído Branco.
- Na prática: Identifica-se o modelo por meio da FAC e FACP; depois usa-se a estatística de Ljung-Box sobre os resíduos estimados.

Teste de Ljung-Box

- Instalando o pacote
  - install.packages("stats")
- require(stats)
- Escrevendo o texto:
  - Para defasagem 1:

    Box.test(y,lag=1,type = "Ljung-Box")

```
        Box-Ljung test

data:  y
X-squared = 51.574, df = 1, p-value = 6.893e-13
```

  - Para a defasagem 5:

    Box.test(y,lag=5,type = "Ljung-Box")

```
        Box-Ljung test

data:  y
X-squared = 109.03, df = 5, p-value < 2.2e-16
```

FAC e FACP – solução única?

- As defasagens obtidas da FAC e FACP muitas vezes não são claras;
  - Difícil de identificar visualmente.
- É possível estimar mais de um modelo “correto”:
  - Resíduo do modelo é um ruído branco
- Como diferenciar?
  - Usar critérios de informação

Estimação e inferência

- A estimação se inicia com o método de máxima verossimilhança condicional.

- É um método que é assintoticamente equivalente ao da máxima verossimilhança exata.

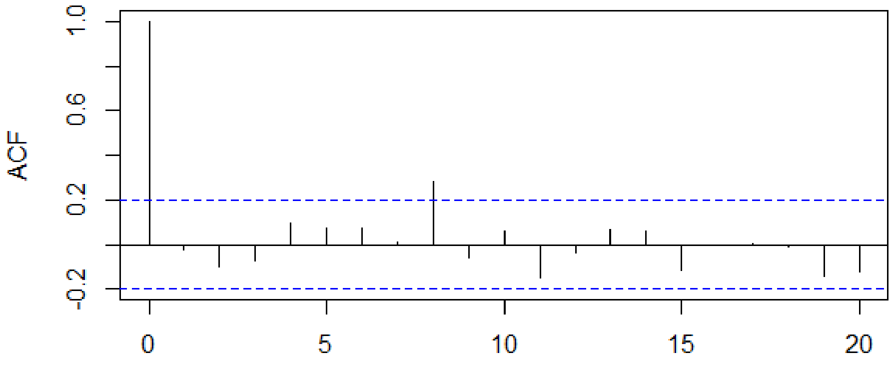
Estimação e inferência – Exemplo AR(2)

- Baixando o pacote para estimação
  - install.packages("FitARMA")
  - require(FitARMA)
- Estimando a equação
  - regarma1<-FitARMA(y_ar2_0_7_0_2,order = c(2,0,0),MeanMLEQ = TRUE)
- Observando os coeficientes estimados:
  - coef(regarma1)

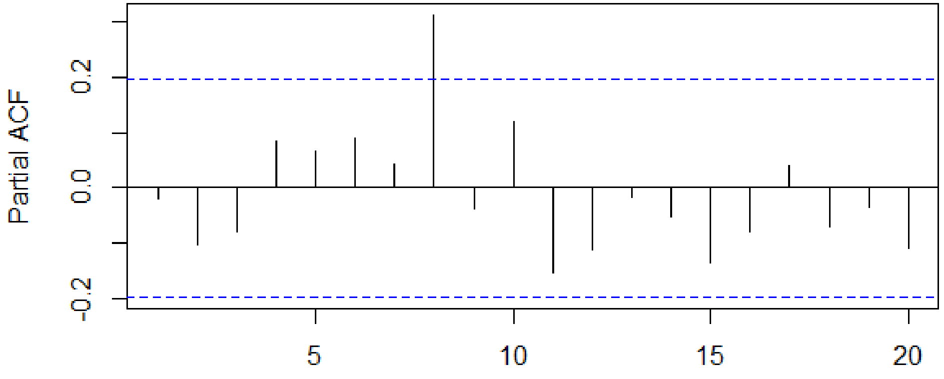
	MLE	sd	Z-ratio
phi(1)	0.5995806	0.0979807	6.11937488
phi(2)	0.1999457	0.0979807	2.04066413
mu	-1.3029027	54.9187532	-0.02372419

Estimação e inferência – Exemplo AR(2)

- erro1<-resid(regarma1)
- acf(erro1)



pacf(erro1)



- Box.test(erro1,lag=3,type = "Ljung-Box")

```
        Box-Ljung test

data:  erro1
X-squared = 1.6921, df = 3, p-value = 0.6387
```

Box.test(erro1,lag=3,type = "Ljung-Box")

```
        Box-Ljung test

data:  erro1
X-squared = 13.366, df = 10, p-value = 0.2039
```

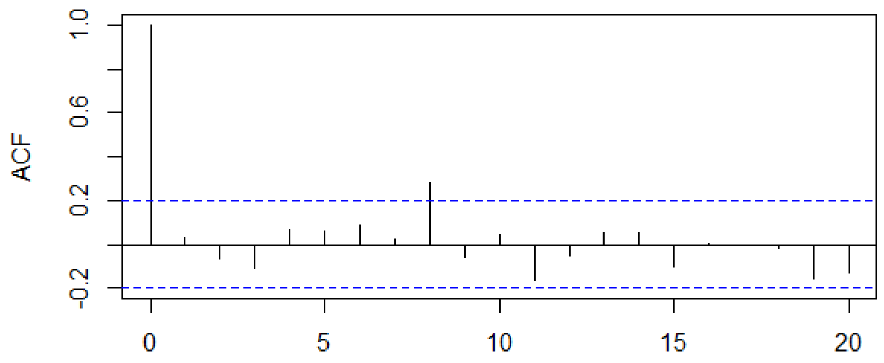
Estimação e inferência – Exemplo ARMA(1,1)

