

# Métodos de Previsão

## Parte II: Métodos Estocásticos

# Contexto

Os métodos estocásticos baseiam-se, entre outros, em

- Modelos autoregressivos (AR), onde a variável  $y_t$  é explicada pelo seu passado:  $y_t = f(y_{t-1}, y_{t-2}, \dots)$
- Modelos de médias móveis (MA) onde  $y_t$  é explicado por um ruído branco  $\varepsilon_t$  e o seu passado:  $y_t = f(\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots)$
- Modelos ARMA, que combinam AR e MA:  $y_t = f(y_{t-1}, y_{t-2}, \varepsilon_t, \varepsilon_{t-1}, \dots)$
- Modelos autorregressivos de defasamentos distribuídos (ADL), onde o presente e, possivelmente, o passado de outras variáveis explicam a variável de interesse:  $y_t = f(y_{t-1}, y_{t-2}, \dots, x_t, x_{t-1}, x_{t-2}, \dots)$
- Modelos VAR, onde se tem um sistema de equações de tipo ADL:

$$\begin{cases} y_t = f(y_{t-1}, y_{t-2}, \dots, x_{t-1}, x_{t-2}, \dots) \\ x_t = f(y_{t-1}, y_{t-2}, \dots, x_{t-1}, x_{t-2}, \dots) \end{cases}$$

# Contexto

Algumas questões adicionais levantadas relativamente aos modelos de regressão standard:

- Inclusão de variáveis desfasadas/lags e/ou às diferenças
- Existência de correlação no tempo entre observações: autocorrelação
- O objectivo principal é a previsão (e não o apuramento de efeitos casuais):
  - A interpretação dos parâmetros estimados não é essencial
  - Promove-se a validação externa do modelo: verifica-se se o modelo estimado com dados do passado, descreve o comportamento num futuro próximo

# Tipos de variáveis

## Lags, primeiras diferenças, logs e taxas de crescimento

- Lag 1:  $y_{t-1}$ ; lag  $j$ :  $y_{t-j}$
- Primeira diferença:  $\Delta y_t = y_t - y_{t-1}$ 
  - Uma aproximação à taxa de crescimento de  $y_t$  é dada por  $\Delta \ln(y_t) = \ln(y_t) - \ln(y_{t-1})$ , quando o valor da taxa é pequeno

### Exemplo:

**TABLE 14.1** Inflation in the United States in 2004 and the First Quarter of 2005

Quarter	U.S. CPI	Rate of Inflation at an Annual Rate ( $Inf_t$ )	First Lag ( $Inf_{t-1}$ )	Change in Inflation ( $\Delta Inf_t$ )
2004:I	186.57	3.8	0.9	2.9
2004:II	188.60	4.4	3.8	0.6
2004:III	189.37	1.6	4.4	-2.8
2004:IV	191.03	3.5	1.6	1.9
2005:I	192.17	2.4	3.5	-1.1

# Autocorrelação

Autocorrelação ou *serial correlation* é a correlação existente entre duas realizações de um determinado processo, desfasadas de  $h$  períodos

Função autocovariância (ACOVF) e função de autocorrelação (ACF) define-se como

$$\gamma_h = Cov(y_t, y_{t-h})$$

$$\rho_h = Cor(y_t, y_{t-h}) = \frac{Cov(y_t, y_{t-h})}{\sqrt{Var(y_t)Var(y_{t-h})}} = \frac{\gamma_h}{\gamma_0},$$

onde  $h = 0, 1, \dots$ ,  $\gamma_0 = Var(y_t)$ ,  $\rho_0 = 1$ , e  $|\rho_h| \leq 1$

# Autocorrelação parcial

A função de autocorrelação parcial (PACF) mede a correlação entre  $y_t$  e  $y_{t-h}$ , excluídos os efeitos das correlações intermédias com  $y_{t-1}, y_{t-2}, \dots, y_{t-h-1}$ .

O coeficiente de autocorrelação  $\phi_{hh}$  é o último coeficiente da regressão que relaciona linearmente  $y_t$  com os seus mais recentes  $h$  valores

$$y_{t-h} = \phi_{k1}y_{t+h-1} + \phi_{k2}y_{t+h-2} + \phi_{hh}y_t + e_{t-h}$$

com  $\phi_{11} = \rho_1$

# Autocorrelação e autocorrelação parcial na amostra

$$\hat{\rho}_h = \frac{\hat{\gamma}_h}{\hat{\gamma}_0} = \frac{\frac{1}{T} \sum_{t=h+1}^T (y_t - \bar{y})(y_{t-h} - \bar{y})}{\frac{1}{T} \sum_{t=1}^T (y_t - \bar{y})^2} = \frac{\sum_{t=h+1}^T (y_t - \bar{y})(y_{t-h} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2}$$

onde a média  $\bar{y}$  é calculada para as observações  $h + 1, \dots, T$

$$\hat{\phi}_{hh} = \frac{\hat{\rho}_h - \sum_{j=1}^{h-1} \hat{\phi}_{h-1,j} \hat{\rho}_{h-j}}{1 - \sum_{j=1}^{h-1} \hat{\phi}_{h-1,j} \hat{\rho}_j}$$

com  $\hat{\phi}_{11} = \hat{\rho}_1$  (inicialização)

# Estacionaridade

**Estacionaridade:** um processo é dito estacionário se a sua distribuição não se altera ao longo do tempo, isto é, não depende de  $t$

**Estacionaridade fraca:** um processo é dito fracamente estacionário quando a sua média e variância são constantes e a covariância e correlação entre  $y_t$  e  $y_{t-h}$  dependem apenas do intervalo de tempo  $h$ :

$$E(y_t) = \mu$$

$$V(y_t) = E(y_t - \mu)^2 = \sigma^2$$

$Cov(y_t, y_{t-h})$  pode depender de  $h$  mas não de  $t$



# Ruido branco

## “White noise”

É um processo cujas observações não são correlacionadas:

$$y_t = \varepsilon_t$$

onde

$$E(\varepsilon_t) = \mu_\varepsilon$$

$$V(\varepsilon_t) = \sigma_\varepsilon^2$$

$$\text{Cov}(\varepsilon_t, \varepsilon_{t-h}) = 0 \text{ para } h \neq 0$$

- A ACF e a PACF são nulas para  $h \neq 0$

# Ruido branco

## Testar o ruido branco

- Em termos gerais, verifica-se se a ACF e a PACF são aproximadamente zero
- Teste Q, versão de Ljung-Box (1978), baseada na original de Box-Pierce (1970)

