Econometria Financeira

MODELOS ARMA

Profa Dra Andreza Palma

UFSCar - Campus Sorocaba

Plano de vôo

- Definição de série temporal
- Estacionariedade
- Modelos ARMA
    - Modelos Autorregressivos
    - Modelos de Médias Móveis
    - Modelos mistos: ARMA
- Identificação, Estimação e Diagnóstico

Série temporal: definição

- Uma **série temporal** consiste em uma sequência de observações de uma variável de interesse.
- As observações vizinhas são dependentes e o nosso interesse é modelar essa dependência.
- Portanto, a ordem das variáveis é importante para análise, ou seja, o ordenamento importa (ao contrário da análise usual de dados cross-section).
- Os modelos para descrever séries temporais são processos estocásticos. Portanto, assume-se que a série temporal é uma realização de um processo estocástico subjacente.
- Um processo estocástico é uma coleção de variáveis aleatórias, ou seja, para cada t, $X(t)$ é uma variável aleatória.

Série temporal: objetivos

- Compreender o comportamento observado da série.
- O que podemos esperar para o futuro?
- Problema fundamental: descobrir o processo gerador dos dados (que é desconhecido)
- O trabalho do econometrista é, portanto, tentar inferir o o processo gerador dos dados a partir do que é observado.

Estacionariedade

O conceito de estacionariedade é fundamental em análise de séries temporais, já que a escolha de um modelo requer uma inferência sobre a distribuição de probabilidade conjunta de uma série:

$$F\left(x_1, x_2, \ldots, x_n\right) = \Pr\left(X_1 \leq x_1, \ldots, X_n \leq x_n\right) \mid$$

Estacionariedade estrita implica que a distribuição da série é invariante a mudanças do tempo.

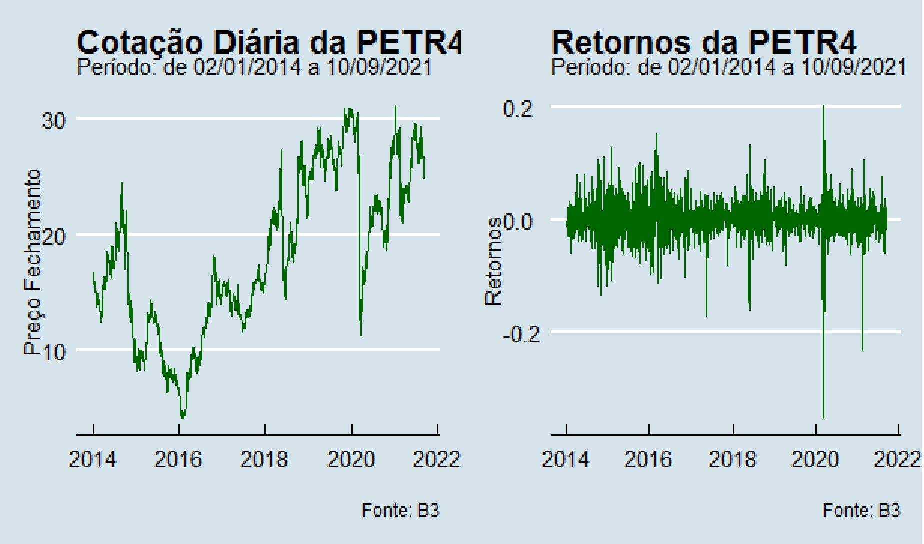
Estacionariedade fraca ou estacionariedade em covariância: possui média constante e a covariância depende apenas da distância entre x_t e x_{t-i}.

Estacionariedade

Em termos gráficos, uma série **estacionária** terá seus valores oscilando de forma constante em torno de um valor fixo.

Já para uma série **não estacionária**, não há uma média de longo prazo para a qual a série retorna.

Estacionariedade



Estacionariedade

Estacionariedade pode ser verificada a partir de testes estatísticos (i.e. testes de raiz unitária, tais como ADF, Phillips-Perron, entre outros) - muito embora esses testes apresentem problemas.

Como em séries temporais financeiras estaremos sempre trabalhando com os retornos, a estacionariedade não será uma preocupação, já que os retornos são em geral estacionários.

