

Capítulo 5

# Previsão

Econometria II · UFBA

Prof. Pablo Castro

## O que veremos

- Por que prever?
- Modelo de previsão: AR(1) passo a passo
- Previsão para  $h$  períodos à frente
- Erro de previsão e variância
- Intervalo de confiança (“cone”)
- Algoritmo geral para qualquer ARIMA
- Previsão dentro vs. fora da amostra
- Medidas de desempenho (RMSE, MAE, MAPE)
- Rolling window (validação temporal)

### Box-Jenkins?

1. Identificação (Cap. 3–4)



2. Estimação (Cap. 4)



3. Diagnóstico (Cap. 4)



4. Previsão (Cap. 5) ← aqui

Seção 5.1

## **Introdução à Previsão**

## Por que prever?

---

**Quase toda análise econômica termina em previsão.**

### Exemplos brasileiros

**Governo:** planejar o orçamento do próximo ano exige prever a arrecadação tributária.

**BCB:** para definir a Selic, considera a projeção do IPCA.

**Empresas:** planejar produção e estoque exige prever demanda e vendas.

**Bancos:** avaliar risco de crédito depende de prever PIB, desemprego e inadimplência.

**O que assumimos:** já ajustamos um modelo ARIMA via Box-Jenkins (Cap. 4).

Agora vamos usá-lo para prever valores futuros – a quarta etapa da metodologia.

### Dois tipos de previsão:

- **Pontual:** um único número (ex.: “IPCA de 2026 = 4,2%”).
- **Intervalar:** um intervalo de confiança (ex.: “IPCA entre 3,8% e 4,6% com 95% de confiança”).

## O que acontece antes de prever: um resumo rápido

---

**Já fizemos** (Cap. 2–4):

1. Testamos raiz unitária  $\rightarrow$  encontramos  $d$ .
2. Diferenciamos  $d$  vezes para estacionarizar.
3. Lemos ACF e PACF  $\rightarrow$  sugerimos  $p$  e  $q$ .
4. Estimamos os candidatos e escolhemos por AIC/BIC.
5. Diagnosticamos resíduos (Ljung-Box).

**Com isso, temos:**

Um modelo  $ARIMA(p, d, q)$  com coeficientes estimados  $\hat{\mu}, \hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\theta}_1, \dots, \hat{\theta}_q$  e resíduos  $\hat{u}_t$ .

**Pergunta central:**

Dado o modelo ajustado, qual é a **melhor previsão** de  $y_{T+h}$  a partir de todas as informações disponíveis até  $T$ ?

**Resposta:** a **esperança condicional**

$$\hat{y}_{T+h} = \mathbb{E}(y_{T+h} \mid y_T, y_{T-1}, \dots)$$

Seção 5.2

## **Modelo de Previsão**

## Princípio da Esperança Condicional

---

### Princípio

A **melhor previsão** de  $y_{T+h}$  é a **esperança condicional**:

$$\hat{y}_{T+h} = \mathbb{E}(y_{T+h} \mid \Omega_T)$$

onde  $\Omega_T = \{y_T, y_{T-1}, \dots, u_T, u_{T-1}, \dots\}$  é toda a informação disponível até  $T$ .

**O que é “melhor” aqui?** Melhor no sentido do **erro quadrático médio** (EQM). Defina o EQM de uma previsão  $f_T$  qualquer como:

$$\text{EQM}(f_T) = \mathbb{E}[y_{T+h} - f_T]^2$$

**Intuição:** ao elevar ao quadrado, o EQM penaliza erros grandes mais do que pequenos. Dentre todas as regras de previsão possíveis, queremos aquela que minimiza essa penalidade em média.

## Por que a esperança condicional é ótima?

---

**Afirmção:** para qualquer outra previsão  $f_T$ ,

$$\mathbb{E}[y_{T+h} - f_T]^2 \geq \mathbb{E}[y_{T+h} - \mathbb{E}(y_{T+h} | \Omega_T)]^2$$

**Ideia da prova:** escreva  $f_T = \mathbb{E}(y_{T+h} | \Omega_T) + d_T$ , onde  $d_T$  é o desvio de  $f_T$  em relação à esperança condicional. Expandindo o quadrado:

$$\mathbb{E}[y_{T+h} - f_T]^2 = \underbrace{\mathbb{E}[y_{T+h} - \mathbb{E}(y_{T+h} | \Omega_T)]^2}_{\text{EQM mínimo}} + \underbrace{\mathbb{E}[d_T^2]}_{\geq 0}$$

O segundo termo é sempre  $\geq 0$ , logo usar qualquer  $f_T \neq \mathbb{E}(y_{T+h} | \Omega_T)$  **só piora o EQM.**

**Propiedades:**

- Erros futuros:  $\mathbb{E}(u_{T+j} | \Omega_T) = 0$  para  $j \geq 1$  — substituímos por zero.
- Observações passadas:  $\mathbb{E}(y_{T-k} | \Omega_T) = y_{T-k}$  — já são conhecidas.

