# Aula 2: Análise de séries temporais univariadas (ARIMA)

Marcus L. Nascimento

6 de outubro de 2025

# Introdução

# Introdução

- **Série temporal**: Sequência de variáveis aleatórias $Y_1, Y_2, \ldots, Y_n$ indexadas no tempo $(t = 1, 2, \ldots, n)$.
    - Por vezes, denotamos a sequência de variáveis aleatórias por $Y_{t_1}, Y_{t_2,\ldots,Y_{t_n}}$;
    - Frequentemente, tal sequência é considerada um subconjunto de uma sequência (possivelmente, infinita) $\{Y_t\}$ de variáveis aleatórias.
- Propósito em análise de séries temporais:
    - **Quantificar dependências** no tempo;
    - **Utilizar as correlações** para **explicar** as observações disponíveis e **inferir** sobre valores não observados.

# Estacionariedade

- Um conceito fundamental em análise de séries temporais recai sobre a ideia de **estacionariedade**.
    - Medidas estatísticas (por exemplo, média, variância e correlação) são constantes ao longo do tempo;
    - Estabilidade no processo gerador subjacente facilita a identificação de padrões e relações nos dados;
    - Modelos estatísticos e de aprendizado de máquinas pressupõem estacionariedade.
- Estacionariedade **estrita** e **fraca**.

# Estacionariedade estrita

DEFINIÇÃO: Uma série temporal $\{Y_t\}$ é dita **estritamente estacionária** se a distribuição conjunta de $(Y_{t_1}, \ldots, Y_{t_k})$ é idêntica à distribuição de $(Y_{t_1+s}, \ldots, Y_{t_k+s})$ para todo $s$, onde $k$ é um número inteiro positivo arbitrário e $(t_1, \ldots, t_k)$ é uma coleção de $k$ inteiros positivos.

OBSERVAÇÕES:

- A distribuição conjunta de $(Y_{t_1}, \ldots, Y_{t_k})$ é invariante sob translações temporais;
- Estacionariedade estrita é uma condição bastante forte;
- Difícil verificação empírica.

# Estacionariedade fraca (segunda ordem)

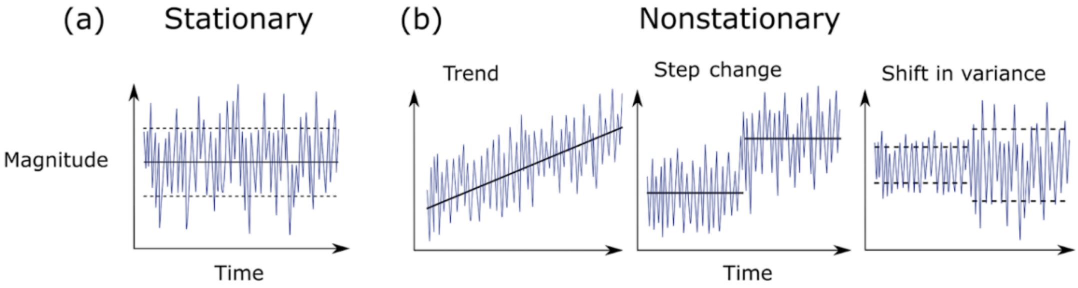
DEFINIÇÃO: Uma série temporal $\{Y_t\}$ é dita **fracamente estacionária** (ou estacionária de segunda ordem) se a média de $Y_t$ e a covariância entre $Y_t$ e $Y_{t-l}$ são invariantes no tempo, onde $l$ é um número inteiro arbitrário.

Mais especificamente, $\{Y_t\}$ é dita fracamente estacionária se:

1. $\mu_Y(t) = \mathrm{E}(Y_t) = \mu$, $|\mu| < \infty$;
2. $\sigma^2_Y(t) = E[(Y_t - \mu)^2] = \gamma_0$, $\gamma_0 < \infty$;
3. $\gamma_Y(t, t-l) = \mathrm{Cov}(Y_t, Y_{t-l}) = E[(Y_t - \mu)(Y_{t-l} - \mu)] = \gamma_l \; \forall l$ com $|\gamma_l| < \infty$.

Estacionariedade de segunda ordem implica que $Cov(Y_{t_1}, Y_{t_1-l}) = Cov(Y_{t_2}, Y_{t_2-l}) = \gamma_l$.