Modelos ARMA(p,q)

Um modelo ARMA (p,q) pode ser descrito pela seguinte equação:

$$y_t = c + \sum_{j=1}^{p} \phi_j y_{t-j} + \sum_{i=0}^{q} \theta_i \epsilon_{t-i}$$

Um ARMA(1,1) será dado, então, por:

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

Processo Auto-regressivo - AR(p)

- Uma característica importante de muitas séries temporais é que seu valor no período atual depende fortemente do valor no período imediatamente anterior, o que reflete o fato que as observações não são independentes, mas autocorrelacionadas.
- Dito de outra forma, a série apresenta algum grau de **persistência**.
- Esse caso pode ser descrito por um processo auto-regressivo de ordem 1, por exemplo, dado por:

$$y_t = \phi y_{t-1} + \epsilon_t$$

- Ou seja, o valor de y está relacionado ao seu valor em $t-1$ mais um erro aleatório, que assumimos ser um ruído branco ,

$$\epsilon_t \sim \text{ i. i. d. } \mathcal{N}\left(0, \sigma^2\right).$$

Nota: Ruído Branco

- Uma série temporal ϵ_t é considerada um ruído branco se $\{\epsilon_t\}$ tem média 0 e variância constante σ^2, ou seja:

$$E(\epsilon_t) = 0 \quad E\left(\epsilon_t^2\right) = \sigma^2$$

onde ϵ_t é não correlacionado temporalmente, $E(\epsilon_t \epsilon_{t-l}) = 0$ para $l \neq 0$

Média, variância e autocovariância de um processo AR(1)

Seja o processo AR(1):

$$y_t = \phi y_{t-1} + \epsilon_t$$

Média: $E(y_t)$

$$E(y_t) = E(\phi y_{t-1} + \epsilon_t)$$

$$E(y_t) = \phi E(y_{t-1}) + E(\epsilon_t))$$

$$\mu = \phi \mu + 0$$

$$\mu = 0$$

Média, variância e autocovariância de um processo AR(1)

Variância: $var(y_t)$

$$var(y_t) = var(\phi y_{t-1} + \epsilon_t)$$

$$var(y_t) = \phi^2 var(y_t + var(\epsilon_t))$$

$$var(y_t) - \phi^2 var(y_t = \sigma^2$$

$$(1 - \phi^2) var(y_t) = \sigma^2$$

$$var(y_t) = \frac{\sigma^2}{(1 - \phi^2)}$$

Média, variância e autocovariância de um processo AR(1)

Autocovariância: $\gamma_k = cov(y_t, y_{t-k})$
Nota: lembrando que a covariância é a média dos produtos menos o produto da média.
Vamos começar calculando a autocovariância de ordem 1:

$$\gamma_1 = cov(y_t, y_{t-1}) = E(y_t y_{t-1}) - E(y_t)E(y_{t-1})$$

$$\gamma_1 = E[(\phi y_{t-1} + \epsilon_t)y_{t-1}] - 0$$

$$\gamma_1 = E[(\phi y_{t-1}^2 + \epsilon_t y_{t-1}]$$

$$\gamma_1 = \phi E[y_{t-1}^2] + E[\epsilon_t y_{t-1}]$$

$$\gamma_1 = \phi var(y_t)$$

$$\gamma_1 = \phi \gamma_0$$

Nota: $var(y_t) = \gamma_0 = cov(y_t, y_t)$

Média, variância e autocovariância de um processo AR(1)

Agora, a autocovariância de ordem 2:

$$\gamma_2 = cov(y_t, y_{t-2}) = E(y_t y_{t-2}) - E(y_t)E(y_{t-2})$$

$$\gamma_2 = E[(\phi y_{t-1} + \epsilon_t)y_{t-2}] - 0$$

Lembrando que $y_{t-1} = \phi y_{t-2} + \epsilon_{t-1}$:

$$\gamma_2 = E[(\phi(\phi y_{t-2} + \epsilon_{t-1}) + \epsilon_t)y_{t-2}]$$

$$\gamma_2 = E[(\phi^2 y_{t-2}^2) + \phi\epsilon_{t-1}y_{t-2} + \phi\epsilon_t y_{t-2})]$$

$$\gamma_2 = \phi^2 var(y_t)$$

$$\gamma_2 = \phi^2 \gamma_0$$

Média, variância e autocovariância de um processo AR(1)