- Estimando a equação
  - regarma2<-FitARMA(y_ar2_0_7_0_2,order = c(1,0,1),MeanMLEQ = TRUE)
- Observando os coeficientes estimados:
  - coef(regarma2)

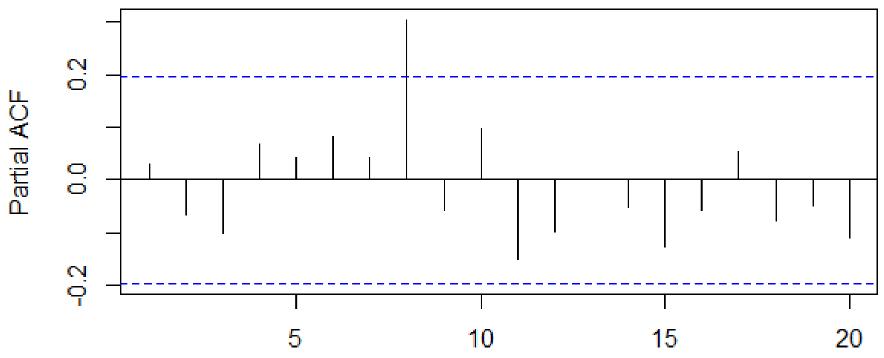
```
                MLE           sd      Z-ratio
phi(1)    0.8930677   0.05734697 15.573058334
theta(1)  0.3624920   0.11879076  3.051517002
mu       -1.2894904 673.32939719 -0.001915096
```

Estimação e inferência – Exemplo ARMA(1,1)

- erro2<-resid(regarma2)
- acf(erro2)



pacf(erro2)



- Box.test(erro2,lag=3,type = "Ljung-Box")

```
        Box-Ljung test

data:  erro2
X-squared = 1.7251, df = 3, p-value = 0.6314
```

Box.test(erro2,lag=10,type = "Ljung-Box")

```
        Box-Ljung test

data:  erro2
X-squared = 12.653, df = 10, p-value = 0.2437
```

Estimação e inferência – Ljung-Box

- Ljung-box – outra forma de obter o resultado
- regarma1$LjungBoxQ   regarma2$LjungBoxQ

AR(2)

m	Qm	pvalue
1	0.04	0.83545072
2	1.12	0.29050079
3	1.69	0.19332334
4	2.67	0.26343070
5	3.29	0.34856619
6	3.86	0.42518886
7	3.87	0.56841120
8	12.57	0.05043674
9	12.98	0.07249434
10	13.37	0.09986736
11	15.93	0.06843426
12	16.11	0.09654949
13	16.67	0.11797004
14	17.06	0.14747951
15	18.62	0.13526711

ARMA(1,1)

m	Qm	pvalue
1	0.09	0.76608823
2	0.55	0.45900059
3	1.73	0.18903746
4	2.16	0.33875038
5	2.53	0.46950107
6	3.32	0.50580931
7	3.39	0.64077305
8	12.06	0.06069784
9	12.41	0.08779662
10	12.65	0.12436572
11	15.71	0.07320355
12	15.99	0.10003036
13	16.30	0.13038049
14	16.65	0.16317891
15	17.90	0.16130392

Estimação e inferência

- Exemplo interessante, aponta que há significância na FAC e FACP na oitava defasagem, mas é espúria!
- Olhar o sentido econômico. Faz sentido? Há sazonalidade no 8º mês ou trimestre?
  - Normalmente olhamos se há significância nas primeiras autocorrelações e para períodos semestrais ou anuais.
- Em modelos simulados, é possível ter divergência de diagnóstico, imagine em variáveis reais (amostra aleatória);
- Como decidir?
- Critério de informação...

Critério de informação

- Método alternativo de identificar um ARMA(p,q).
  - Além disso, é comum dois ou mais modelos possíveis gerando resíduos cujos testes indiquem ser um ruído branco.
  - O melhor modelo será o mais parcimonioso. Por quê? O modelo com menor número de parâmetros deverá gerar menos imprecisão de estimativas.
- Ideia: balancear a redução dos erros e o aumento do número de regressores.
  - Normalmente, o aumento de regressores reduz a soma dos quadrados dos resíduos.
- O critério de informação associa uma penalidade ao aumento do número de regressores. Se a penalidade for menor do que a diminuição da soma de resíduos, o regressor adicional deve ser incorporado ao modelo.