- $H_0: \rho_1 = \rho_2 = \dots = \rho_h = 0$

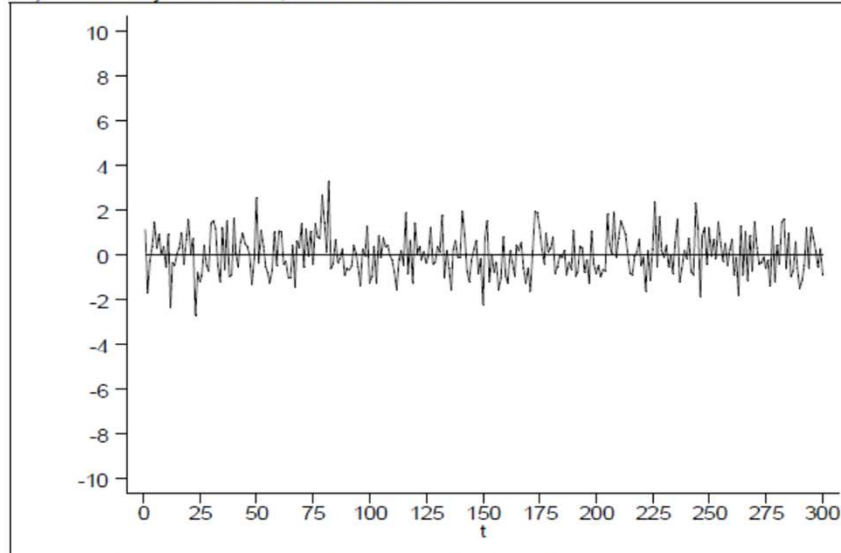
- Estatística de teste  $Q = T(T + 2) \sum_{j=1}^h \frac{\hat{\rho}_j^2}{T-j}$

- Segue uma distribuição  $\chi_{j-m}^2$ , onde  $m$  é o número de parâmetros estimados