Em resumo, a autocovariância de ordem k para um processo AR(1) será dada por:

$$\gamma_k = \phi^k \gamma_0$$

Função de autocorrelação (FAC) de um processo AR(1)

A FAC é dada por:

$$\rho_k = corr(y_t, y_{t-k})$$

$$\rho_k = \frac{cov(y_t, y_{t-k})}{var(y_t)}$$

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

Portanto, para um AR(1) com $\phi < 0$, a FAC será declinante:

$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\phi\gamma_0}{\gamma 0} = \phi$

$\rho_2 = \phi^2$

$\rho_3 = \phi^3$

...

$\rho_k = \phi^k$

Função de autocorrelação Parcial (FACP) de um processo AR(1)

A FACP determina a correlação entre duas variáveis eliminando o efeito de outras variáveis, o que pode ser feito através de uma regressão linear. Portanto, a função de correlação parcial será dada pelos coeficientes $\varphi_1, \varphi_2, ..., \varphi_k$ das regressões a seguir:

- φ_1: $y_t = \varphi_1 y_{t-1} + v_t$
- φ_2: $y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + v_t$
- ...
- φ_k: $y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + ... + \varphi_k y_{t-k} + v_t$

Função de autocorrelação Parcial (FACP) de um processo AR(1)

- É imediato notar que para um processo AR(1) o único coeficiente que será significativo em uma regressão será φ_1. Portanto, a FACP será **truncada** na defasagem (lag) 1.

- De forma geral, a FACP de um AR(p) terá $\varphi_k = 0$ para $k > p$ e $\varphi_k \neq 0$ para $k \leq p$

FAC e FACP de um processo AR(1)

FAC e FACP para um AR(1) com $\phi = 0.9$

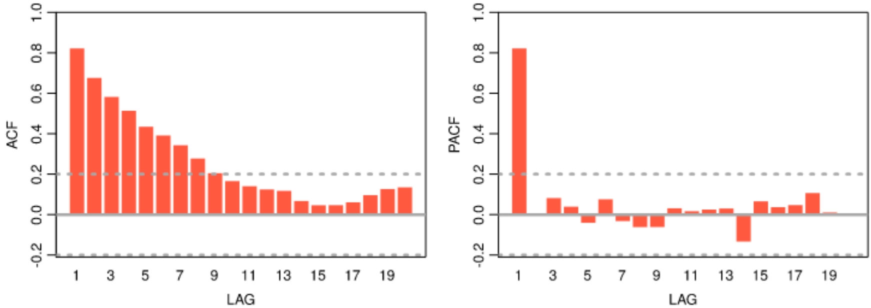


Figure: Fonte: Kotzé, K. (2019)

FAC e FACP de um processo AR(1)

FAC e FACP para um AR(1) com $\phi = -0.8$

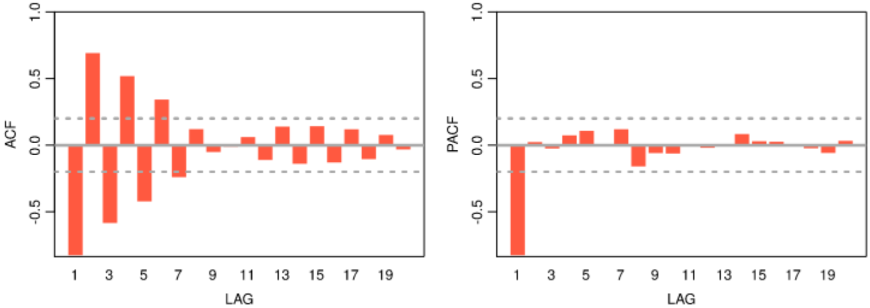


Figure: Fonte: Kotzé, K. (2019)

Exemplo de processo AR(1)