Critério de informação

- A especificação geral:

$$C = \ln \hat{\sigma}^2(T) + c_T \varphi(T)$$

- $\hat{\sigma}^2(T) = T^{-1}\sum_{t=1}^{T} \hat{\varepsilon}_t^2$ é a variância estimada dos resíduos;
- c_T é o número de parâmetros estimados, $\varphi(T)$ é a função de penalização;
- O número total de observações T é invariante.
  - É necessário comparar séries com o mesmo número de observações.
- **OBJETIVO:** minimizar a **função C**

Critério de informação

- n=p+q ou n=p+q+1 se houver constante. T é o total de observações.
- Estatística de Schwarz ou BIC- Baysian Information Criterion;

$$BIC(p,q) = \ln \hat{\sigma}^2 + n\frac{\ln T}{T}$$

- Estatística de Akaike ou AIC – Akaike Information Criterion;

$$AIC(p,q) = \ln \hat{\sigma}^2 + n\frac{2}{T}$$

- Estatística de Hanna-Quinn, HQ;

$$HQ(p,q) = \ln \hat{\sigma}^2 + n\frac{2}{T}\ln \ln T$$

- BIC é consistente assintoticamente e o AIC para pequenas amostras;
- De forma geral, T>16

$$\hat{p}_{BIC} \leq \hat{p}_{HQ} \leq \hat{p}_{AIC}$$

Critério de informação

- summary(regarma1)

```
ARIMA(2,0,0)  with mean MLE.
length of series = 100 ,  number of parameters = 3
loglikelihood = 11.18 ,  aic = -16.4 ,  bic =  -8.5
```

- regarma2 <enter>

```
ARIMA(1,0,1)  with mean MLE.
length of series = 100 ,  number of parameters = 3
loglikelihood = 11.95 ,  aic = -17.9 ,  bic =  -10.1
```

Revisão: processo de uma série ARMA(p,q)

1. Identificar as ordens p e q do modelo.
2. Estimá-lo.
3. Verificar se os resíduos estimados são um ruído branco.
   - Sim, próximo passo;
   - Não, volte ao passo 1.
4. Projetar!

- Se uma série for considerada não estacionária, deve ser diferenciada.

Testes dos resíduos

- Erro de especificação
  - Teste Reset
- Teste para heterocedasticidade
  - LM ou Breusch-Pagan-Godfrey
- Autocorrelação dos resíduos
  - Teste Breusch-Godfrey
- Teste de normalidade
  - Jarque-Bera – verifica se os momentos da série estimada são semelhantes aos da normal (assimetria e curtose);

$$\begin{aligned} H_0 &: E\left(\varepsilon_t^s\right)^3 = 0 \wedge E\left(\varepsilon_t^s\right)^4 = 3 \\ &\times \\ H_1 &: E\left(\varepsilon_t^s\right)^3 \neq 0 \vee E\left(\varepsilon_t^s\right)^4 \neq 3 \end{aligned}$$

  - Teorema do Limite Central - Quando o tamanho amostral é suficientemente grande, a distribuição da média é uma distribuição aproximadamente normal.

Testes dos resíduos - normalidade

- Instalando o pacote
  - install.packages("tseries");
  - require(tseries);
- Resultado do teste de normalidade
  - jarque.bera.test(erro2)

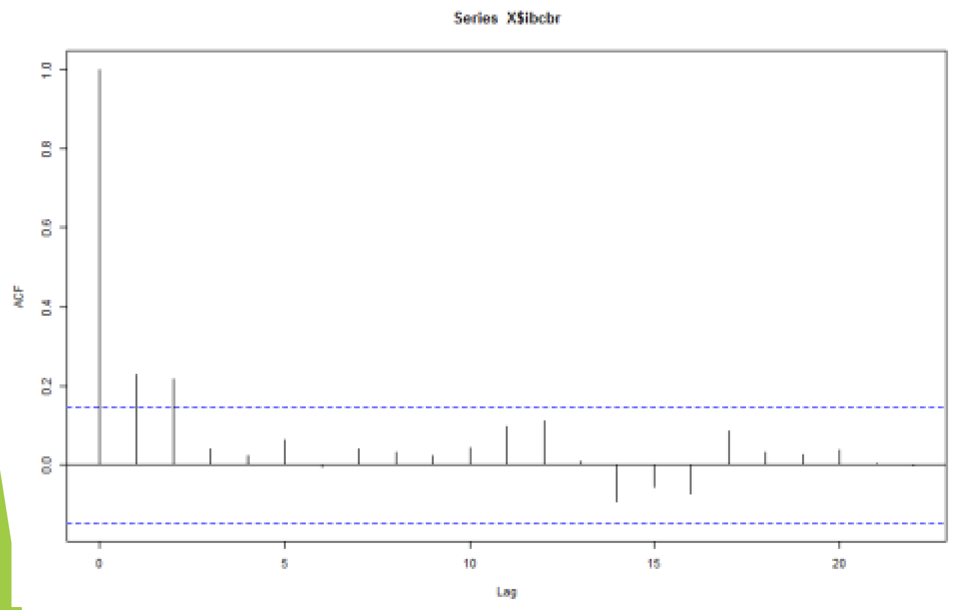
```
        Jarque Bera Test

data:  x1
X-squared = 0.82314, df = 2, p-value = 0.6626
```