## AR(1): previsão 1 período à frente

---

Considere o **AR(1)** estimado:

$$y_t = \mu + \phi y_{t-1} + u_t, \quad u_t \sim \text{RB}(0, \sigma^2), \quad |\phi| < 1.$$

**Passo 1 – Escrever a equação para  $t = T + 1$ :**

$$y_{T+1} = \mu + \phi y_T + u_{T+1}$$

**Passo 2 – Aplicar a esperança condicional em  $\Omega_T$ .** Considere

$$\hat{y}_{T+1} = \mathbb{E}(y_{T+1} \mid \Omega_T):$$

$$\mathbb{E}(y_{T+1} \mid \Omega_T) = \mu + \phi \underbrace{\mathbb{E}(y_T \mid \Omega_T)}_{= y_T} + \underbrace{\mathbb{E}(u_{T+1} \mid \Omega_T)}_{= 0}$$

$$\boxed{\hat{y}_{T+1} = \mu + \phi y_T}$$

## AR(1): previsão 2 períodos à frente

---

Queremos  $\hat{y}_{T+2} = \mathbb{E}(y_{T+2} \mid \Omega_T)$ .

**Passo 1 – Escrever a equação do modelo para  $t = T + 2$ :**

$$y_{T+2} = \mu + \phi y_{T+1} + u_{T+2}$$

**Passo 2 – Aplicar a esperança condicional:**

$$\mathbb{E}(y_{T+2} \mid \Omega_T) = \mu + \phi \underbrace{\mathbb{E}(y_{T+1} \mid \Omega_T)}_{=\hat{y}_{T+1}} + \underbrace{\mathbb{E}(u_{T+2} \mid \Omega_T)}_{=0}$$

**Passo 3 – Substituir  $\hat{y}_{T+1} = \mu + \phi y_T$ :**

$$\hat{y}_{T+2} = \mu + \phi(\mu + \phi y_T) = \mu(1 + \phi) + \phi^2 y_T$$

$$\boxed{\hat{y}_{T+2} = \mu(1 + \phi) + \phi^2 y_T}$$

### Insight

A previsão de  $T + 2$  usa a **previsão** de  $T + 1$  – não um valor observado.

## AR(1): previsão 3 períodos à frente

---

Queremos  $\hat{y}_{T+3} = \mathbb{E}(y_{T+3} \mid \Omega_T)$ .

**Passo 1 – Escrever a equação do modelo para  $t = T + 3$ :**

$$y_{T+3} = \mu + \phi y_{T+2} + u_{T+3}$$

**Passo 2 – Aplicar a esperança condicional:**

$$\mathbb{E}(y_{T+3} \mid \Omega_T) = \mu + \phi \underbrace{\mathbb{E}(y_{T+2} \mid \Omega_T)}_{=\hat{y}_{T+2}} + \underbrace{\mathbb{E}(u_{T+3} \mid \Omega_T)}_{=0}$$

**Passo 3 – Substituir  $\hat{y}_{T+2} = \mu(1 + \phi) + \phi^2 y_T$ :**

$$\begin{aligned}\hat{y}_{T+3} &= \mu + \phi [\mu(1 + \phi) + \phi^2 y_T] = \mu + \mu\phi(1 + \phi) + \phi^3 y_T \\ &= \mu(1 + \phi + \phi^2) + \phi^3 y_T\end{aligned}$$

**Padrão:**  $h = 1: \mu + \phi y_T$ .  $h = 2: \mu(1 + \phi) + \phi^2 y_T$ .  $h = 3: \mu(1 + \phi + \phi^2) + \phi^3 y_T$ . A cada passo, acrescenta-se um termo  $\mu\phi^{h-1}$  na soma e aumenta-se o expoente de  $y_T$ .

## AR(1): generalização para $h$ períodos à frente

---

Padrão geral – previsão  $h$  períodos à frente:

$$\hat{y}_{T+h} = \mu(1 + \phi + \phi^2 + \dots + \phi^{h-1}) + \phi^h y_T = \mu \cdot \frac{1 - \phi^h}{1 - \phi} + \phi^h y_T$$

Quando  $h \rightarrow \infty$ ?

Sob  $|\phi| < 1$ :

- $\phi^h \rightarrow 0 \Rightarrow$  o efeito de  $y_T$  **some** — o passado perde relevância.
- $\mu \cdot \frac{1 - \phi^h}{1 - \phi} \rightarrow \frac{\mu}{1 - \phi} \Rightarrow$  a soma da PG converge.

**Conclusão:**

$$\hat{y}_{T+h} \xrightarrow{h \rightarrow \infty} \frac{\mu}{1 - \phi} = \mathbb{E}(y_t) = \bar{y}$$

A previsão **converge para a média incondicional** do processo.

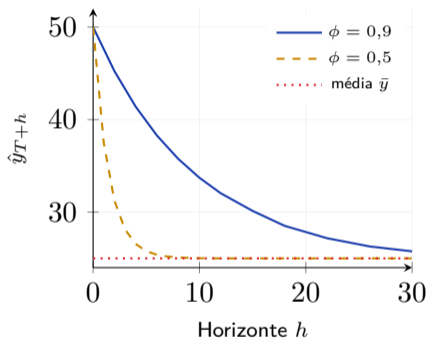
## Intuição:

---

### Imagine o processo como uma borracha:

- Cada período,  $y_t$  é “puxado” para a média  $\bar{y} = \mu/(1 - \phi)$ .
- O coeficiente  $\phi$  controla a **força do elástico**.
- $\phi = 0,9$ : elástico fraco – retorno lento para a média.
- $\phi = 0,5$ : elástico forte – retorno rápido.