# Estacionariedade Fraca



# Função de autocorrelação (ACF)

- Sob estacionariedade fraca, a função de autocorrelação de defasagem $l$ (lag $l$) é dada por:

$$\rho_l = \rho_Y(t, t-l) = \text{Corr}(Y_t, Y_{t-l}) = \frac{\gamma_Y(t, t-l)}{\sqrt{\sigma_Y^2(t)\sigma_Y^2(t-l)}} = \frac{\gamma_l}{\gamma_0}$$

- A autocorrelação amostral de defasagem $l$ (lag $l$), por sua vez, é definida como

$$\hat{\rho}_l = \frac{\sum_{t=l+1}^{n}(y_t - \bar{y})(y_{t-l} - \bar{y})}{\sum_{t=1}^{n}(y_t - \bar{y})^2}, \ 0 \leq l < n-1.$$

# Função de autocorrelação (ACF) - Teste de Portmanteau

- A verificação quanto à presença ou não de autocorrelação em uma série temporal pode ser realizada através de testes de hipótese da seguinte forma:

$$\begin{aligned} \mathrm{H}_0 &: \rho_1 = \rho_2 = \ldots = \rho_m = 0 \\ \mathrm{H}_1 &: \rho_i \neq 0 \text{ para algum } i \in \{1, 2, \ldots, m\}. \end{aligned}$$

- Proposto por Box and Pierce (1970), o teste de Portmanteau aplica a seguinte estatística de teste:

$$Q^*(m) = n \sum_{l=1}^{m} \hat{\rho}_l^2.$$

- Assumindo que $\{Y_t\}$ é uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d) com certas condições acerca dos momentos, $Q^*(m)$ é assintoticamente $\chi_m^2$.

# Função de autocorrelação (ACF) - Teste de Ljung-Box

- Com intuito de aumentar o poder do teste em amostras finitas, Ljung and Box (1978) propuseram uma modificação na estatística $Q^*(m)$ da seguinte forma:

$$Q(m) = n(n+2) \sum_{l=1}^{m} \frac{\hat{\rho}_l^2}{n-1}.$$

- Neste caso, a regra de decisão é rejeitar $\mathrm{H}_0$ se $Q(m) > q_\alpha^2$, onde:
  - $q_\alpha^2$ denota o $100(1-\alpha)$-ésimo percentil da distribuição $\chi_m^2$;
  - $\alpha$ é um nível de significância fixado.

# Função de autocorrelação parcial (PACF)

- Em geral, o conceito de correlação parcial se refere a uma correlação condicional.
- No contexto de séries temporais, a autocorrelação parcial entre $Y_t$ e $Y_{t-l}$ é definida como a correlação entre $Y_t$ e $Y_{t-l}$ condicional em $Y_{t-l+1}, \ldots, Y_{t-1}$ (conjunto de observação entre os tempos $t$ e $t-l$.

$$\phi_{l,l} = \frac{\mathrm{Cov}(Y_t, Y_{t-l} | Y_{t-l+1}, \ldots, Y_{t-1})}{\sqrt{\mathrm{Var}(Y_t | Y_{t-l+1} = y_{t-l+1}, \ldots, Y_{t-1} = y_{t-1}) \mathrm{Var}(Y_{t-l} | Y_{t-l+1} = y_{t-l+1}, \ldots, Y_{t-1} = y_{t-1})}}$$

Exemplos:

- A autocorrelação parcial de primeira ordem é definida para ser igual à autocorrelação de primeira ordem;
- Autocorrelação de segunda ordem: $\phi_{2,2} = \frac{\mathrm{Cov}(Y_t, Y_{t-2} | Y_{t-1})}{\sqrt{\mathrm{Var}(Y_t | Y_{t-1} = y_{t-1}) \mathrm{Var}(Y_{t-2} | Y_{t-1} = y_{t-1})}}$;

# Séries Temporais Lineares

# Ruído branco

- Uma série temporal $\{W_t\}$ é denominada **ruído branco** se $\{W_t\}$ é uma sequência de variáveis aleatórias independentes e identicamente distribuídas com média 0 e variância constante $\tau^2$ finita.

  - $\{W_t\}$ é estacionária;
  - Se $s \neq t$, $\text{Cov}(W_s, W_t) = \mathrm{E}[(W_s - \mu)(W_t - \mu)] = 0$;
  - $\rho_l = 0$ para todo $l$.