AR(1) com $\phi = -0.8$

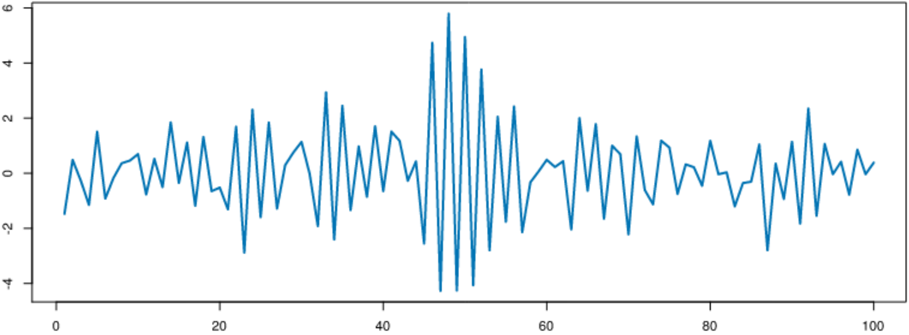


Figure: Fonte: Kotzé, K. (2019)

Operador Defasagem (*Lag*)

- $Ly_t = y_{t-1}$
- $L^2 y_t = LLy_y = L(Ly_t) = L(y_{t-1}) = y_{t-2}$
- $L^n = y_{t-n}$

Operador Defasagem (*Lag*)

Usando o operador defasagem, podemos escrever o AR(1) de forma alternativa:

$$y_t = \phi y_{t-1} + \epsilon_t$$
$$y_t - \phi y_{t-1} = \epsilon_t$$

$$y_t - \phi L y_t = \epsilon_t$$

$$(1 - \phi L) y_t = \epsilon_t$$

Temos um polinômio na variável L que pode ser chamado apenas de $\Theta_p(L)$ e, portanto:

$$\Theta_p(L) y_t = \epsilon_t$$

Operador Defasagem (*Lag*)

$$\Theta_p(L)y_t = \epsilon_t$$

Para o caso de um AR(1):

$\Theta_p(L) = (1 - \phi_1 L)$

Para o caso de um AR(p):

$\Theta_p(L) = (1 - \phi_1 L - \phi_2 L^2 - ...\phi_p L^p)$

Modelos de Média Móvel - MA (q)

O modelo de média móvel consiste em uma combinação linear dos erros ruído branco. O MA(1) pode ser escrito da seguinte forma:

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

onde μ é uma constante, y_t é a soma ponderada de uma constante e os dois valores mais recentes do termo de erro, ε_t e ε_{t-1} e
$\varepsilon_t \sim$ i. i. d. $\mathcal{N}\left(0, \sigma^2\right)$

Média, variância e autocovariância de um MA (q)

Média:

$$\begin{aligned}\mathbb{E}\left[y_t\right] &= \mathbb{E}\left[\mu + \varepsilon_t + \theta\varepsilon_{t-1}\right] \\ &= \mu + \mathbb{E}\left[\varepsilon_t\right] + \theta\mathbb{E}\left[\varepsilon_{t-1}\right] \\ &= \mu\end{aligned}$$

Média, variância e autocovariância de um MA (q)

Variância:

$$
\begin{aligned}
\text{var}\left[y_t\right] &= \mathbb{E}\left[y_t - \mathbb{E}\left[y_t\right]\right]^2 \\
&= \mathbb{E}\left[(\mu + \varepsilon_t + \theta\varepsilon_{t-1}) - \mu\right]^2 \\
&= \mathbb{E}\left[\varepsilon_t\right]^2 + 2\theta\mathbb{E}\left[\varepsilon_t\varepsilon_{t-1}\right] + \mathbb{E}\left[\theta\varepsilon_{t-1}\right]^2 \\
&= \sigma^2 + 0 + \theta\sigma^2 \\
&= \left(1 + \theta^2\right)\sigma^2
\end{aligned}
$$

Lembrando que:

$$\mathbb{E}\left[\varepsilon_t\right] = 0 \text{ and } \mathbb{E}\left[\left(\varepsilon_t\right)^2\right] = \sigma^2$$

Média, variância e autocovariância de um MA (q)

Autocovariância:

Vamos começar com a autocovariância de ordem 1:

$$
\begin{aligned}
\text{cov}\left[y_t, y_{t-1}\right] &= \mathbb{E}\left[\left(y_t - \mathbb{E}\left[y_t\right]\right)\left(y_{t-1} - \mathbb{E}\left[y_{t-1}\right]\right)\right] \\
&= \mathbb{E}\left[\left(\varepsilon_t + \theta\varepsilon_{t-1}\right)\left(\varepsilon_{t-1} + \theta\varepsilon_{t-2}\right)\right] \\
&= \mathbb{E}\left[\varepsilon_t\varepsilon_{t-1}\right] + \theta\mathbb{E}\left[\varepsilon_{t-1}^2\right] + \mathbb{E}\left[\theta\varepsilon_t\varepsilon_{t-2}\right] + \mathbb{E}\left[\theta^2\varepsilon_{t-1}\varepsilon_{t-2}\right] \\
&= 0 + \theta\sigma^2 + 0 + 0 \\
&= \theta\sigma^2
\end{aligned}
$$

Média, variância e autocovariância de um MA (q)

Autocovariância:

E agora de ordem j:

$$\begin{aligned}
\text{cov}\left[y_t, y_{t-j}\right] &= \mathbb{E}\left[\left(y_t - \mathbb{E}\left[y_t\right]\right)\left(y_{t-j} - \mathbb{E}\left[y_{t-j}\right]\right)\right] \\
&= \mathbb{E}\left[\left(\varepsilon_t + \theta\varepsilon_{t-1}\right)\left(\varepsilon_{t-j} + \theta\varepsilon_{t-j}\right)\right] \\
&= 0 \text{ para } j > 1
\end{aligned}$$

FAC de um MA (1)

Portanto, a FAC de um MA(1) será dada por:

$$\rho(1) = \frac{\theta\sigma^2}{(1+\theta^2)\,\sigma^2} = \frac{\theta}{(1+\theta^2)}$$
$$\rho(j) = 0 \text{ para } j > 1$$

A FAC de um MA(1) será, então, truncada em 1.

E generalizando, a FAC de um MA(1) será truncada em q.

- Um MA(1) pode ser escrito como uma AR(∞). Vejamos:

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

Isolando ε_t, *temos* : $\varepsilon_t = y_t + \theta\varepsilon_{t-1}$
Mas sabemos que :

$$\varepsilon_{t-1} = y_{t-1} + \theta\varepsilon_{t-2}$$

Substituindo, temos:

$$\varepsilon_t = y_t - \theta(y_{t-1} + \theta\varepsilon_{t-2}$$

$$\varepsilon_t = y_t - \theta y_{t-1} + \theta\varepsilon_{t-2}$$

Fazendo sucessivamente, chegaremos ao seguinte resultado:

$$\varepsilon_t = y_t - \theta y_{t-1} + \theta^2 y_{t-2} - -\theta^3 y_{t-3} - \dots$$

Que é um AR(∞). Portanto, a FACP de um MA(1) será declinante, já que $\theta < 1$.

FAC e FACP para um MA(1) com $\theta = 0.9$

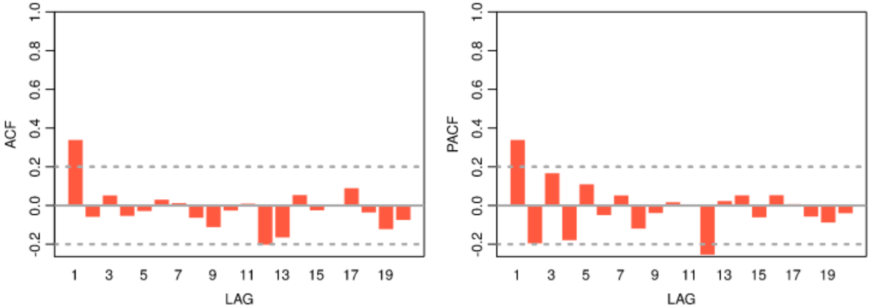


Figure: Fonte: Kotzé, K. (2019)

FAC e FACP de um MA (1)