**Para a previsão:** quanto mais longe no futuro, mais incerto. A única coisa “certa” no longo prazo é a média.



## Exemplo numérico: previsão com AR(1)

**Modelo estimado:**  $\hat{y}_t = 10 + 0,6 y_{t-1}$ , média de longo prazo =  $10/(1 - 0,6) = 25$ .

$h$	Cálculo	$\hat{y}_{T+h}$	Distância à média
1	$10 + 0,6 \times 50$	40,0	15,0
2	$10 + 0,6 \times 40$	34,0	9,0
3	$10 + 0,6 \times 34$	30,4	5,4
4	$10 + 0,6 \times 30,4$	28,2	3,2
5	$10 + 0,6 \times 28,2$	26,9	1,9
$\infty$	$10/(1 - 0,6)$	25,0	0,0

**Lendo a tabela:** cada período, a série perde cerca de 40% da distância à média (porque  $\phi = 0,6$ ). Em 5 períodos, a previsão já está a menos de 2 unidades da média de longo prazo.

### Definição

O erro de previsão  $h$  períodos à frente é a diferença entre o valor realizado e o previsto:

$$e_{T+h} \equiv y_{T+h} - \hat{y}_{T+h}$$

**Exemplo: AR(1)**  $y_t = \mu + \phi y_{t-1} + u_t$

Para  $h = 1$ :

$$e_{T+1} = \underbrace{(\mu + \phi y_T + u_{T+1})}_{y_{T+1}} - \underbrace{(\mu + \phi y_T)}_{\hat{y}_{T+1}} = u_{T+1}$$

## Erro de previsão do AR(1)

---

Para  $h = 2$ :

$$\begin{aligned}e_{T+2} &= y_{T+2} - \hat{y}_{T+2} \\&= (\mu + \phi y_{T+1} + u_{T+2}) - (\mu + \phi \hat{y}_{T+1}) \\&= \phi \underbrace{(y_{T+1} - \hat{y}_{T+1})}_{= e_{T+1} = u_{T+1}} + u_{T+2} \\&= \phi u_{T+1} + u_{T+2}\end{aligned}$$

## Erro de previsão do AR(1)

---

Para  $h = 3$ :

$$\begin{aligned}e_{T+3} &= y_{T+3} - \hat{y}_{T+3} \\ &= (\mu + \phi y_{T+2} + u_{T+3}) - (\mu + \phi \hat{y}_{T+2}) \\ &= \phi \underbrace{(y_{T+2} - \hat{y}_{T+2})}_{= e_{T+2} = \phi u_{T+1} + u_{T+2}} + u_{T+3} \\ &= \phi^2 u_{T+1} + \phi u_{T+2} + u_{T+3}\end{aligned}$$

**Padrão geral – AR(1):**

$$e_{T+h} = u_{T+h} + \phi u_{T+h-1} + \phi^2 u_{T+h-2} + \dots + \phi^{h-1} u_{T+1}$$

O erro  $h$  períodos à frente é uma **média ponderada dos choques futuros**  $u_{T+1}, \dots, u_{T+h}$  — todos imprevisíveis hoje. Quanto maior  $h$ , mais choques se acumulam e maior a incerteza.

## Propriedades do erro de previsão do AR(1)

---

### 1. Esperança (viés):

$$\mathbb{E}(e_{T+h}) = \sum_{j=0}^{h-1} \phi^j \underbrace{\mathbb{E}(u_{T+h-j})}_{=0} = 0$$

⇒ **Previsão não viesada.** Em média, não erramos sistematicamente para cima nem para baixo.

### 2. Variância (precisão):

$$\text{Var}(e_{T+h}) = \sigma^2 \sum_{j=0}^{h-1} \phi^{2j} = \sigma^2 \frac{1 - \phi^{2h}}{1 - \phi^2}$$

### Observações:

- $h = 1$ :  $\text{Var}(e_{T+1}) = \sigma^2$ .
- $h$  cresce: variância **umenta** monotonicamente.
- $h \rightarrow \infty$ :  $\text{Var}(e_{T+h}) \rightarrow \sigma^2 / (1 - \phi^2) = \text{Var}(y_t)$  (variância incondicional).

**Interpretação:** previsões de curto prazo aproveitam a informação recente e têm menor incerteza. No longo prazo, perdemos tudo o que sabemos e a incerteza iguala a variância incondicional da série.

## Intervalo de confiança para a previsão

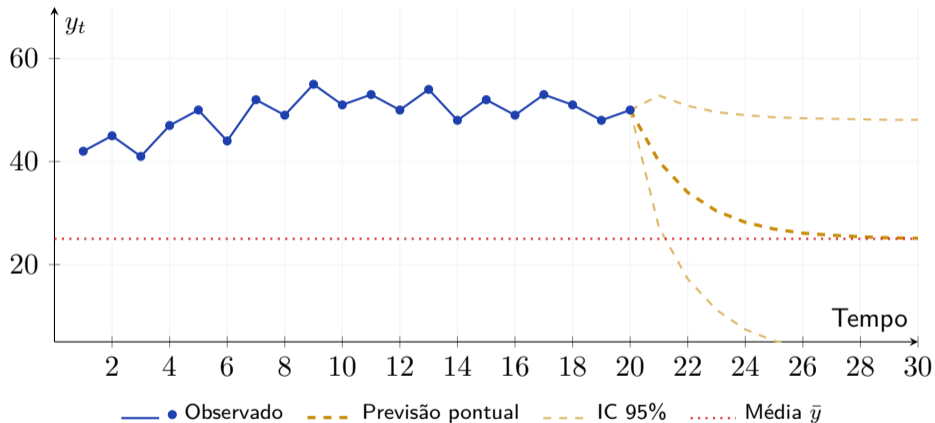
---

Intervalo de confiança de 95% para  $\hat{y}_{T+h}$

$$\hat{y}_{T+h} \pm 1,96 \sqrt{\text{Var}(e_{T+h})}$$

O “cone” de previsão: o intervalo fica cada vez mais largo. Graficamente, isso forma um cone que se abre à medida que avançamos no tempo.