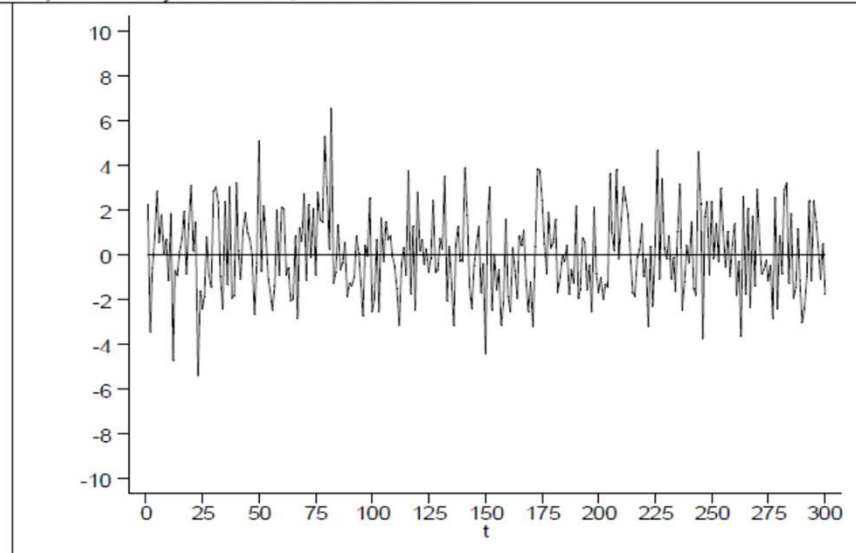
# Ruido branco

## Simulações para diferentes “white noises”

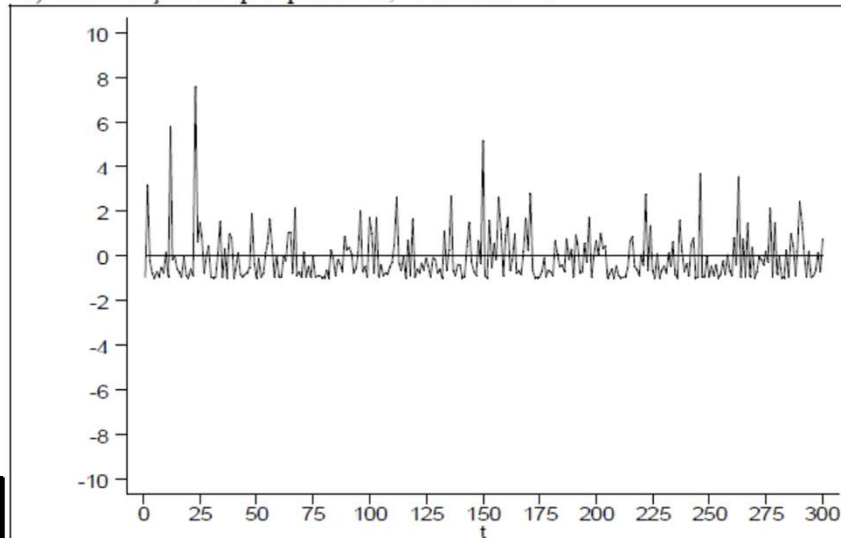
a) distribuição normal, variância = 1



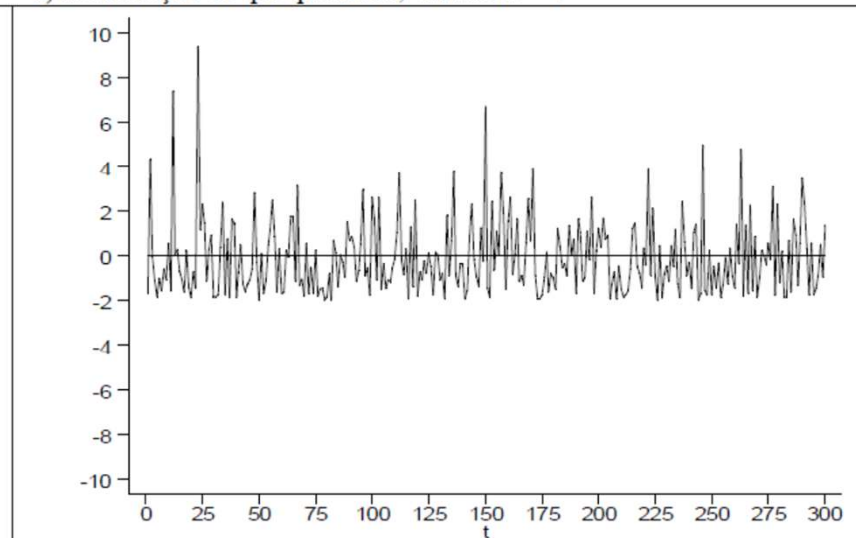
b) distribuição normal, variância = 4



c) distribuição de qui-quadrado, variância = 1



d) distribuição de qui-quadrado, variância = 4



# Ruido branco

“White noises”: correlogramas

a) distribuição normal, variância = 1

LAG	AC	PAC	Q	Prob>Q	-1	0	1	-1	0	1
					[Autocorrelation]			[Partial Autocor]		
1	-0.0245	-0.0245	.18137	0.6702						
2	0.0962	0.0962	2.9978	0.2234						
3	0.0162	0.0205	3.0778	0.3798						
4	-0.0604	-0.0703	4.1952	0.3802						
5	-0.0054	-0.0130	4.2041	0.5204						
6	-0.0164	-0.0037	4.2873	0.6379						
7	-0.0666	-0.0650	5.6582	0.5802						
8	-0.0031	-0.0078	5.6612	0.6851						
9	-0.0715	-0.0612	7.2535	0.6107						
10	0.1016	0.1060	10.477	0.3997						
11	-0.0242	-0.0162	10.662	0.4720						
12	-0.0050	-0.0217	10.669	0.5574						
13	-0.0262	-0.0400	10.886	0.6204						
14	0.0352	0.0507	11.277	0.6641						
15	-0.0328	-0.0319	11.618	0.7077						
16	-0.0141	-0.0323	11.682	0.7656						
17	-0.1143	-0.1174	15.867	0.5333						
18	-0.0502	-0.0487	16.676	0.5455						
19	-0.0370	-0.0158	17.117	0.5820						
20	-0.0465	-0.0571	17.817	0.5994						

# Ruido branco

“White noises”: correlogramas

c) distribuição de qui-quadrado, variância = 1

LAG	AC	PAC	Q	Prob>Q	-1 [Autocorrelation]	0	1	-1 [Partial Autocor]	0	1
1	-0.0515	-0.0516	.80474	0.3697						
2	0.0694	0.0675	2.2697	0.3215						
3	-0.0262	-0.0202	2.4793	0.4790						
4	-0.0655	-0.0735	3.7914	0.4350						
5	-0.0418	-0.0460	4.3287	0.5031						
6	-0.0601	-0.0561	5.4405	0.4887						
7	-0.0334	-0.0376	5.786	0.5649						
8	-0.0261	-0.0293	5.9973	0.6475						
9	-0.0599	-0.0695	7.1159	0.6251						
10	0.0536	0.0404	8.0138	0.6275						
11	0.1339	0.1454	13.638	0.2537			-		-	
12	-0.0046	0.0001	13.645	0.3240						
13	-0.0001	-0.0351	13.645	0.3993						
14	0.0895	0.0974	16.185	0.3022						
15	-0.0719	-0.0486	17.829	0.2718						
16	0.0097	-0.0053	17.859	0.3322						
17	-0.0308	-0.0089	18.164	0.3786						
18	-0.1111	-0.1146	22.13	0.2262						
19	-0.0336	-0.0423	22.495	0.2603						
20	-0.0961	-0.0668	25.48	0.1837						

# Notação para lags

O operador “backshift” ou operador de “lag”, designado de B ou L na literatura, é utilizado para representar lags:

$$\begin{aligned}By_t &= y_{t-1} \\ B^2 y_t &= B y_{t-1} = y_{t-2} \\ B^h y_t &= y_{t-h}\end{aligned}$$

- Em dados mensais, o mês correspondente a  $y_t$  no ano anterior é  $B^{12} y_t = y_{t-12}$
- Em dados trimestrais, o trimestre correspondente a  $y_t$  no ano anterior é  $B^4 y_t = y_{t-4}$

# Modelos autorregressivos

## Modelo AR(1):

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

com  $\varepsilon_t$  ruído branco com  $E(\varepsilon_t) = 0$ ,  $V(\varepsilon_t) = \sigma_\varepsilon^2$  e  $Cov(\varepsilon_t, \varepsilon_{t-h}) = 0$  para  $h \neq 0$

- O comportamento de  $y_t$  resulta do seu passado e de um termo erro  $\varepsilon_t$ , que inclui todos os outros determinantes de  $y_t$  para além de  $y_{t-1}$
- Caso  $\rho = 0$  então  $y_{t-1}$  não é útil para prever  $y_t$
- Caso  $\rho = 1$  tem-se um passeio aleatório – processo não estacionário a estudar mais tarde

# Modelos autorregressivos

Dado que

$$\begin{aligned}y_t &= \rho y_{t-1} + \varepsilon_t \\y_t &= \rho B y_t + \varepsilon_t \\y_t &= [1/(1 - \rho B)] \varepsilon_t\end{aligned}$$

e  $\frac{1}{(1-\rho B)} = 1 + \rho B + \rho^2 B^2 + \dots$  o modelo AR(1) pode ser visto como um modelo de médias móveis de ordem infinita

$$y_t = \varepsilon_t + \rho \varepsilon_{t-1} + \rho^2 \varepsilon_{t-2} + \dots$$

De onde resulta

$$E(y_t) = E(\varepsilon_t) + \rho E(\varepsilon_{t-1}) + \rho^2 E(\varepsilon_{t-2}) + \dots = 0$$

$$V(y_t) = E(y_t^2) = \frac{\sigma^2}{1 - \rho^2}$$

$$\text{Cov}(y_t, y_{t-h}) = \gamma_h = \rho^h \sigma^2$$

$$\text{Cor}(y_t, y_{t-h}) = \frac{\gamma_h}{\gamma_0} = \rho^h$$

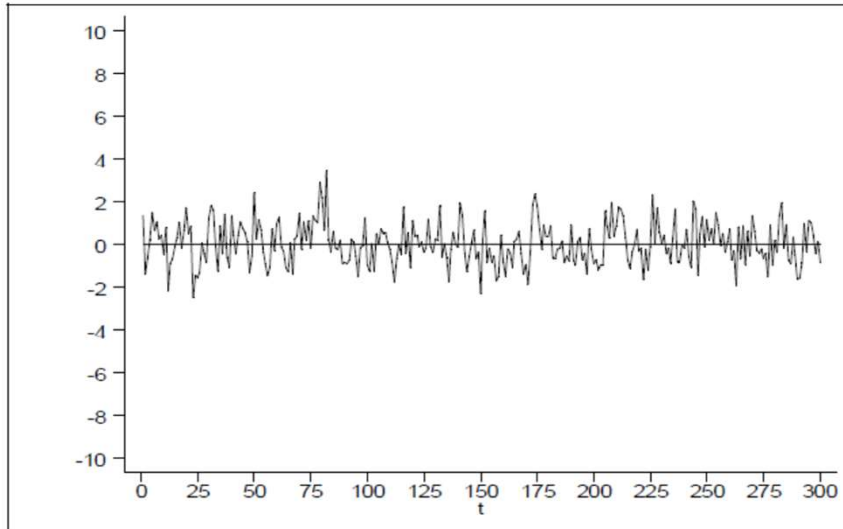
Se  $|\rho| < 1$   $y_t$  é fracamente dependente pois  $\lim_{h \rightarrow \infty} \text{Cor}(y_t, y_{t-h}) = 0$