MA(1) com $\theta = 0.9$

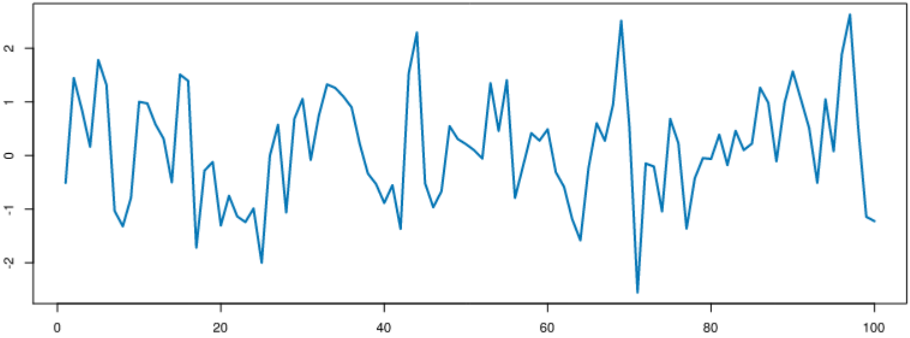


Figure: Fonte: Kotzé, K. (2019)

Modelos ARMA(p,q)

Um modelo ARMA(p,q) é a soma de processos AR(p) e MA(q). Um ARMA(1,1) pode ser escrito como:

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

Ou, usando o operador defasagem, podemos escrever o ARMA(p,q):

$$\phi(L) y_t = \theta(L) \varepsilon_t$$

$$\phi(L) = 1 - \sum_{i=1}^{p} \phi_i L^i \text{ and } \theta(L) = 1 + \sum_{i=1}^{q} \theta_i L^i$$

E um ARMA(2,1) seria representado por:

$$\begin{aligned} \left(1 - \phi_1 L - \phi_2 L^2\right) y_t &= (1 + \theta_1 L)\, \varepsilon_t \\ y_t &= \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t + \theta_1 \varepsilon_{t-1} \end{aligned}$$

FAC e FACP de modelos ARMA(p,q)

- ARMA(p,q) é a soma dos processos AR(p) e MA(q).

- Portanto, sua FAC será declinate, bem como sua FACP.

FAC e FACP de modelos ARMA(p,q)

FAC e FACP para um ARMA(1,1) com $\phi = 0.5$ $\theta = 0.6$

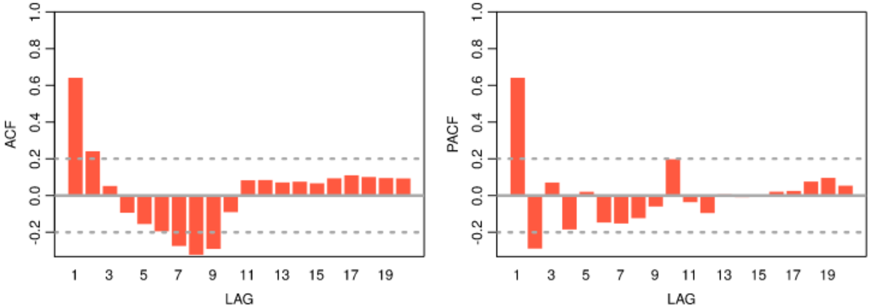


Figure: Fonte: Kotzé, K. (2019)

Estimação

- Modelos AR podem ser estimados por Mínimos Quadrados Ordinários (MQO).

- Modelos MA e ARMA são estimados através do método da Máxima Verossimilhança (MV), pois os erros que atuam como variáveis independentes no modelo não são diretamente observáveis.

Diagnóstico: resíduos

- Quando realizamos a etapa de identificação do modelo, estamos trabalhando com funções amostrais e não populacionais.
- Portanto, a análise das FAC e FACP não dá uma resposta definitiva sobre qual modelo deve ser estimado.
- Após a estimação, deve ser realizado uma análise dos resíduos (diagnóstico) para saber se o modelo de fato é adequado.
- Se o modelo for adequado, os resíduos terão comportamento de ruído branco, ou seja, média zero, variância constante e ausência de autocorrelação.

Diagnóstico: resíduos

- Podemos fazer isso através da análise das FAC e FACP dos resíduos, além dos testes de Ljung-Box e Box-Pierce.

- Se houver qualquer evidência de correlação serial nos resíduos, ele deve ser revisto.

- Modelos parcimoniosos são preferidos. Portanto, se um modelo com menos parâmetros tiver um diagnóstico adequado, ele é preferível a um modelo com mais parâmetros.

Teste de Ljung-Box

A hipótese nula do teste de Ljung-Box é que todas as autocorrelação são nulas, contra a hipótese alternativa que pelo menos uma das autocorrelações é diferente de zero.