## Visualizando: o “cone” da previsão



**Previsão pontual** converge para  $\bar{y}$  (linha vermelha). O **cone** alarga conforme o horizonte cresce.

Seção 5.3

## **Algoritmo Geral para Qualquer ARIMA**

## Por que precisamos de um algoritmo geral?

---

Até agora trabalhamos com o AR(1). Na prática, o modelo ajustado pode ser:

- ARMA(2,1), ARIMA(1,1,2), SARIMA(1, 1, 0)(2, 0, 2)<sub>12</sub>, ...

A ideia é sempre a mesma – esperança condicional – mas a álgebra fica mais complexa.

### Algoritmo de 4 passos (Hyndman & Athanasopoulos, 2018)

1. **Reescrever** a equação ARIMA com  $y_t$  isolado no lado esquerdo.
2. **Substituir**  $t$  por  $T + h$ .
3. **Substituir** no lado direito:
  - Observações futuras  $y_{T+j}$  ( $j \geq 1$ )  $\rightarrow$  previsões  $\hat{y}_{T+j}$  já calculadas.
  - Erros futuros  $u_{T+j}$  ( $j \geq 1$ )  $\rightarrow$  zero (esperança nula).
  - Erros passados  $u_{T-j}$  ( $j \geq 0$ )  $\rightarrow$  resíduos estimados  $\hat{u}_{T-j}$ .
4. **Iterar:** começar com  $h = 1$ , depois  $h = 2$ ,  $h = 3$ , ...

## ARMA(1,1): previsão para $T + 1$ e $T + 2$

**Modelo:**  $y_t = \mu + \phi y_{t-1} + u_t + \theta u_{t-1}$

**Para  $t = T + 1$ :**

$$y_{T+1} = \mu + \phi y_T + u_{T+1} + \theta u_T$$

$$\hat{y}_{T+1} = \mu + \phi y_T + \underbrace{\mathbb{E}(u_{T+1} | \Omega_T)}_{=0} + \theta \underbrace{\mathbb{E}(u_T | \Omega_T)}_{=\hat{u}_T} \Rightarrow \boxed{\hat{y}_{T+1} = \mu + \phi y_T + \theta \hat{u}_T}$$

**Para  $t = T + 2$ :**

$$y_{T+2} = \mu + \phi y_{T+1} + u_{T+2} + \theta u_{T+1}$$

$$\hat{y}_{T+2} = \mu + \phi \underbrace{\mathbb{E}(y_{T+1} | \Omega_T)}_{=\hat{y}_{T+1}} + \underbrace{\mathbb{E}(u_{T+2} | \Omega_T)}_{=0} + \theta \underbrace{\mathbb{E}(u_{T+1} | \Omega_T)}_{=0} \Rightarrow \boxed{\hat{y}_{T+2} = \mu + \phi \hat{y}_{T+1}}$$

**Diferença:** em  $T + 1$ , o resíduo  $\hat{u}_T$  é *passado* – conhecido. Em  $T + 2$ ,  $u_{T+1}$  é *futuro* – esperança zero. O componente MA só importa no **primeiro passo**.

## ARMA(1,1): previsão para $T + 3$ e padrão geral

---

Para  $t = T + 3$ :

$$y_{T+3} = \mu + \phi y_{T+2} + u_{T+3} + \theta u_{T+2}$$
$$\hat{y}_{T+3} = \mu + \phi \underbrace{\mathbb{E}(y_{T+2} \mid \Omega_T)}_{=\hat{y}_{T+2}} + \underbrace{\mathbb{E}(u_{T+3} \mid \Omega_T)}_{=0} + \theta \underbrace{\mathbb{E}(u_{T+2} \mid \Omega_T)}_{=0} \Rightarrow \boxed{\hat{y}_{T+3} = \mu + \phi \hat{y}_{T+2}}$$

Padrão geral – para  $h \geq 2$ :

$$\boxed{\hat{y}_{T+h} = \mu + \phi \hat{y}_{T+h-1}}$$

A partir de  $h = 2$ , o ARMA(1,1) se comporta **exatamente como um AR(1) puro**.

### Resumo das regras de previsão do ARMA(1,1)

---

Horizonte	Fórmula
$h = 1$	$\hat{y}_{T+1} = \mu + \phi y_T + \theta \hat{u}_T$ (resíduo entra)
$h \geq 2$	$\hat{y}_{T+h} = \mu + \phi \hat{y}_{T+h-1}$ (MA some)
$h \rightarrow \infty$	$\hat{y}_{T+h} \rightarrow \mu / (1 - \phi) = \bar{y}$

---

Seção 5.4

## **Previsão Dentro da Amostra**

## Dentro vs. fora da amostra

---

### Previsão **dentro** da amostra

Usa dados que **já conhecemos** para avaliar a qualidade do modelo.

**Como:** “finja” que não conhece os últimos  $H$  pontos; ajuste o modelo com o restante; compare previsão com o valor real.

**Objetivo:** validar o modelo antes de usá-lo para o futuro.

### Previsão **fora** da amostra

Prever valores **genuinamente futuros**.