- Se $W_t$ tem distribuição normal com média 0 e variância $\tau^2$, a série temporal é denominada ruído branco Gaussiano.

# Série temporal linear

- Uma série temporal $\{Y_t\}$ é dita linear se pode ser escrita da seguinte forma:

$$y_t = \mu + \sum_{i=0}^{\infty} \psi_i w_{t-i},$$

onde $\mu$ é a média de $Y_t$, $\psi_0 = 1$ e $\{W_t\}$ é ruído branco.

- $w_t$ denota a nova informação da série temporal no tempo $t$ e é frequentemente chamado de inovação ou choque em $t$.

- Se $\{Y_t\}$ é fracamente estacionária, pela independência de $\{W_t\}$, $\mathrm{E}(Y_t) = \mu$ e $\mathrm{Var}(Y_t) = \tau^2 \sum_{i=0}^{\infty} \psi_i^2$.

# Série temporal linear

- A covariância de defasagem $l$, $\gamma_Y(t, t-l)$, é dada por:

$$
\begin{aligned}
\gamma_Y(t, t-l) &= \text{Cov}(Y_t, Y_{t-l}) \\
&= \mathrm{E}\left[\left(\sum_{i=0}^{\infty} \psi_i w_{t-i}\right)\left(\sum_{j=0}^{\infty} \psi_j w_{t-l-j}\right)\right] \\
&= \mathrm{E}\left(\sum_{i,j=0}^{\infty} \psi_i \psi_j w_{t-i} w_{t-l-j}\right) \\
&= \sum_{j=0}^{\infty} \psi_{j+l} \psi_j \mathrm{E}(w_{t-l-j}^2) = \tau^2 \sum_{j=0}^{\infty} \psi_{j+l} \psi_j .
\end{aligned}
$$

- Modelos lineares de séries temporais são modelos estatísticos e econométricos utilizados para descrever o padrão dos pesos $\psi$ de $Y_t$.

# Processos Autorregressivos (AR)

# Processo Autorregressivo de Ordem 1 - AR(1)

- Seja $\{W_t\}$ ruído branco, um processo AR(1) pode ser descrito como a seguir:

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

- Por recursão, podemos mostrar que um processo AR(1) pode ser descrito como um modelo linear de série temporal.

$$\begin{aligned}
y_t &= \phi_1 y_{t-1} + w_t \\
&= \phi_1(\phi_1 y_{t-2} + w_{t-1}) + w_t \\
&= \phi_1(\phi_1(\phi_1 y_{t-3} + w_{t-2}) + w_{t-1}) + w_t \\
&\vdots \\
&= \sum_{i=0}^{\infty} \phi_1^i w_{t-i}.
\end{aligned}$$

# Processo Autorregressivo de Ordem 1 - AR(1)

- Sob estacionariedade, temos que

$$
\begin{aligned}
\text{Var}(Y_t) &= \text{Var}(\phi_1 Y_{t-1} + W_t) \\
&= \text{Var}(\phi_1 Y_{t-1}) + \text{Var}(W_t) \\
&= \phi^2 \text{Var}(Y_{t-1}) + \tau^2 \\
&= \frac{\tau^2}{1-\phi_1^2},
\end{aligned}
$$

$\phi_1^2 < 1$. Em outras palavras, a estacionariedade fraca de um modelo AR(1) implica que $|\phi_1| < 1$.

# Processo Autorregressivo de Ordem 1 - AR(1)

- Ainda sob estacionariedade, temos que

$$
\begin{aligned}
\gamma_l &= \gamma_Y(t, t-l) = \text{Cov}(Y_t, Y_{t-l}) = \mathrm{E}(Y_t Y_{t-l}) \\
&= \mathrm{E}((\phi_1 Y_{t-1} + W_t) Y_{t-l}) \\
&= \phi_1 \mathrm{E}(Y_{t-1} Y_{t-l}) + \mathrm{E}(W_t Y_{t-l}) \\
&= \phi_1 \mathrm{E}(Y_{t-1} Y_{t-l}) \\
&= \phi_1^2 \mathrm{E}(Y_{t-2} Y_{t-l}) \\
&\vdots \\
&= \phi_1^l \mathrm{E}(Y_{t-l} Y_{t-l}) \\
&= \phi_1^l \mathrm{E}(Y_t Y_t) \\
&= \phi_1^l \gamma_0,
\end{aligned}
$$

para $l = -1, 1, -2, 2, \ldots$.