A estatística do teste é dada por:

$$Q = T(T+2)\frac{\sum_{k=1}^{s}\rho_j^2}{(T-j)} \sim \chi^2(s)$$

Se o valor da estatística Q calculada for maior que o valor crítico da distribuição $\chi^2(s)$, então rejeita-se a hipótese nula de que todas as autocorrelações são nulas.

Teste de Ljung-Box

- Para o caso do diagnóstico do modelo, para que este seja adequado, não pode haver nenhuma correlação significativa.

- Portanto, nenhuma estatística Q deve ser estatisticamente significativa (não rejeitar a hipótese nula).

Critérios de informação

Os critérios de informação avaliam o ajuste (fit) de um modelo penalizando pelo número de parâmetros. Os dois principais são o Critério de Informação de Akaike e o Critério de Informação Bayesiano, descritos abaixo:

$$AIC = \log \hat{\sigma}_k^2 + \frac{T + 2k}{T}$$
$$BIC = \log \hat{\sigma}_k^2 + \frac{k \log T}{T}$$

em que

$$\hat{\sigma}_k^2 = \frac{SSR_k}{T}$$

Critérios de informação

- Os critérios de informação penalizam a variância do erro por um termo proporcional aos graus de liberdade.
- A interpretação é que quanto menor o valor calculado do critério de informação, "melhor" é o modelo.
- Se a estatística for negativa, você deve selecionar o modelo com a estatística de teste mais próxima de $-\infty$
- Entre vários modelos adequados, ou seja, que passaram no teste de diagnóstico, você pode escolher o "melhor" através dos critérios de informação descritos aqui.

Previsão (*Forecasting*)

- A habilidade de um modelo prever o comportamento futuro de uma variável com razoável grau de acurácia é um importante tópico em séries temporais.
- Uma decisão tomada hoje é baseada no que se espera que ocorrerá no futuro.
- O futuro pode ser o próximo minuto, próximo dia, mês, ano , etc.
- Por exemplo, um *day-trader* quer prever o preço de uma ação no próximo minuto, hora, dia, etc.
- Já o Banco Central está interessado em prever a taxa de inflação no próximo ano.
- Previsão não é uma tarefa fácil.

Previsão (*Forecasting*)

Considere o modelo AR(1) que vimos anteriormente:

$$y_t = \phi_1 y_{t-1} + \epsilon_t$$

com

$$\epsilon_t \sim \text{i. i. d. } \mathcal{N}\left(0, \sigma^2\right)$$

Para prever os valores futuros, nós simplesmente iteramos o modelo para frente (*forward*) quantas vezes forem necessárias para calcular os próximos valores. Abaixo, fazemos isso para H passos a frente:

$$\begin{aligned}
y_{t+1} &= \phi_1 y_t + \epsilon_{t+1} \\
y_{t+2} &= \phi_1 y_{t+1} + \epsilon_{t+2} \\
\vdots &= \vdots \\
y_{t+H} &= \phi_1 y_{t+H-1} + \epsilon_{t+H}
\end{aligned}$$

Previsão (*Forecasting*)

Usando as equações anteriores recursivamente (inserindo a primeira linha na segunda, a segunda na terceira e assim por diante), podemos escrever a seguinte expressão resumida para o valor previsto de y, H passos a frente:

$$y_{t+h} = \phi_1^h y_t + \sum_{i=0}^{h-1} \phi_1^i \epsilon_{t+h-i}$$

Previsão (*Forecasting*)

- Note que o valor previsto de y, y_{t+h} depende de y_t, que representa o conjunto de informação até a data t, mais uma soma dos choques esperados futuros.

- Infelizmente, não temos informação sobre o valor realizado dos choques futuros, ou seja, não faz parte do conjunto de informação até a data t.

- Portanto, pode haver previsões bastante ruins dependendo do valor realizado desses choques.

Metodologia de Box-Jenkins

- Identificação do modelo através das FAC e FACP
  - AR(p): FAC truncada em p e FACP declinante
  - MA(q): FAC declinante e FACP truncada em q
  - ARMA(p,q): FAC declinante e FACP declinate
- Estimação
- Diagnóstico
- Previsão