Exemplo real – IBC-br

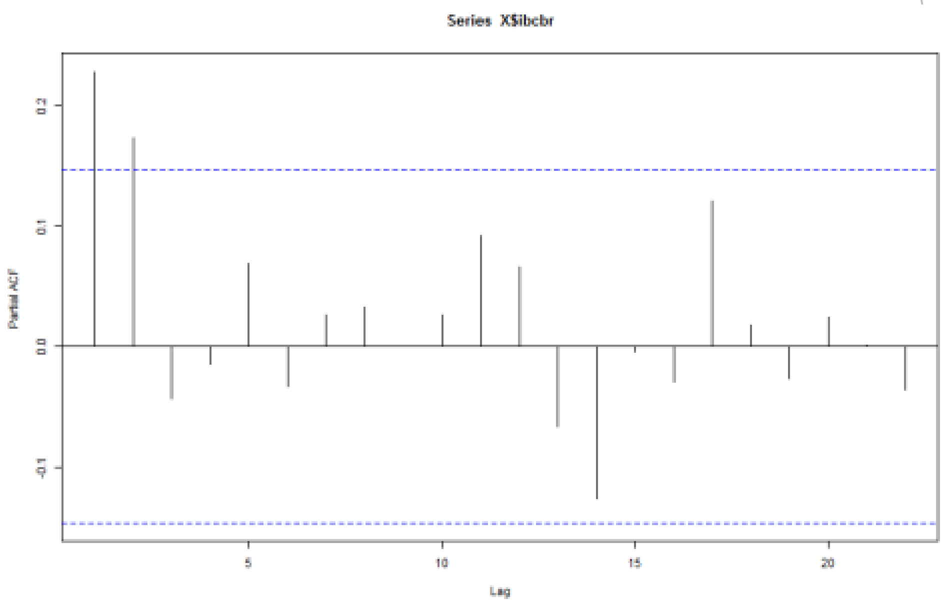
- Baixando os dados.
- setwd("C:/diretorio/diretorio")
- list.files()
- X<-read.csv("aula2_dados.csv",sep=";", dec=".", head=TRUE)
- X$data<-as.Date(X$data,'%d/%b/%y')
- head(X)

Exemplo real – IBC-br

acf(X$ibcbr)



pacf(X$ibcbr)



- Que modelo sugere?

Exemplo real – IBC-br

- Possíveis modelos:
  - ARMA(2,2)
  - AR(2)
  - MA(2)

Exemplo real – IBC-br – ARMA(2,2)

- Estimando o modelo
  - ibcbr_arma22<-FitARMA(X$ibcbr,order = c(2,0,2),MeanMLEQ = TRUE)a
  - coef(ibcbr_arma22)

	MLE	sd	Z-ratio
phi(1)	0.02791488	0.40722151	0.06854961
phi(2)	0.02681891	0.33437672	0.08020567
theta(1)	-0.17575284	0.40166412	-0.43756171
theta(2)	-0.20180264	0.28398058	-0.71062126
mu	0.19182657	0.07083413	2.70810931

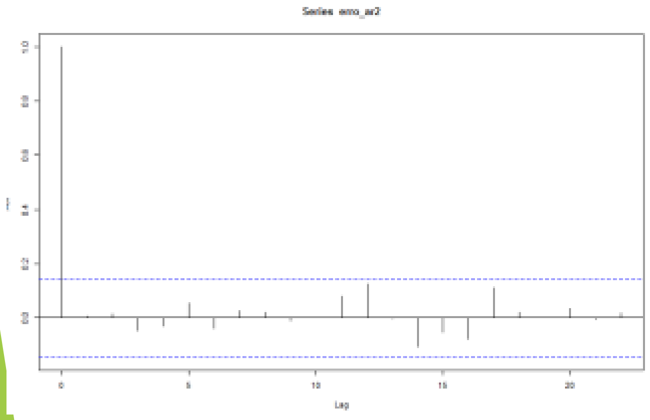
Exemplo real – IBC-br – AR(2)

- Estimando o modelo
  - ibcbr_ar2<-FitARMA(X$ibcbr,order = c(2,0,0),MeanMLEQ = TRUE)
  - coef(ibcbr_ar2)

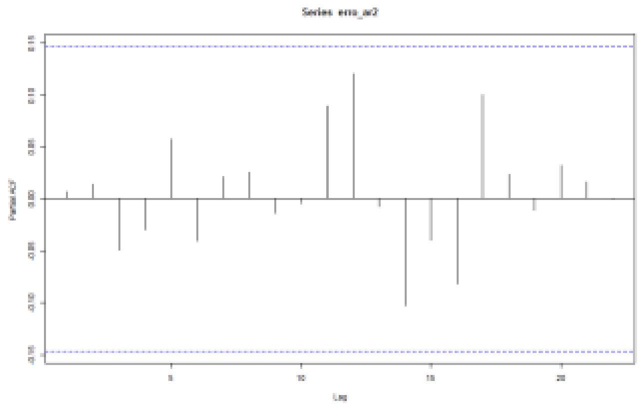
	MLE	sd	Z-ratio
phi(1)	0.1916543	0.07358821	2.6044153
phi(2)	0.1751424	0.07358821	2.3800335
mu	0.1926979	0.35278910	0.5462127

  - erro_ar2<-resid(ibcbr_ar2)

acf(erro_ar2)



pacf(erro_ar2)



ibcbr_ar2$LjungBoxQ

m	Qm	pvalue
1	0.01	0.9305878
2	0.04	0.8363149
3	0.48	0.4872894
4	0.65	0.7236980
5	1.22	0.7484525
6	1.49	0.8283449
7	1.61	0.8999974
8	1.69	0.9461123
9	1.71	0.9739582
10	1.71	0.9885560
11	2.95	0.9664136
12	5.95	0.8191315
13	5.95	0.8764191
14	8.21	0.7682411
15	8.80	0.7875653