# Processo Autorregressivo de Ordem 1 - AR(1)

- Para encontrar $\gamma_0$, basta seguirmos de forma similar.

$$\begin{aligned}
\gamma_0 &= \mathrm{E}(Y_t Y_t) \\
&= \mathrm{E}((\phi_1 Y_{t-1} + W_t) Y_t) \\
&= \phi_1 \mathrm{E}(Y_{t-1} Y_t) + \mathrm{E}(W_t Y_t) \\
&= \phi_1 \gamma_1 + \tau^2.
\end{aligned}$$

- Como $\gamma_1 = \phi_1 \gamma_0$ e $\gamma_0 = \phi_1 \gamma_1 + \tau^2$, segue que $\gamma_0 = \frac{\tau^2}{1-\phi_1^2}$.

- A função de autocorrelação de um modelo AR(1), portanto, é dada por $\rho_l = \frac{\gamma_l}{\gamma_0} = \phi_l^l$, para $l = -1, 1, -2, 2, \ldots$.

# Processo Autorregressivo de Ordem 2 - AR(2)

- Seja $\{W_t\}$ ruído branco, um processo AR(2) assume a seguinte forma:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + w_t.$$

- Para encontrar a função de covariância do modelo AR(2), aplicamos ideias similares às aplicadas no modelo AR(1).

$$\begin{aligned}
\gamma_l &= E(Y_t Y_{t-l}) \\
&= E((\phi_1 Y_{t-1} + \phi_2 Y_{t-1} + W_t) Y_{t-l}) \\
&= \phi_1 E(Y_{t-1} Y_{t-l}) + \phi_2 E(Y_{t-2} Y_{t-l}) \\
&= \phi_1 \gamma_{l-1} + \phi_2 \gamma_{l-2} + \mathrm{E}(W_t Y_{t-l}).
\end{aligned}$$

$$
\begin{aligned}
(l=0): \gamma_0 &= \phi_1\gamma_{-1} + \phi_2\gamma_{-2} + \mathrm{E}(W_t Y_{t-l}) \\
\gamma_0 &= \phi_1\gamma_1 + \phi_2\gamma_2 + \tau^2
\end{aligned}
$$

$$
\begin{aligned}
(l=1): \gamma_1 &= \phi_1\gamma_0 + \phi_2\gamma_1 + \mathrm{E}(W_t Y_{t-l}) \\
&= \phi_1\gamma_0 + \phi_2\gamma_1 \\
&= \frac{\phi_1}{1-\phi_2}\gamma_0
\end{aligned}
$$

$$
\begin{aligned}
(l=2): \gamma_2 &= \phi_1\gamma_1 + \phi_2\gamma_0 + \mathrm{E}(W_t Y_{t-l}) \\
&= \phi_1\gamma_1 + \phi_2\gamma_0 \\
&= \left[\frac{\phi_1^2}{1-\phi_2} + \phi_2\right]\gamma_0
\end{aligned}
$$

# Processo Autorregressivo de Ordem 2 - AR(2)

Substituindo $\gamma_1$ e $\gamma_2$ em $\gamma_0$, temos:

$$
\begin{aligned}
\gamma_0 &= \frac{\phi_1^2}{1-\phi_2}\gamma_0 + \left[\frac{\phi_2\phi_1^2}{1-\phi_2} + \phi_2^2\right]\gamma_0 + \tau^2 \\
&= \frac{(1-\phi_2)\tau^2}{(1-\phi_2) - \phi_1^2(1+\phi_2) + \phi_2^2(1-\phi_2)}.
\end{aligned}
$$

A função de autocorrelação, por sua vez, é dada por:

$$
\begin{aligned}
\rho_l &= \frac{\gamma_l}{\gamma_0} \\
&= \phi_1\frac{\gamma_{l-1}}{\gamma_0} + \phi_2\frac{\gamma_{l-2}}{\gamma_0} \\
&= \phi_1\rho_{l-1} + \phi_2\rho_{l-2}.
\end{aligned}
$$

# Processo Autorregressivo de Ordem 2 - AR(2)

- Note que, a partir dos resultados encontrados para as covariâncias, conhecendo os valores para as funções $\gamma_0$, $\gamma_1$ e $\gamma_2$, é possível um sistema com três equações e três parâmetros desconhecidos ($\phi_1$, $\phi_2$ e $\tau^2$)

- Equações de Yule-Walker.