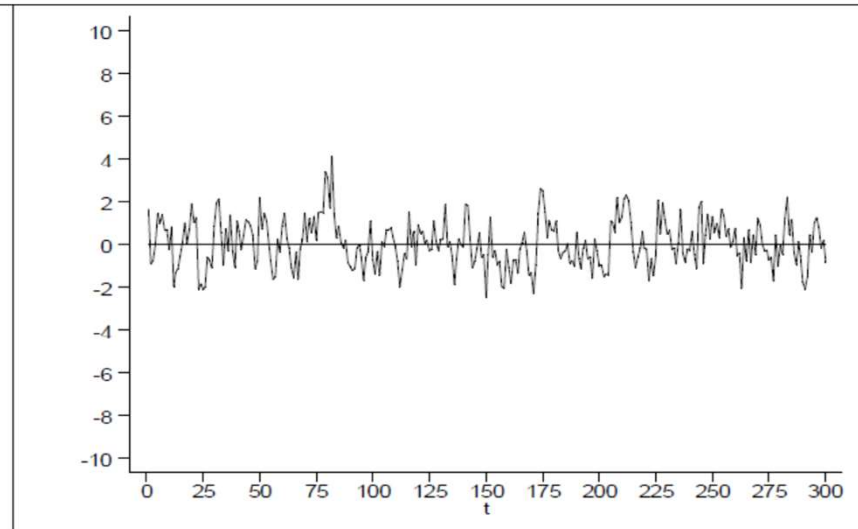
# Modelos autorregressivos

## Exemplos de processos AR(1)

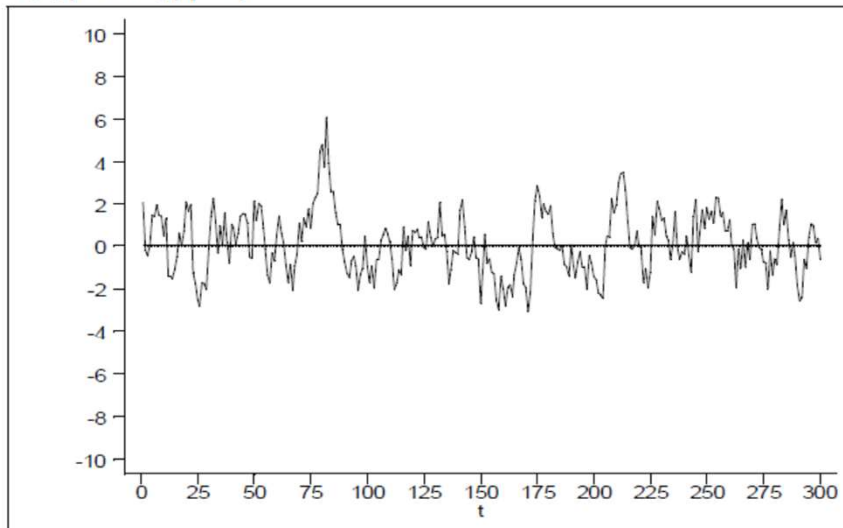
a)  $y_t = 0.25y_{t-1} + \varepsilon_t$



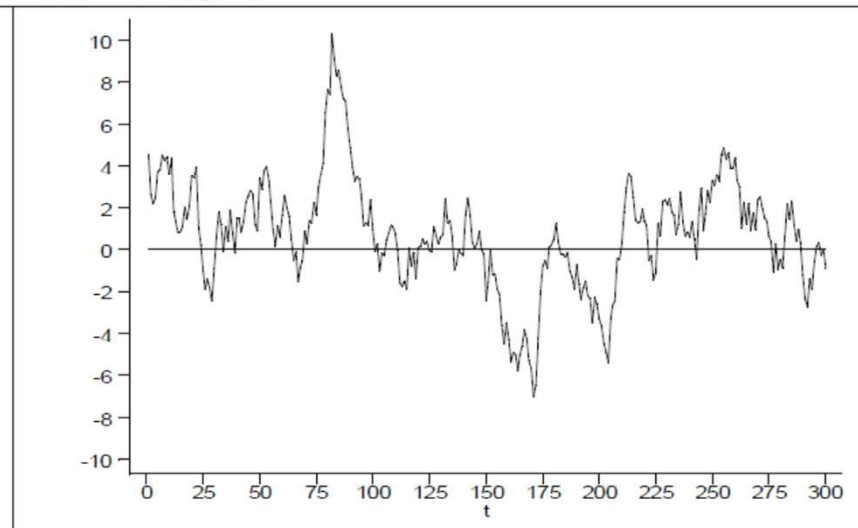
b)  $y_t = 0.5y_{t-1} + \varepsilon_t$



c)  $y_t = 0.75y_{t-1} + \varepsilon_t$



d)  $y_t = 0.95y_{t-1} + \varepsilon_t$



# Modelos autorregressivos

Exemplos de processos AR(1): correlograma

a)  $y_t = 0.5y_{t-1} + \varepsilon_t$

LAG	AC	PAC	Q	Prob>Q	-1 [Autocorrelation]	0	1 -1 [Partial Autocor]	0	1
1	0.5102	0.5112	78.871	0.0000	-----		-----		
2	0.3182	0.0726	109.66	0.0000	--				
3	0.1442	-0.0639	116.01	0.0000	-				
4	0.0189	-0.0713	116.12	0.0000					
5	-0.0121	0.0143	116.16	0.0000					
6	-0.0436	-0.0221	116.75	0.0000					
7	-0.0778	-0.0515	118.62	0.0000					
8	-0.0515	0.0198	119.45	0.0000					
9	-0.0501	-0.0101	120.23	0.0000					
10	0.0342	0.0949	120.59	0.0000					
11	-0.0134	-0.0825	120.65	0.0000					
12	-0.0221	-0.0265	120.8	0.0000					
13	-0.0321	-0.0168	121.13	0.0000					
14	-0.0195	0.0263	121.25	0.0000					
15	-0.0667	-0.0877	122.67	0.0000					