**Como:** ajusta-se o modelo em *toda* a amostra disponível; geram-se projeções para períodos que ainda não ocorreram.

**Objetivo:** tomar decisões sobre o futuro.

## Procedimento: previsão dentro da amostra

---

**Cenário:**  $T$  observações da série.

### Procedimento – 4 passos

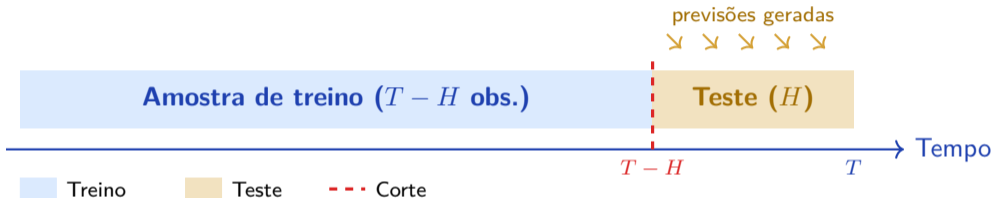
1. **Reservar** as últimas  $H$  observações como **conjunto de teste**.
2. **Ajustar** o modelo usando **apenas** as primeiras  $T - H$  observações.
3. **Gerar** previsões  $\hat{y}_{T-H+1}, \hat{y}_{T-H+2}, \dots, \hat{y}_T$ .
4. **Comparar** com os valores reais  $y_{T-H+1}, \dots, y_T$  pelas medidas de erro.

### Exemplo: IPCA acumulado em 12 meses

Série: jan/2000 a jan/2024 ( $T = 289$  meses). Reservar os últimos 12 meses (fev/2023 a jan/2024) como conjunto de teste. Ajustar o SARIMA em dados até jan/2023. Gerar previsões para fev/2023–jan/2024 e comparar com os valores reais já realizados, calculando o Erro Quadrático Médio, por exemplo.

## Visualizando a divisão treino/teste

---



**Escolha de  $H$ :** sem regra rígida. Prática comum:  $H = 12$  (mensal,  $\approx 1$  ano) ou  $H = 4$  (trimestral).

Seção 5.5

## **Previsão Fora da Amostra e Medidas de Desempenho**

## Previsão fora da amostra

---

Para prever valores genuinamente futuros (que ainda não se realizaram), aplicamos o algoritmo de 4 passos visto antes – agora com o modelo ajustado em **toda** a amostra.

### Regras de substituição:

Termo	Posição	Substituir por
$y_{T+j}, j \geq 1$	futuro (lado direito)	previsão $\hat{y}_{T+j}$
$u_{T+j}, j \geq 1$	futuro (lado direito)	zero
$y_{T-j}, j \geq 0$	passado	valor observado $y_{T-j}$
$u_{T-j}, j \geq 0$	passado	resíduo $\hat{u}_{T-j}$

## Medidas de desempenho: definição

---

Erro de previsão  $h$  passos à frente:  $e_h = y_{T+h} - \hat{y}_{T+h}$ , calculado para  $h = 1, \dots, H$ .

### Três medidas principais

$$\mathbf{MAE} = \frac{1}{H} \sum_{h=1}^H |e_h| \quad (\text{Mean Absolute Error})$$

$$\mathbf{MAPE} = \frac{100}{H} \sum_{h=1}^H \left| \frac{e_h}{y_{T+h}} \right| \quad (\text{Mean Absolute Percentage Error})$$

$$\mathbf{RMSE} = \sqrt{\frac{1}{H} \sum_{h=1}^H e_h^2} \quad (\text{Root Mean Squared Error})$$

**Regra: menor é melhor.** Para comparar modelos, prefira o de menor RMSE.

## Qual medida usar? MAE vs. RMSE vs. MAPE

---

### MAE

**Intuitivo:** erro médio em unidades da série.

- + Robusto a outliers.
- + Fácil de comunicar.
- Não penaliza erros grandes.

### RMSE

**Penaliza erros grandes** (eleva ao quadrado).

- + O mais usado na literatura.
- + Penaliza erros maiores.
- Sensível a outliers.

### MAPE

**Erro em %:** útil para comparar séries de escalas diferentes.

- + Interpretação imediata.
- Explode quando  $y \approx 0$ .

### Exemplo: IPCA vs. PIB

Comparar MAE do IPCA ( $\approx 0,01$  p.p.) com MAE do PIB em bilhões ( $\approx 50$ ) não faz sentido – unidades diferentes. O **MAPE** resolve: ambos expressos em %.

No entanto, se o IPCA passar fica próximo de zero, o MAPE explode. Aí use MAE ou RMSE.

## Exemplo numérico: comparando dois modelos para o IPCA

**Cenário:** previsões de 5 meses do IPCA mensal.

$h$	<b>Modelo A</b> (ARIMA(1,1,0))			<b>Modelo B</b> (SARIMA)		
	$y_h$	$\hat{y}_h^A$	$e_h^A$	$y_h$	$\hat{y}_h^B$	$e_h^B$
1	0,50	0,40	+0,10	0,50	0,30	+0,20
2	0,30	0,40	-0,10	0,30	0,40	-0,10
3	0,60	0,50	+0,10	0,60	0,60	+0,00
4	0,40	0,50	-0,10	0,40	0,45	-0,05
5	0,50	0,40	+0,10	0,50	0,40	+0,10

**Modelo A:** MAE = 0,100, RMSE = 0,100.

**Modelo B:** MAE =  $\frac{0,20+0,10+0+0,05+0,10}{5} = 0,090$ , RMSE =  $\sqrt{0,0085} \approx 0,092$ .

**Conclusão:** Modelo B tem MAE e RMSE menores – é preferido. O modelo com sazonalidade explícita captura melhor os dados mensais.