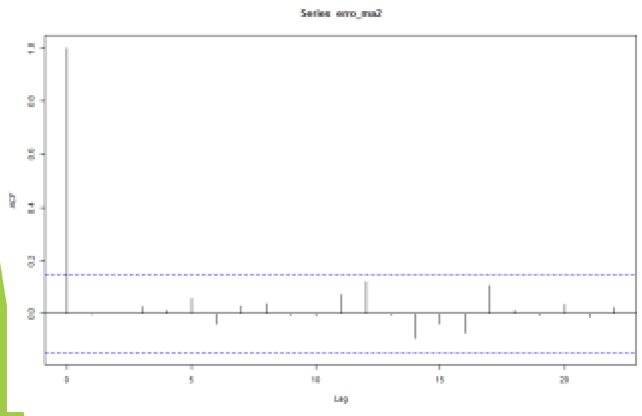
Exemplo real – IBC-br – MA(2)

- Estimando o modelo
  - ibcbr_ma2<-FitARMA(X$ibcbr,order = c(0,0,2),MeanMLEQ = TRUE)
  - coef(ibcbr_ma2)

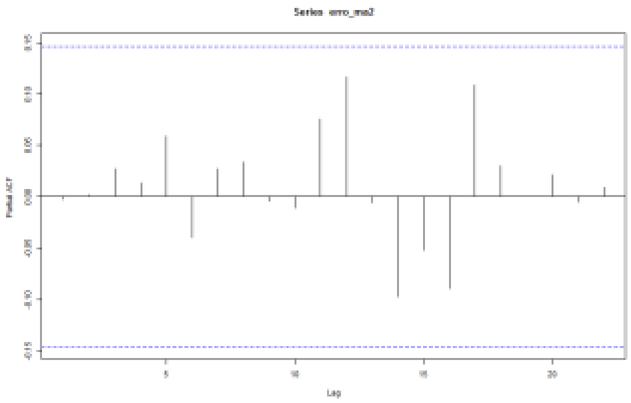
```
                 MLE          sd   Z-ratio
theta(1) -0.2026664 0.07270271 -2.787605
theta(2) -0.2320831 0.07270271 -3.192221
mu        0.1915622 0.05655121  3.387411
```

  - erro_ma2<-resid(ibcbr_ma2)

acf(erro_ma2)



pacf(erro_ma2)



ibcbr_ma2$LjungBoxQ

m	Qm	pvalue
1	0.00	0.9683351
2	0.00	0.9647442
3	0.14	0.7086904
4	0.17	0.9173560
5	0.81	0.8462653
6	1.10	0.8935265
7	1.25	0.9396079
8	1.51	0.9587208
9	1.52	0.9817585
10	1.52	0.9922921
11	2.56	0.9793594
12	5.35	0.8667979
13	5.35	0.9128684
14	7.06	0.8536997
15	7.38	0.8817300

Exemplo real – IBC-br – AR(2)

- Critério de informação
  - AR(2)

```
> ibcbr_ar2
ARIMA(2,0,0)  with mean MLE.
length of series = 179 ,  number of parameters = 3
loglikelihood = 49.26 ,  aic = -92.5 ,  bic =  -83
```

  - MA(2)

```
> ibcbr_ma2
ARMA(0,2)  with mean MLE.
length of series = 179 ,  number of parameters = 3
loglikelihood = 49.64 ,  aic = -93.3 ,  bic =  -83.7
```

Introdução a previsão

- Horizonte de previsão que se deseja prever, h: período de tempo entre hoje, t, e h-passos à frente, t + h.
- Distinção: **previsões** ***h*-passos** à frente e **extrapolação** ***h*-passos** à frente.
  - Previsões de um ponto de futuro com dados até t: t + h, t + 1 + h.
  - Extrapolação: previsões consecutivas em t + 1, t + 2, . . . , t + h.
- Previsão dinâmica: não usa a informação adicional para realimentar as previsões. Isto é, para prever y_{T+2}, usa-se $E_t\ (y_{T+1})$.
- Previsão estática: usa na previsão de $_e\ y_{T+h}$ a observação $y_{T+h-1.}$Também chamada de previsão 1-passao à frente.
- A previsão estática tende a ter um erro de previsão menor do que a a previsão dinâmica, haja vista que as informações são atualizadas a cada passo.
- A previsão dinâmica representa uma extrapolação H-passos a frente.

Previsão dinâmica: valor médio – AR(1)

$$y_{t+1} = c + \phi y_t + \varepsilon_{t+1}.$$

Logo:

$$\begin{aligned}
E_t\left(y_{t+1}\right) &= c + \phi y_t = y_{t+1} - \varepsilon_{t+1}; \\
E_t\left(y_{t+2}\right) &= c + \phi E_t\left(y_{t+1}\right) = c + \phi\left(c + \phi y_t\right); \\
&\vdots \\
E_t\left(y_{t+h}\right) &= c \sum_{i=1}^{h} \phi^{i-1} + \phi^h y_t.
\end{aligned}$$

Previsão dinâmica: Erro de previsão – AR(1)

O erro de previsão no período h, $e_t(h)$, é dado por:

$$
\begin{aligned}
e_t(1) &= y_{t+1} - E_t(y_{t+1}) = \varepsilon_{t+1}; \\
e_t(2) &= y_{t+2} - E_t(y_{t+2}) = c + \phi y_{t+1} + \varepsilon_{t+2} - c - \phi E_t(y_{t+1}) = \\
&= \phi \varepsilon_{t+1} + \varepsilon_{t+2}; \\
e_t(3) &= y_{t+3} - E_t(y_{t+3}) = c + \phi y_{t+2} + \varepsilon_{t+3} - c - \phi E_t(y_{t+2}) = \\
&= \varepsilon_{t+3} + \phi \varepsilon_{t+2} + \phi^2 \varepsilon_{t+1}; \\
&\vdots \\
e_t(h) &= y_{t+h} - E_t(y_{t+h}) = \\
&= \varepsilon_{t+h} + \phi \varepsilon_{t+h-1} + \phi^2 \varepsilon_{t+h-2} + \cdots + \phi^{h-1} \varepsilon_{t+1}.
\end{aligned}
$$

Tomando as esperanças dos erros de previsão, verifica-se que são iguais a zero.