# Modelos autorregressivos

Exemplos de processos AR(1): correlograma

b)  $y_t = 0.75y_{t-1} + \varepsilon_t$

LAG	AC	PAC	Q	Prob>Q	-1	0	1	-1	0	1
					[Autocorrelation]			[Partial Autocor]		
1	0.7522	0.7528	171.43	0.0000		-----		-----		
2	0.5852	0.0342	275.54	0.0000		----				
3	0.4122	-0.1004	327.37	0.0000		---				
4	0.2643	-0.0629	348.75	0.0000		--				
5	0.1704	0.0298	357.67	0.0000		-				
6	0.0925	-0.0192	360.3	0.0000						
7	0.0289	-0.0339	360.56	0.0000						
8	0.0065	0.0346	360.58	0.0000						
9	-0.0155	-0.0084	360.65	0.0000						
10	0.0038	0.0614	360.66	0.0000						
11	-0.0349	-0.1176	361.04	0.0000						
12	-0.0568	-0.0344	362.05	0.0000						
13	-0.0791	-0.0250	364.03	0.0000						
14	-0.0927	-0.0007	366.75	0.0000						
15	-0.1334	-0.1075	372.4	0.0000						

# Modelos autorregressivos

Exemplos de processos AR(1): correlograma

c)  $y_t = 0.95y_{t-1} + \varepsilon_t$

LAG	AC	PAC	Q	Prob>Q	-1	0	1	-1	0	1
					[Autocorrelation]			[Partial Autocor]		
1	0.9294	0.9305	261.71	0.0000		-----		-----		
2	0.8676	0.0037	490.55	0.0000		-----				
3	0.7953	-0.1159	683.5	0.0000		-----				
4	0.7229	-0.0448	843.45	0.0000		-----				
5	0.6586	0.0492	976.66	0.0000		-----				
6	0.5958	-0.0105	1086.1	0.0000		-----				
7	0.5348	-0.0193	1174.5	0.0000		-----				
8	0.4847	0.0415	1247.4	0.0000		-----				
9	0.4349	-0.0126	1306.2	0.0000		-----				
10	0.3972	0.0407	1355.5	0.0000		-----				
11	0.3464	-0.1248	1393.1	0.0000		-----				
12	0.3042	-0.0175	1422.2	0.0000		-----				
13	0.2645	-0.0041	1444.3	0.0000		-----				
14	0.2294	0.0168	1461	0.0000		-----				
15	0.1907	-0.0747	1472.6	0.0000		-----				

Em qualquer um dos exemplos, a ACF apresenta uma redução progressiva e a PACF decai para zero para  $h \geq 2$

# Modelos autorregressivos

## Modelo AR(2):

$$\begin{aligned}y_t &= \rho_1 y_{t-1} + \rho_2 y_{t-2} + \varepsilon_t \\y_t &= \rho_1 B y_t + \rho_2 B^2 y_t + \varepsilon_t \\y_t &= [1/(1 - \rho_1 B - \rho_2 B^2)] \varepsilon_t\end{aligned}$$

- A ACF deste processo apresenta uma redução progressiva e a PACF decai para zero para  $h \geq 3$

## Modelo AR(p):

$$\begin{aligned}y_t &= \rho_1 y_{t-1} + \rho_2 y_{t-2} + \dots + \rho_p y_{t-p} + \varepsilon_t \\y_t &= [1/(1 - \rho_1 B - \rho_2 B^2 - \dots - \rho_p B^p)] \varepsilon_t\end{aligned}$$

- A ACF deste processo apresenta uma redução progressiva e a PACF decai para zero para  $k \geq p + 1$ , o que permite identificar o modelo

# Modelos autorregressivos

## Exemplos de AR(2): correlograma

a)  $y_t = 0.5y_{t-1} + 0.35y_{t-2} + \varepsilon_t$

LAG	AC	PAC	Q	Prob>Q	-1	0	1	-1	0	1
					[Autocorrelation]			[Partial Autocor]		
1	0.7332	0.7344	162.89	0.0000		-----		-----		
2	0.7168	0.3818	319.11	0.0000		-----		-----		
3	0.5806	-0.0713	421.92	0.0000		-----		-----		
4	0.4855	-0.1055	494.05	0.0000		-----		-----		
5	0.4058	0.0056	544.63	0.0000		-----		-----		
6	0.3324	0.0162	578.67	0.0000		-----		-----		
7	0.2529	-0.0521	598.44	0.0000		-----		-----		
8	0.2171	0.0350	613.07	0.0000		-----		-----		
9	0.1489	-0.0324	619.97	0.0000		-----		-----		
10	0.1475	0.0697	626.76	0.0000		-----		-----		
11	0.0723	-0.0950	628.4	0.0000		-----		-----		
12	0.0462	-0.0610	629.07	0.0000		-----		-----		
13	-0.0018	-0.0416	629.07	0.0000		-----		-----		
14	-0.0198	0.0246	629.19	0.0000		-----		-----		
15	-0.0727	-0.0789	630.88	0.0000		-----		-----		
16	-0.0938	-0.0410	633.68	0.0000		-----		-----		
17	-0.1410	-0.0466	640.04	0.0000		-----		-----		
18	-0.1373	0.0465	646.1	0.0000		-----		-----		
19	-0.1371	0.0592	652.16	0.0000		-----		-----		
20	-0.1382	-0.0192	658.34	0.0000		-----		-----		

# Modelos autorregressivos

## Exemplos de AR(2): correlograma

b)  $y_t = 0.75y_{t-1} + 0.25y_{t-2} + \varepsilon_t$

LAG	AC	PAC	Q	Prob>Q	-1 [Autocorrelation]	0	1	-1 [Partial Autocor]	0	1
1	0.9467	0.9470	271.58	0.0000	-----			-----		
2	0.9206	0.2661	529.22	0.0000	-----			--		
3	0.8788	-0.0951	764.81	0.0000	-----					
4	0.8383	-0.0781	979.88	0.0000	-----					
5	0.8050	0.0413	1178.9	0.0000	-----					
6	0.7709	0.0168	1362	0.0000	-----					
7	0.7387	-0.0247	1530.7	0.0000	-----					
8	0.7126	0.0492	1688.3	0.0000	-----					
9	0.6838	-0.0120	1833.9	0.0000	-----					
10	0.6643	0.0685	1971.8	0.0000	-----					
11	0.6330	-0.1006	2097.4	0.0000	-----					
12	0.6032	-0.0350	2211.8	0.0000	-----					
13	0.5716	-0.0105	2315	0.0000	-----					
14	0.5435	0.0372	2408.6	0.0000	-----					
15	0.5102	-0.0635	2491.3	0.0000	-----					
16	0.4815	-0.0037	2565.3	0.0000	-----					
17	0.4522	-0.0032	2630.7	0.0000	-----					
18	0.4318	0.0890	2690.6	0.0000	-----					
19	0.4148	0.0603	2746.1	0.0000	-----					
20	0.4019	0.0019	2798.4	0.0000	-----					