## Rolling window: motivação

---

Suponha que você reserva os últimos 12 meses como teste. O desempenho do modelo pode ser muito bom (ou muito ruim) por acaso – pode ser que o período teste escolhido seja atípico.

### Solução: rolling forecast (janela deslizante)

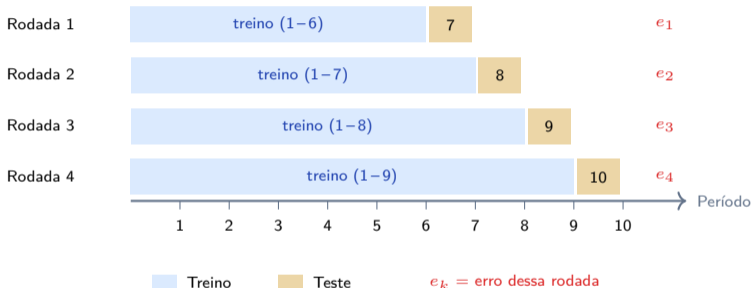
Em vez de um único corte, fazemos o processo **várias vezes**, avançando um período por vez:

1. Estima o modelo nos primeiros  $m$  períodos. Prevê o período  $m + 1$ . Calcula o erro  $e_1$ .
2. Estima o modelo nos primeiros  $m + 1$  períodos. Prevê o período  $m + 2$ . Calcula o erro  $e_2$ .
3. Repete até o final da amostra, acumulando  $K$  erros.

É como simular, retrospectivamente, todas as previsões que você teria feito ao longo do tempo – uma por mês, com a informação disponível naquele momento. O RMSE final resume o desempenho **médio** do modelo ao longo de todo o período.

## Rolling window: o diagrama

Cada linha abaixo representa uma **rodada** de estimação + previsão. A janela de treino cresce um período a cada rodada; o bloco de teste é sempre **um** período à frente.



Ao final das  $K$  rodadas, resumimos a performance num único número:

$$\text{RMSE rolling} = \sqrt{\frac{1}{K} \sum_{k=1}^K e_k^2}$$

## Boas práticas e limitações

---

### Sempre faça:

- **Inspecione o gráfico** da previsão vs. valores reais – o olho humano detecta padrões que as métricas não captam.
- **Reporte mais de uma métrica** (MAE + RMSE).
- **Compare com um benchmark simples**: a média histórica ou a previsão ingênua ( $\hat{y}_{T+1} = y_T$ ) como linha de base.

### Cuidados:

- **Quebras estruturais**: crises (pandemia, Plano Real) fazem a estrutura do modelo mudar. Previsões baseadas em períodos antes da quebra serão ruins.
- **MAPE** explode quando  $y \approx 0$  (ex: variação do PIB em recessão).
- **Horizonte longo**: a incerteza cresce rápido. Uma previsão de 1–3 meses é muito mais confiável que de 2 anos.

**Regra da parcimônia**: um modelo mais simples que funciona bem é preferível a um modelo complexo que funciona marginalmente melhor.

# Exercícios

## Enunciado

Considere  $Y_t = 10 + 0,6 Y_{t-1} + e_t$ , com  $e_t$  ruído branco com  $\mathbb{E}(e_t) = 0$  e  $\mathbb{E}(e_t e_s) = 0$  para  $t \neq s$ .

Sabendo que  $Y_0 = 0$ , obtenha  $\mathbb{E}(Y_2)$ .

**Dica:** Aplique a lei das expectativas iteradas.  $\mathbb{E}(Y_1)$  depende de  $Y_0 = 0$ ; depois  $\mathbb{E}(Y_2)$  depende de  $\mathbb{E}(Y_1)$ .

**Passo 1 – Calcular  $\mathbb{E}(Y_1)$  a partir de  $Y_0 = 0$ :**

$$\mathbb{E}(Y_1) = 10 + 0,6 \cdot 0 + \underbrace{\mathbb{E}(e_1)}_{=0} = 10$$

**Passo 2 – Calcular  $\mathbb{E}(Y_2)$ :**

$$\mathbb{E}(Y_2) = 10 + 0,6 \cdot \mathbb{E}(Y_1) + \underbrace{\mathbb{E}(e_2)}_{=0} = 10 + 0,6 \times 10 = \boxed{16}$$

**Média de longo-prazo do processo:**  $\mu/(1 - \phi) = 10/(1 - 0,6) = 25$ . Ainda estamos longe (partimos de 0).

## Enunciado

Suponha que  $y_t = 35 + \frac{3}{5}y_{t-1} + u_t$ , com  $u_t \sim N(0, 2)$  i.i.d.

Você é informado de que  $y_2 = 50$ .

- (a) Determine a melhor previsão para  $y_4$ .
- (b) Qual é a variância do erro de previsão para  $y_4$ ?

(a) Previsão pontual:

Passo 1:

$$\hat{y}_3 = 35 + \frac{3}{5} \cdot 50 = 35 + 30 = 65$$

Passo 2:

$$\hat{y}_4 = 35 + \frac{3}{5} \cdot 65 = 35 + 39 = \boxed{74}$$

(b) Temos um AR(1), e vimos em aula que, para esse caso, a variância do erro de previsão 2 passos à frente (de  $y_2$  para  $y_4$ ,  $h = 2$ ):

$$\text{Var}(e_4) = \sigma^2(1 + \phi^2) = 2 \cdot \left(1 + \left(\frac{3}{5}\right)^2\right) = 2 \cdot \left(1 + \frac{9}{25}\right) = 2 \cdot \frac{34}{25} = \frac{68}{25} = \boxed{2,72}$$

## Enunciado

Considere o AR(1):  $Y_t = a + bY_{t-1} + u_t$ , com  $|b| < 1$  e  $u_t$  i.i.d. com média zero e variância  $\sigma^2$ .