- Substituindo os valores teóricos de $\gamma_0$, $\gamma_1$, $\gamma_2$ por suas estimativas $\hat{\gamma}_0$, $\hat{\gamma}_1$, $\hat{\gamma}_2$, obtemos estimativas $\hat{\phi}_1$, $\hat{\phi}_2$ e $\hat{\tau}^2$.

- Tal método é de simples implementação e é utilizado no ajuste de modelos autorregressivos (uma vez que a ordem tenha sido fixada).

- Também aplicado em modelos autorregressivos mais gerais (ordem $p$). Neste caso, temos um sistema linear com $(p+1)$ equações e $(p+1)$ parâmetros.

# Processo Autorregressivo de Ordem $p$ - AR($p$)

- Os resultados para os modelos AR(1) e AR(2) podem ser generalizados para um modelo AR($p$):

$$y_t = \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} + w_t,$$

onde $p$ é um inteiro não negativo e $\{W_t\}$ é um ruído branco.

- De forma mais geral, $\{Y_t\}$ é AR($p$) se existe $\phi_0$ tal que $\{(Y_t - \phi_0)\}$ é AR($p$).

- Neste caso, o valor esperado de $Y_t$ é dado por:

$$\mathrm{E}(Y_t) = \frac{\phi_0}{1 - \phi_1 - \ldots - \phi_p}.$$

# Processo Autorregressivo de Ordem $p$ - AR($p$)

- Vimos que é possível estimar os parâmetros do modelo a partir de um sistema de equações uma vez fixada a ordem do processo autorregresivo.
- Uma pergunta relevante neste caso seria como definimos a ordem do processo autorregressivo neste caso.
- Considerando um modelo AR($p$) estacionário Gaussiano temos os seguintes resultados para a função de autocorrelação parcial (PACF) amostral:
  - $\hat{\phi}_{p,p} \to \phi_p$, $n \to \infty$;
  - $\hat{\phi}_{l,l} \to 0$ para todo $l > p$;
  - $\text{Var}(\hat{\phi}_{l,l}) \to 1/n$ para todo $l > p$.
- Tais resultados indicam que o PACF de um modelo AP(p) zera na defasagem $p$.

# Processo Autorregressivo de Ordem $p$ - AR($p$)

- Uma alternativa a identificação da ordem de um processo autorregressivo via função de autocorrelação parcial (PACF) se dá através de critérios de seleção de modelos.

- Para identificação da ordem de um processo AR via critérios de seleção de modelos, seguimos os seguintes passos:

  1. Ajuste de modelos AR com diferentes ordens;
  2. Calcule o critério para cada um dos modelos;
  3. Selecione o modelo com melhor métrica.

- Dentre os critérios de seleção de modelos, destacamos o AIC (Akaike Information Criterion) e o BIC (Bayesian information criterion).

# Processo Autorregressivo de Ordem $p$ - AR($p$)

- O AIC (Akaike Information Criterion) é dado pela seguinte equação;

$$\text{AIC} = \underbrace{-2 \times \log(\text{verossimilhança})}_{\text{bondade do ajuste}} + \underbrace{2 \times (\text{número de parâmetros})}_{\text{penalização}},$$

  onde a verossimilhança é avaliada considerando o modelo estimado.

- Considerando um modelo AR($p$) estimado utilizando Yule-Walker, ao invés de avaliar a função de verossimilhança, podemos aplicar:

$$\text{AIC} = n \times (\log(\tilde{\tau}^2) + 1) + 2(p + 1),$$

  onde $n$ é o tamanho da amostra e $\tilde{\tau}^2$ é o valor estimado para a variância do ruído branco ($\tau^2$).

# Processo Autorregressivo de Ordem $p$ - AR($p$)

- O BIC (Bayesian information criterion), por sua vez, é dado pela seguinte equação;

$$\text{BIC} = \underbrace{-2 \times \log(\text{verossimilhança})}_{\text{bondade do ajuste}} + \underbrace{(\text{número de parâmetros}) \times \log(\text{tamanho da amostra})}_{\text{penalização}},$$

  onde a verossimilhança é avaliada considerando o modelo estimado.