# Modelos de médias móveis

## Modelo MA(1)

$$y_t = \varepsilon_t + \theta\varepsilon_{t-1}, |\theta| < 1$$

onde  $\varepsilon_t$  é um ruído branco com média zero,  $E(\varepsilon_t) = 0$

Decorre que

$$E(y_t) = E(\varepsilon_t) + \theta E(\varepsilon_{t-1}) = 0$$

$$V(y_t) = V(\varepsilon_t + \theta\varepsilon_{t-1}) = V(\varepsilon_t) + \theta^2 V(\varepsilon_{t-1}) + 2\theta \text{Cov}(\varepsilon_t \varepsilon_{t-1}) = \sigma^2(1 + \theta^2)$$

$$\text{Cov}(y_t, y_{t-h}) = \begin{cases} \theta\sigma^2 & \text{se } h = 1 \\ 0 & \text{se } h > 1 \end{cases}$$

$$\text{Cor}(y_t, y_{t-h}) = \begin{cases} \frac{\theta\sigma^2}{\sigma^2(1+\theta^2)} = \frac{\theta}{(1+\theta^2)} & \text{se } h = 1, \\ 0 & \text{se } h > 1 \end{cases},$$

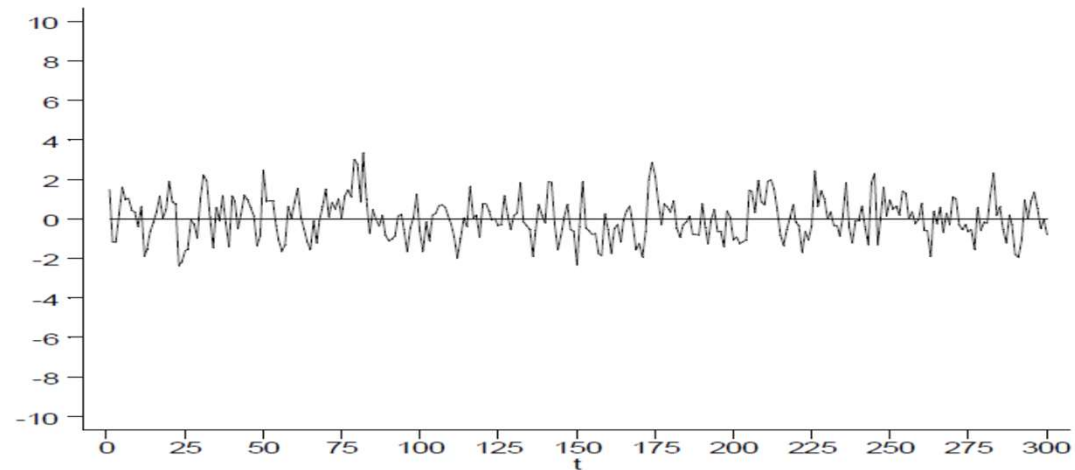
sendo fracamente dependente



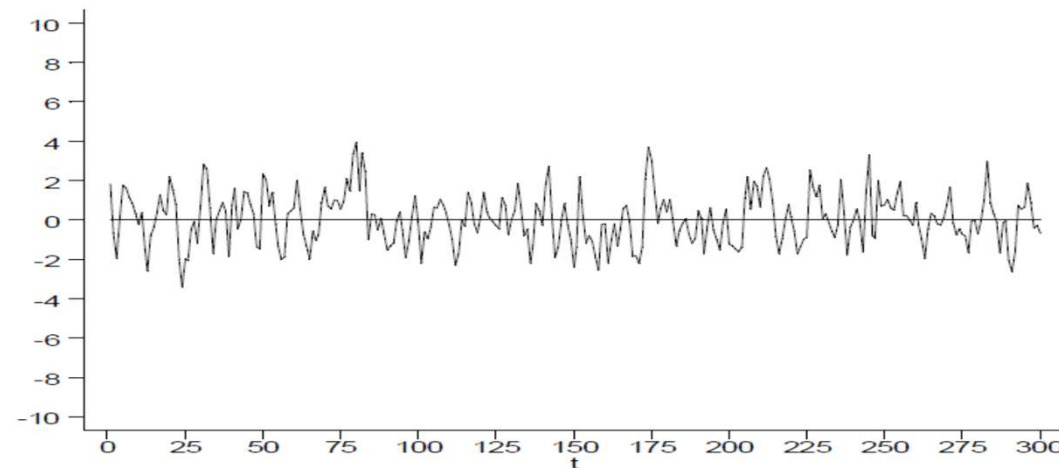
# Modelos de médias móveis

## Exemplos de modelo MA(1)

a)  $y_t = \varepsilon_t + 0.5\varepsilon_{t-1}$



b)  $y_t = \varepsilon_t + 0.95\varepsilon_{t-1}$



# Modelos de médias móveis

## Exemplos de modelo MA(1)

LAG	AC	PAC	Q	Prob>Q	-1	0	1	-1	0	1
					[Autocorrelation]			[Partial Autocor]		
1	0.4190	0.4198	53.194	0.0000		---			---	
2	0.0932	-0.1047	55.833	0.0000						
3	0.0305	0.0367	56.117	0.0000						
4	-0.0547	-0.0902	57.033	0.0000						
5	-0.0339	0.0360	57.387	0.0000						
6	-0.0437	-0.0521	57.976	0.0000						
7	-0.0753	-0.0400	59.729	0.0000						
8	-0.0577	-0.0177	60.763	0.0000						
9	-0.0346	0.0001	61.137	0.0000						
10	0.0624	0.0972	62.351	0.0000						
11	0.0100	-0.0767	62.383	0.0000						
12	-0.0278	-0.0117	62.626	0.0000						
13	-0.0149	-0.0107	62.696	0.0000						
14	0.0098	0.0392	62.727	0.0000						
15	-0.0252	-0.0667	62.929	0.0000						
16	-0.0723	-0.0541	64.595	0.0000						
17	-0.1424	-0.1116	71.085	0.0000		-				
18	-0.1091	-0.0014	74.91	0.0000						
19	-0.0734	-0.0388	76.645	0.0000						
20	-0.0429	-0.0087	77.241	0.0000						