A previsão  $n$  passos à frente para  $Y$  converge para:

- (a)  $a$
- (b) a média de  $u_t$
- (c)  $a/(1 - b)$
- (d)  $\mathbb{E}(Y_t)$
- (e)  $\infty$

Respostas corretas: (c) e (d) (são a mesma coisa).

**Justificativa.** Vimos nessa aula que para o processo AR(1), temos:

$$\hat{Y}_{T+n} = a \cdot \frac{1 - b^n}{1 - b} + b^n Y_T$$

Sob  $|b| < 1$ , quando  $n \rightarrow \infty$ :

- $b^n \rightarrow 0 \Rightarrow$  o efeito de  $Y_T$  desaparece.
- $\frac{1 - b^n}{1 - b} \rightarrow \frac{1}{1 - b}$

$$\hat{Y}_{T+n} \longrightarrow \frac{a}{1 - b} = \mathbb{E}(Y_t)$$

**Por que não (a)?** Sabemos que  $\mathbb{E}(Y_t) = \mathbb{E}(Y_{t-1}) = \bar{y}$ , então, tomando o operador esperança no processo,  $\bar{y} = a + b\bar{y} \iff \bar{y} = a/(1 - b) \neq a$ .

**Por que não (b)?** A média de  $u_t$  é zero.

**Por que não (e)?**  $|b| < 1$  garante estabilidade – sem divergência.

## Enunciado

Uma série  $Y$  mensal (jan/2015 a fev/2022) é descrita por:

Variável	Estimativa	Erro-padrão
Constante	7,37	6,65
$Y(-1)$	0,77	0,07

A sazonalidade é determinística, com coeficiente para março = 73,96.

Sabendo que  $Y_{\text{fev}/22} = 26,4$ , obtenha a previsão para março/22.

**Passo 1 – Escrever o modelo (AR(1) com sazonalidade aditiva):**

$$\hat{Y}_t = 7,37 + 0,77 Y_{t-1} + \delta_{\text{mês}}$$

**Passo 2 – Substituir para março/22:**

$$\hat{Y}_{\text{mar}/22} = 7,37 + 0,77 \times 26,4 + 73,96 = 7,37 + 20,33 + 73,96 = \boxed{101,66}$$

**Nota:** o coeficiente sazonal de março (73,96) entra como um *shift* aditivo da previsão – ele captura o efeito sistemático de estar em março.

## Exercício 5

---

### Enunciado

Considere o ARMA(1,1) estimado:

$$y_t = 2 + 0,7 y_{t-1} + u_t - 0,4 u_{t-1}$$

Sabe-se que  $y_T = 15$  e que o último resíduo estimado foi  $\hat{u}_T = -2$ .

- (a) Calcule  $\hat{y}_{T+1}$  e  $\hat{y}_{T+2}$ .
- (b) Para qual valor converge  $\hat{y}_{T+h}$  quando  $h \rightarrow \infty$ ?
- (c) Qual é a variância do erro de previsão 1 passo à frente? (use  $\sigma^2 = 4$ )

## Exercício 5 – Gabarito

---

(a) Previsões:

$$\hat{y}_{T+1} = 2 + 0,7 \times 15 + (-0,4) \times (-2) = 2 + 10,5 + 0,8 = \boxed{13,3}$$

$$\hat{y}_{T+2} = 2 + 0,7 \times 13,3 + \underbrace{(-0,4) \times 0}_{u_{T+1}=0} = 2 + 9,31 = \boxed{11,31}$$

(b) **Convergência:** para o ARMA(1,1) com  $|\phi| < 1$ , a previsão converge para a média incondicional:

$$\bar{y} = \frac{\mu}{1 - \phi} = \frac{2}{1 - 0,7} = \frac{2}{0,3} \approx \boxed{6,67}$$

## Exercício 5 – Gabarito (c): derivando o erro para $h = 1$

---

**Passo 1 – Escrever o valor realizado  $y_{T+1}$ :**

$$y_{T+1} = 2 + 0,7 y_T + u_{T+1} - 0,4 u_T$$

**Passo 2 – Escrever a previsão  $\hat{y}_{T+1}$  (que já calculamos no item a):**

$$\hat{y}_{T+1} = 2 + 0,7 y_T - 0,4 \hat{u}_T$$

**Passo 3 – Subtrair:**

$$\begin{aligned} e_{T+1} &= y_{T+1} - \hat{y}_{T+1} \\ &= (2 + 0,7 y_T + u_{T+1} - 0,4 u_T) - (2 + 0,7 y_T - 0,4 \hat{u}_T) \\ &= u_{T+1} - 0,4 u_T + 0,4 \hat{u}_T \\ &= u_{T+1} - 0,4 \underbrace{(u_T - \hat{u}_T)}_{\approx 0} \end{aligned}$$

O resíduo  $\hat{u}_T$  é a estimativa de  $u_T$  obtida na etapa de diagnóstico. Assumimos que a estimativa é precisa:  $\hat{u}_T \approx u_T$ . O termo  $0,4(u_T - \hat{u}_T)$  cancela e sobra apenas:  $e_{T+1} = u_{T+1}$

## Exercício 5 – Gabarito (c): por que $\text{Var}(e_{T+1}) = \sigma^2$

---

Acabamos de mostrar que  $e_{T+1} = u_{T+1}$ . Portanto:

$$\text{Var}(e_{T+1}) = \text{Var}(u_{T+1}) = \sigma^2 = \boxed{4}$$

### Por que $u_T$ desapareceu?

Em  $h = 1$ , todos os termos do modelo que envolvem o **passado** ( $y_T, u_T$ ) são **conhecidos** no momento da previsão – eles entram tanto em  $y_{T+1}$  quanto em  $\hat{y}_{T+1}$  e se cancelam na subtração. A única peça que **não conhecemos** é o choque futuro  $u_{T+1}$ .