Previsão dinâmica: Variância – AR(1)

$$
\begin{aligned}
Var\left(e_t\left(h\right)\right) &= Var\left(\varepsilon_{t+h} + \phi\varepsilon_{t+h-2} + \phi^2\varepsilon_{t+h-3} + \cdots + \phi^{h-1}\varepsilon_{t+1}\right) = \\
&= \sigma^2\left(1 + \phi^2 + \phi^4 + \cdots + \phi^{2(h-1)}\right).
\end{aligned}
$$

- A variância aumenta com o horizonte de previsão a taxas decrescentes. Quando h—> ∞ , a variância de previsão converge à variância não condicional $\frac{\sigma^2}{1-\phi^2}$

- O intervalo de conança para resíduos normais é dado da seguinte forma:

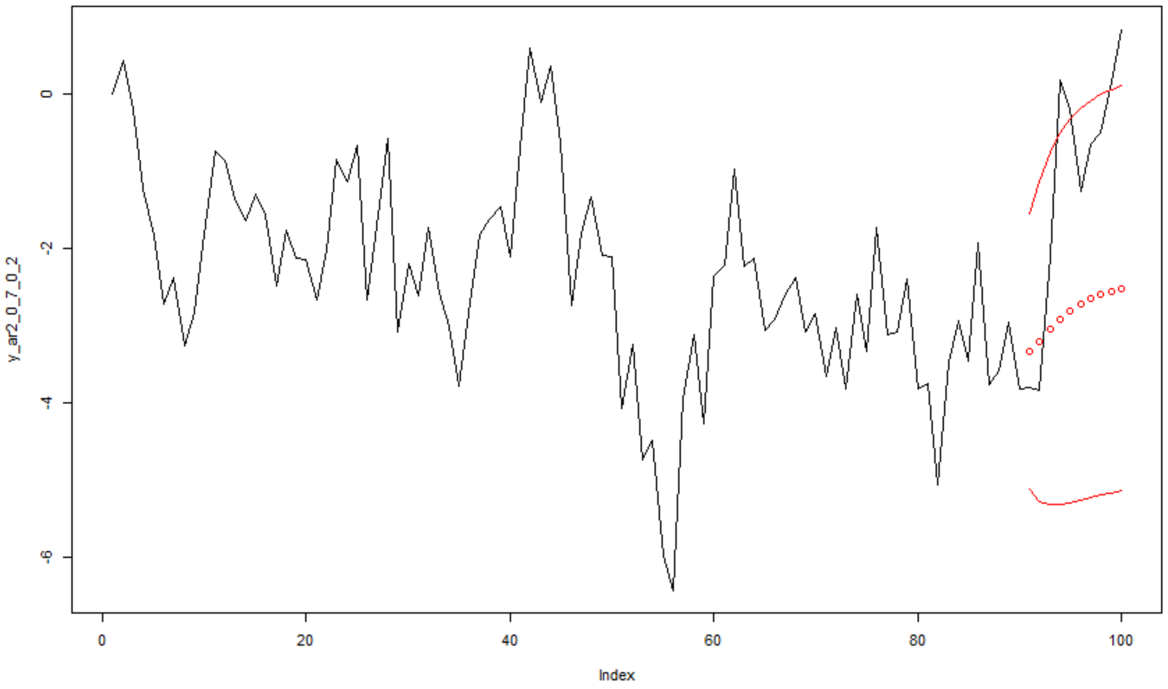
$$
c\sum_{i=1}^{h}\phi^{i-1} + \phi^h y_t \pm 2\sigma\left(1 + \phi^2 + \phi^4 + \cdots + \phi^{2(h-1)}\right)^{\frac{1}{2}}
$$

Previsão do modelo teórico

- Vamos selecionar as últimas 10 observações para projeção.
- Treinaremos o modelo com as 90 primeiras observações.
- Coeficientes estimados serão diferentes – número de observação é diferente.
- Projeção dinâmica!
- Regarma1 – ARMA(2,0)
  - previsao1<-predict(arima(y_ar2_0_7_0_2[1:90], order = c(2,0,0)), n.ahead = 10)
- Calculando o intervalo de projeção – baseado no desvio padrão.
  - Superior: previsao1$pred+previsao1$se*1.96
  - Inferior: previsao1$pred-previsao1$se*1.96