# Modelos de médias móveis

## Exemplos de modelo MA(1)

LAG	AC	PAC	Q	Prob>Q	-1 [Autocorrelation]	0	1	-1	0	1	[Partial Autocor]
1	0.5326	0.5331	85.939	0.0000			----			----	
2	0.0909	-0.2745	88.45	0.0000					--		
3	0.0341	0.1750	88.805	0.0000						-	
4	-0.0525	-0.2107	89.649	0.0000						-	
5	-0.0394	0.1662	90.126	0.0000						-	
6	-0.0510	-0.1967	90.927	0.0000						-	
7	-0.0761	0.1054	92.716	0.0000						-	
8	-0.0720	-0.1541	94.326	0.0000						-	
9	-0.0260	0.1729	94.537	0.0000						-	
10	0.0509	-0.0704	95.345	0.0000							
11	0.0188	0.0247	95.456	0.0000							
12	-0.0346	-0.0945	95.833	0.0000							
13	-0.0129	0.0923	95.886	0.0000							
14	0.0023	-0.0679	95.887	0.0000							
15	-0.0226	0.0031	96.05	0.0000							
16	-0.0874	-0.1398	98.488	0.0000						-	
17	-0.1478	-0.0262	105.48	0.0000			-				
18	-0.1233	-0.0395	110.36	0.0000							
19	-0.0826	-0.0240	112.56	0.0000							
20	-0.0408	0.0204	113.1	0.0000							

# Modelos de médias móveis

## Modelo MA(2)

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$
$$y_t = (1 + \theta_1 B + \theta_2 B^2) \varepsilon_t$$

A ACF deste processo decai para zero para  $k \geq 3$  e a PACF apresenta uma redução progressiva

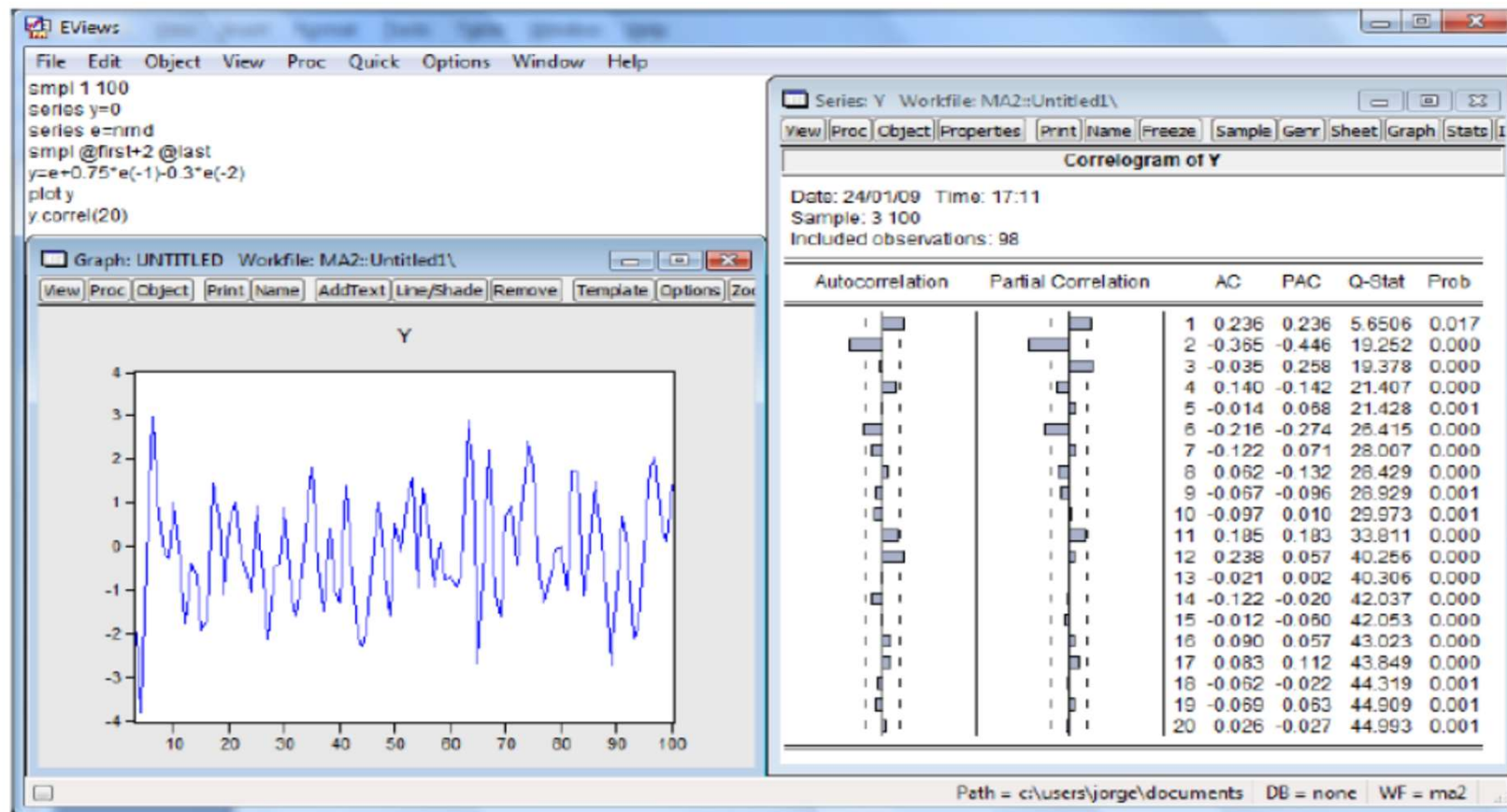
## Modelo MA(q)

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$
$$y_t = (1 + \theta_1 B + \dots + \theta_q B^q) \varepsilon_t$$

A ACF deste processo decai para zero para  $k \geq q + 1$  e a PACF apresenta uma redução progressiva

# Modelos de médias móveis

## Exemplo de modelo MA(2)



ACF and PACF of the simulated MA(2) model:  $Y_t = \varepsilon_t + 0,75\varepsilon_{t-1} - 0,2\varepsilon_{t-2}$