- As penalizações para o AIC e o BIC são 2 e $\log(n)$ respectivamente.
  - BIC tende a selecionar modelos AR de ordens menores quando o tamanho da amostra é moderado ou grande.

- A ordem do processo é identificada através dos modelos que produzem os menores valores para o AIC ou para o BIC.

# Processo Autorregressivo de Ordem $p$ - AR($p$)

- Em análise de séries temporais, muito do nosso interesse está em fazer previsões a partir dos dados observados.

- Suponha que estejamos no tempo $h$ (**origem**) e tenhamos interesse em prever $Y_{h+l}$ onde $l \geq 1$ (**horizonte**).

  - $F_h$ é a informação disponível em $h$;

  - $\hat{y}_h(l)$ é a previsão $l$-passos a frente de $Y_t$ com origem $h$.

# Processo Autorregressivo de Ordem $p$ - AR($p$)

- Previsão 1-passo a frente:

$$\hat{y}_h(1) = \mathrm{E}(Y_{h+1}|F_n) = \phi_0 + \phi_1 y_h + \phi_2 y_{h-1} + \ldots + \phi_p y_{h+1-p}.$$

- Previsão 2-passos a frente:

$$\begin{aligned} \hat{y}_h(2) &= \mathrm{E}(Y_{h+2}|F_n) \\ &= \phi_0 + \phi_1 \hat{y}_h(1) + \phi_2 y_h + \ldots + \phi_p y_{h+2-p} \\ &= \phi_0 + \phi_1(\phi_0 + \phi_1 y_h + \ldots + \phi_p y_{h+1-p}) + \phi_2 y_h + \ldots + \phi_p y_{t+2-p}. \end{aligned}$$

- Podemos aplicar o mesmo procedimento para computar $\hat{y}_h(k)$ para qualquer $k$.
- As previsões são combinações lineares dos $p$ valores observados mais recentes da série com coeficientes computados indutivamente a partir dos coeficientes do modelo.
- Previsões convergem rapidamente para a média da série ($\phi_0$).

# Processos de Médias Móveis (MA)

# Processo de Média Móvel de Ordem 1 - MA(1)

- Seja $\{W_t\}$ ruído branco, um processo MA(1) pode ser descrito como a seguir:

$$y_t = c_0 + \theta_1 w_{t-1} + w_t.$$

- Por conveniência, assumimos $c_0 = 0$ para computar as autocovariâncias.

$$\begin{aligned}
\gamma_l &= \mathrm{E}(Y_t Y_{t-l}) \\
&= \mathrm{E}((\theta_1 W_{t-1} + W_t)(\theta_1 W_{t-l-1} + W_{t-l}) \\
&= \mathrm{E}(\theta_1^2 W_{t-1} W_{t-l-1}) + \mathrm{E}(\theta_1 W_{t-1} W_{t-l}) + \mathrm{E}(\theta_1 W_t W_{t-l-1}) + \mathrm{E}(W_t W_{t-l}).
\end{aligned}$$

- Para $l = 0$, temos:

$$\begin{aligned}
\gamma_0 &= \mathrm{E}(\theta_1^2 W_{t-1} W_{t-1}) + \mathrm{E}(\theta_1 W_{t-1} W_t) + \mathrm{E}(\theta_1 W_t W_{t-1}) + \mathrm{E}(W_t W_t) \\
&= \theta_1^2 \tau^2 + \tau_2 \\
&= (1 + \theta_1^2)\tau^2.
\end{aligned}$$

# Processo Média Móvel de Ordem 1 - MA(1)

- Para $l = 1$, temos:

$$\begin{aligned}\gamma_1 &= \mathrm{E}(\theta_1^2 W_{t-1} W_{t-2}) + \mathrm{E}(\theta_1 W_{t-1} W_{t-1}) + \mathrm{E}(\theta_1 W_t W_{t-2}) + \mathrm{E}(W_t W_{t-1}) \\ &= \theta_1 \tau^2.\end{aligned}$$

- Para $l > 1$, temos $\gamma_l = 0$.

- Note que a função de autocovariância será zero para toda defasagem maior que 1 em um modelo MA(1).