Previsão do modelo teórico - Gráfico

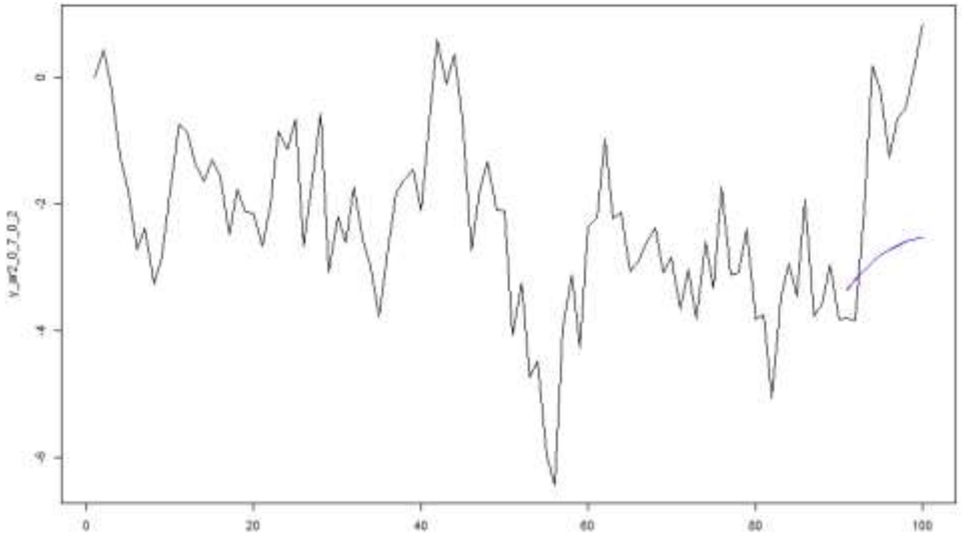
```
plot(y_ar2_0_7_0_2,type="l")
points(previsao1$pred,col="red")
points(previsao1$pred+previsao1$se*1.96,type="l",col="red")
points(previsao1$pred-previsao1$se*1.96,type="l",col="red")
```



Previsão do modelo teórico

- Regarma2 – ARMA(1,1)
  - previsao2<-predict(arima(y_ar2_0_7_0_2[1:90], order = c(1,0,1)), n.ahead = 10)
- Gráfico:

```
plot(y_ar2_0_7_0_2,type="l")

points(previsao1$pred,type="l",col="red")

points(previsao2$pred,type="l",col="blue")
```



Avaliação da Previsão

- Divide-se a amostra:
  1. Estimação do modelo
  2. Treinamento das projeções
  3. Comparação das projeções com o restante da amostra – *pseudo out-of-sample*.
- Se a amostra for suficientemente grande, pode-se deixar de 1/4 a 1/3 fora da amostra por razões de previsão.
- Métodos:

$$MSE_{t,H} = \sqrt{\frac{\sum_{h=1}^{H} e_t^2(h)}{H}};$$

$$MAE_{t,H} = \frac{\sum_{h=1}^{H} |e_t(h)|}{H};$$

$$MAPE_{t,H} = \sum_{h=1}^{H} \left| \frac{e_t(h)}{Hy_{t+h}} \right|.$$

Função de Autocorrelação Parcial - FACP

- MSE

```
MSE1<- (mean((y_ar2_0_7_0_2[91:100] - previsao1$pred )^2))^(0.5)

MSE2<- (mean((y_ar2_0_7_0_2[91:100] – previsao2$pred )^2))^(0.5)
```

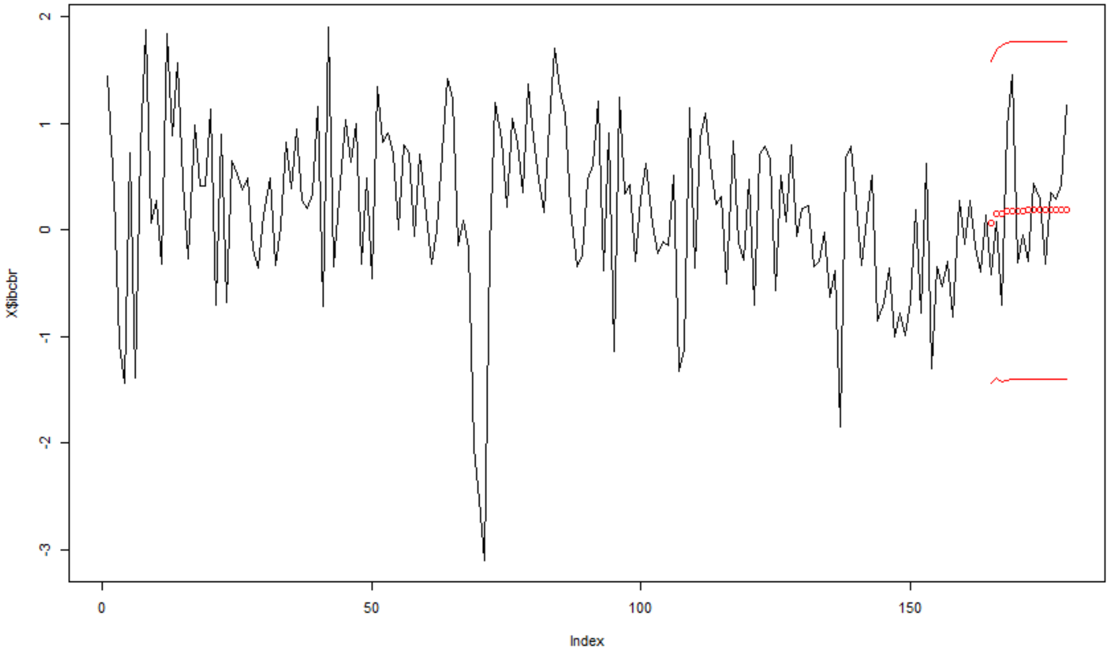
- MAE

```
MAE1<-mean(abs(y_ar2_0_7_0_2[91:100] - previsao1$pred ))

MAE2<-mean(abs(y_ar2_0_7_0_2[91:100] - previsao2$pred ))
```