### Vale para qualquer modelo ARIMA

Independente de o modelo ser AR(1), MA(1), ARMA(1,1) ou ARIMA(2,1,1):

$$e_{T+1} = u_{T+1} \implies \text{Var}(e_{T+1}) = \sigma^2$$

O componente MA afeta a *previsão pontual* via  $\hat{u}_T$  (como vimos no item a), mas não reduz a incerteza de  $h = 1$  abaixo de  $\sigma^2$  – esse é o **piso da incerteza**.

## Exercício 5 – Gabarito (c): extrapolando para $h = 2$ ?

---

Para contrastar, veja o que acontece em  $h = 2$ .

**Valor realizado:**

$$y_{T+2} = 2 + 0,7 y_{T+1} + u_{T+2} - 0,4 u_{T+1}$$

**Previsão (item a):**

$$\hat{y}_{T+2} = 2 + 0,7 \hat{y}_{T+1}$$

**Erro:**

$$\begin{aligned} e_{T+2} &= y_{T+2} - \hat{y}_{T+2} \\ &= 0,7 \underbrace{(y_{T+1} - \hat{y}_{T+1})}_{= e_{T+1} = u_{T+1}} + u_{T+2} - 0,4 u_{T+1} \\ &= 0,7 u_{T+1} + u_{T+2} - 0,4 u_{T+1} \\ &= \underbrace{(0,7 - 0,4)}_{= 0,3} u_{T+1} + u_{T+2} \end{aligned}$$

**Variância do erro para  $h = 2$ :**

$$\text{Var}(e_{T+2}) = (0,3)^2 \sigma^2 + \sigma^2 = (1 + 0,09) \sigma^2 = 1,09 \sigma^2$$

Com  $\sigma^2 = 4$ :  $\text{Var}(e_{T+2}) = 4,36 > 4 = \text{Var}(e_{T+1})$ . A variância **crece** com o horizonte.

## Enunciado

Dois modelos foram ajustados para a taxa Selic mensal. As previsões e valores reais para os últimos 4 meses foram:

Mês	$y_h$	$\hat{y}_h^A$	$y_h$	$\hat{y}_h^B$
1	10,5	10,0	10,5	10,8
2	11,0	10,5	11,0	10,6
3	11,25	11,5	11,25	11,0
4	11,75	11,0	11,75	11,5

- Calcule o MAE de cada modelo.
- Calcule o RMSE de cada modelo.
- Qual modelo você escolheria? Justifique.

## Exercício 6 – Gabarito

---

**Erros:**  $e_h^A$ : 0,5, 0,5, -0,25, 0,75.     $e_h^B$ : -0,3, 0,4, 0,25, 0,25.

**(a) MAE:**

$$\text{MAE}^A = \frac{0,5+0,5+0,25+0,75}{4} = \frac{2,0}{4} = 0,500$$

$$\text{MAE}^B = \frac{0,3+0,4+0,25+0,25}{4} = \frac{1,2}{4} = 0,300$$

**(b) RMSE:**

$$\text{RMSE}^A = \sqrt{\frac{0,25+0,25+0,0625+0,5625}{4}} = \sqrt{\frac{1,125}{4}} \approx 0,530$$

$$\text{RMSE}^B = \sqrt{\frac{0,09+0,16+0,0625+0,0625}{4}} = \sqrt{\frac{0,375}{4}} \approx 0,306$$

**(c) Modelo B:** tanto MAE quanto RMSE são menores. Modelo B é preferido. Observe que o erro maior de A no mês 4 ( $e = 0,75$ ) é penalizado pelo RMSE.

## Resumo do Capítulo 5

---

### Conceitos fundamentais

- Previsão = esperança condicional
- AR(1):  $\hat{y}_{T+h} = \mu \frac{1-\phi^h}{1-\phi} + \phi^h y_T$
- Iteração: previsão usa previsões anteriores
- Convergência para  $\mu/(1-\phi)$  quando  $h \rightarrow \infty$
- Erro não viesado; variância cresce com  $h$
- IC fica mais largo (“cone”)
- MA some após  $h = 1$  nos erros futuros

### Algoritmo geral (qualquer ARIMA):

1. Isolar  $y_t$  no lado esquerdo
2. Substituir  $t \rightarrow T + h$
3. Futuro  $\rightarrow$  previsão / zero; passado  $\rightarrow$  real / resíduo
4. Iterar de  $h = 1$  em diante

### Avaliação:

- Dentro  $\times$  fora da amostra
- MAE, RMSE, MAPE
- Rolling window para robustez

**Resumo:** para  $h = 1$ , o erro de previsão  $e_{T+1}$  é  $u_{T+1}$  (não há como melhorar). Para  $h$  grande, a previsão converge para a média incondicional – o passado distante perde efeito. Tudo no meio é uma questão de **quão persistente** é a série.

## Dúvidas?

Próxima aula: Defasagens Distribuídas e Cointegração (Cap. 6)