- $\rho_0 = 1$, $\rho_1 = \theta_1/(1 + \theta_1^2)$, $\rho_l = 0$ para $l > 1$.

# Processo Média Móvel de Ordem 2 - MA(2)

- Seja $\{W_t\}$ ruído branco, um processo MA(2) é dado pela seguinte equação:

$$y_t = c_0 + \theta_1 w_{t-1} + \theta_2 w_{t-2} + w_t.$$

- Assumindo $c_0 = 0$, teremos as autocovariâncias como a seguir:

$$\begin{aligned}
\gamma_0 &= (1 + \theta_1^2 + \theta_2^2)\tau^2 \\
\gamma_1 &= (\theta_1 + \theta_1\theta_2)\tau^2 \\
\gamma_2 &= \theta_2\tau^2 \\
\gamma_l &= 0. \text{ para } l > 2.
\end{aligned}$$

- Consequentemente, temos:

$$\begin{aligned}
\rho_0 &= 1 \\
\rho_1 &= (\theta_1 + \theta_1\theta_2)/(1 + \theta_1^2 + \theta_2^2) \\
\rho_2 &= \theta_2/(1 + \theta_1^2 + \theta_2^2) \\
\rho_l &= 0. \text{ para } l > 2.
\end{aligned}$$

# Processo Média Móvel de Ordem $q$ - MA($q$)

- Seja $\{W_t\}$ ruído branco, generalizando os modelos MA(1) e MA(2), descrevemos um modelo MA($q$):

$$y_t = c_0 + \theta_1 w_{t-1} + \ldots + \theta_q w_{t-q} + w_t.$$

- Processos de médias móveis sempre fracamente estacionários.
  - $\mathrm{E}(Y_t) = c_0$;
  - $\mathrm{Var}(Y_t) = (1 + \theta_1^2 + \theta_2^2 + \ldots + \theta_q^2)\tau^2$.
- Assumindo $c_0 = 0$, a função de autocovariância $\gamma_l$ e, consequentemente, a função de autocorrelação $\rho_l$ serão iguais a 0 para $l > q$.
- Tal resultado dá uma indicação acerca da escolha da ordem de um processo de médias móveis.

# Processo Média Móvel de Ordem $q$ - MA($q$)

- Assim como verificamos que um processo AR(1) é um processo MA($\infty$), também por recursão, podemos verificar que um processo MA(1) é um processo AR($\infty$).

$$
\begin{aligned}
y_t &= \theta_1 w_{t-1} + w_t \\
&= \theta_1 (x_{t-1} - \theta_1 w_{t-2}) + w_t \\
&= \theta_1 x_{t-1} - \theta_1^2 w_{t-2} + w_t \\
&= \theta_1 x_{t-1} - \theta_1^2 (x_{t-2} - \theta_1 w_{t-3}) + w_t \\
&= \theta_1 x_{t-1} - \theta_1^2 x_{t-2} - \theta_1^3 w_{t-3} + w_t \\
&\vdots \\
&= \theta_1 x_{t-1} - \theta_1^2 x_{t-2} - \theta_1^3 w_{t-3} - \ldots + w_t
\end{aligned}
$$

- Ressalta-se que o resultado vale para $|\theta_1| < 1$

# Processos Autorregressivos de Médias Móveis (ARMA)

# Processos Autorregressivos de Médias Móveis

- Seja $\{W_t\}$ ruído branco, um processo autorregressivo de médias móveis de ordem $p$ e $q$, ARMA($p$,$q$), é descrito da seguinte forma:

$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j w_{t-j} + w_t.$$

- Uma série temporal é dita um processo ARIMA se, quando tomada a diferença um número finito de vezes, a mesma se torna um processo ARMA.

- $Y \sim$ ARIMA$(p, d, q)$ se $Y \sim$ ARMA$(p, q)$ após $d$ diferenças serem tomadas.
  - $Y \sim \text{ARIMA}(p, d, q) \iff \nabla^d Y \sim \text{ARMA}(p, q)$.

# Referências


Box, G. E. P., and David A. Pierce. 1970. "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models." *Journal of the American Statistical Association* 65 (332): 1509–1526.

Ljung, G. M., and G. E. P. Box. 1978. "On a measure of lack of fit in time series models." *Biometrika* 65 (2): 297–303.