Voltando ao exercício do IBC-br

- Projetando o modelo do AR(2)

```
previbcbr_ar2<-predict(arima(X$ibcbr[1:164], order = c(2,0,0)), n.ahead = 15)

plot(X$ibcbr,type="l")

points(previbcbr_ar2 $pred,col="red")

points(previbcbr_ar2$pred+ previbcbr_ar2$se*1.96,type="l",col="red")

points(previbcbr_ar2$pred- previbcbr_ar2$se*1.96,type="l",col="red")
```



Voltando ao exercício do IBC-br

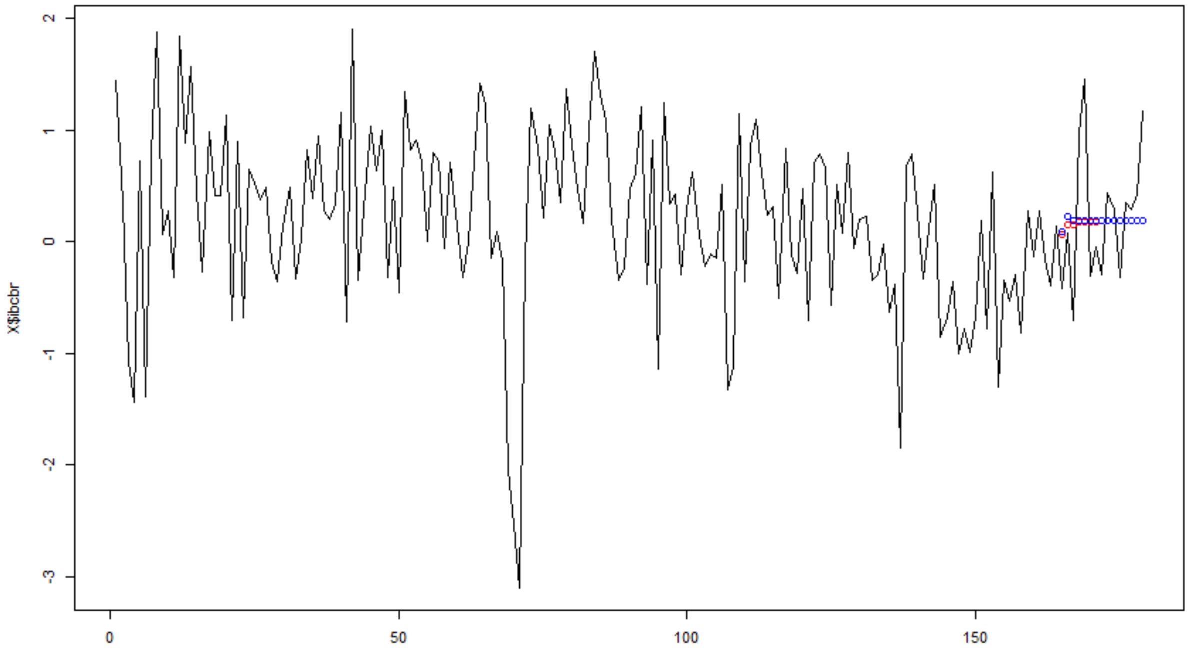
- Projetando o modelo do MA(2)

previbcbr_ma2<-predict(arima(X$ibcbr[1:164], order = c(0,0,2)), n.ahead = 15)

- plot(X$ibcbr,type="l") #gráfico

points(previbcbr_ar2$pred,col="red")

points(previbcbr_ma2$pred,col="blue")



Voltando ao exercício do IBC-br

- MSE

```
MSE1ibcbr<- (mean((X$ibcbr[165:179] - previbcbr_ar2$pred )^2))^(0.5)

MSE2ibcbr<- (mean((X$ibcbr[165:179] - previbcbr_ma2$pred )^2))^(0.5)
```

- Conclusão:
  - Os modelos AR(2) e MA(2) tiveram um bom ajuste no IBC-br
  - Segundo o AIC e BIC, o modelo MA(2) minimizou os critérios
  - No entanto, para a projeção, o modelo AR(2) teve o menor erro quadrático médio.

Exercício em aula

```
> head(x)
        data    varejo varejo_ampl      ibcbr      ipca
1 01/fev/03  2.364066   -3.611738  1.4439355 1.4918278
2 01/mar/03 -9.006928    0.234192  0.3533916 1.2039650
3 01/abr/03  4.060914    2.102804 -1.0760051 0.9450566
4 01/mai/03 -1.707317    4.347826 -1.4337981 0.9637544
5 01/jun/03 -3.722084   -5.701754  0.7223114 0.2512400
6 01/jul/03  6.958763    7.441860 -1.3844622 0.2701869
```

- Escolher alguma dessas variáveis e estimar pelo menos 2 modelos ARMA;
  - Baseando-se na FAC e FACP
- Apresentar o teste Ljung-Box e funções de autocorrelação do resíduo.
- Exibir o melhor modelo baseado no critério de informação.
- Projetar os dois modelos e mostrar qual tem o menor erro quadrático médio.