# Modelos ARMA

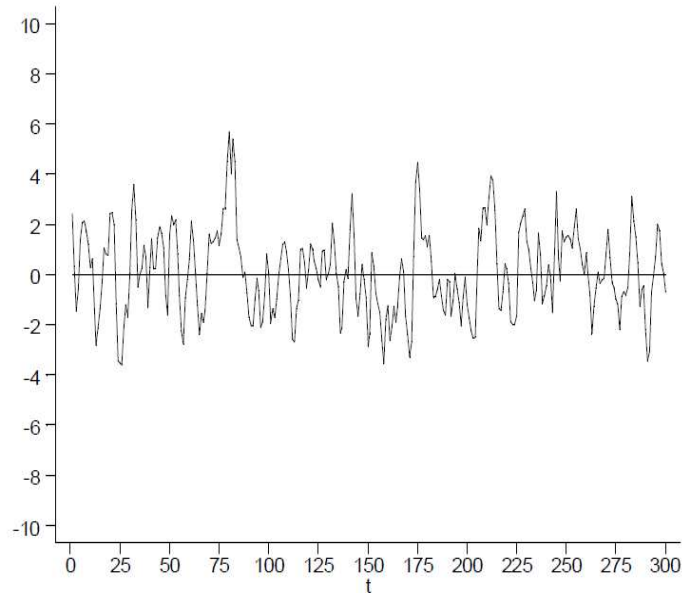
## Modelo ARMA(1,1)

$$y_t = \rho y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$$
$$y_t(1 - \rho B) = \varepsilon_t(1 + \theta B)$$

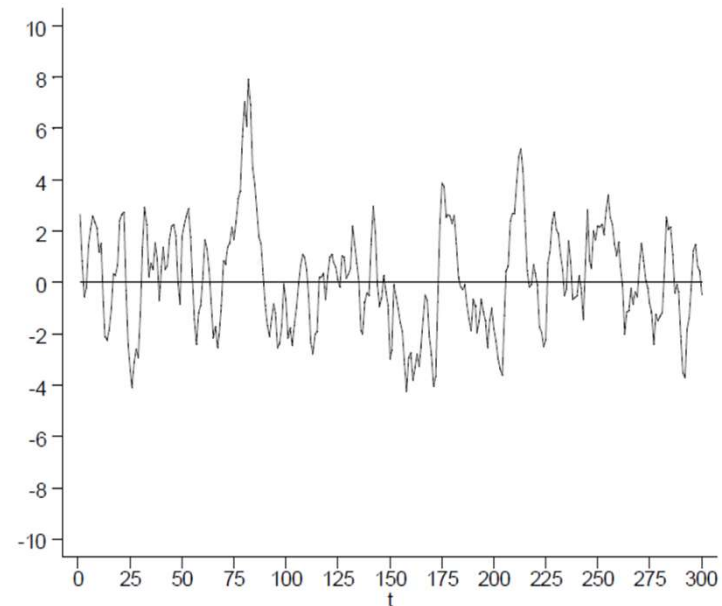
onde a estacionaridade ocorre se  $|\rho| < 1$

Exemplo:

a)  $y_t = 0.5y_{t-1} + \varepsilon_t + 0.75\varepsilon_{t-1}$



b)  $y_t = 0.75y_{t-1} + \varepsilon_t + 0.5\varepsilon_{t-1}$



# Modelos ARMA

## Correlograma

LAG	AC	PAC	Q	Prob>Q	-1	0	1	-1	0	1
					[Autocorrelation]			[Partial Autocor]		
1	0.7646	0.7652	177.15	0.0000		-----		-----		
2	0.4219	-0.4050	231.27	0.0000		---		---		
3	0.2045	0.1616	244.03	0.0000		-		-		
4	0.0574	-0.1704	245.04	0.0000				-		
5	-0.0129	0.1261	245.09	0.0000				-		
6	-0.0558	-0.1504	246.05	0.0000				-		
7	-0.0813	0.0779	248.09	0.0000						
8	-0.0752	-0.0457	249.85	0.0000						
9	-0.0413	0.0959	250.38	0.0000						
10	0.0004	-0.0407	250.38	0.0000						
11	-0.0080	-0.0710	250.4	0.0000						
12	-0.0323	0.0045	250.73	0.0000						
13	-0.0373	-0.0021	251.17	0.0000						
14	-0.0471	-0.0391	251.87	0.0000						
15	-0.0824	-0.1047	254.03	0.0000						
16	-0.1365	-0.0530	259.98	0.0000		-				
17	-0.1798	-0.0466	270.33	0.0000		-				
18	-0.1719	0.0392	279.82	0.0000		-				
19	-0.1364	-0.0494	285.81	0.0000		-				
20	-0.0954	0.0442	288.76	0.0000						



# Modelos ARMA

## Correlograma

LAG	AC	PAC	Q	Prob>Q	-1	0	1	-1	0	1
					[Autocorrelation]			[Partial Autocor]		
1	0.8640	0.8642	226.19	0.0000		-----		-----		
2	0.6547	-0.3801	356.49	0.0000		-----		---		
3	0.4686	0.0548	423.47	0.0000		---				
4	0.3109	-0.0745	453.05	0.0000		--				
5	0.1963	0.0613	464.89	0.0000		-				
6	0.1084	-0.0667	468.5	0.0000						
7	0.0433	0.0188	469.08	0.0000						
8	0.0077	0.0210	469.1	0.0000						
9	-0.0077	0.0130	469.12	0.0000						
10	-0.0106	-0.0061	469.16	0.0000						
11	-0.0360	-0.1287	469.56	0.0000				-		
12	-0.0650	0.0116	470.89	0.0000						
13	-0.0877	-0.0318	473.32	0.0000						
14	-0.1118	-0.0392	477.27	0.0000						
15	-0.1478	-0.1140	484.22	0.0000		-				
16	-0.1873	-0.0233	495.41	0.0000		-				
17	-0.2185	-0.0228	510.69	0.0000		-				
18	-0.2180	0.0567	525.96	0.0000		-				
19	-0.1989	-0.0306	538.72	0.0000		-				
20	-0.1733	0.0153	548.44	0.0000		-				



# Modelos ARMA

## Modelo ARMA(p,q)

$$y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \dots + \rho_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

$$(1 - \rho_1 B - \rho_2 B^2 - \dots - \rho_p B^p) y_t = (1 + \theta_1 B + \dots + \theta_q B^q) \varepsilon_t$$

$$\rho_p(B) y_t = \theta_q(B) \varepsilon_t$$

# Exercício

Seja  $e_t, t = -1, 0, 1, \dots$  uma sucessão de variáveis aleatórias independentes e identicamente distribuídas com média zero e variância unitária,  $e_t \sim i. i. d. (0, 1)$ . Defina o processo estocástico

$$x_t = e_t - 0.5e_{t-1} + 0.5e_{t-2}, t = 1, 2, \dots$$

- Determine  $E(x_t)$  e  $V(x_t)$  e verifique se dependem de  $t$ .
- Mostre que  $\text{Cor}(x_t, x_{t+1}) = -0.5$  e  $\text{Cor}(x_t, x_{t+2}) = 1/3$
- Qual é a  $\text{Cor}(x_t, x_{t+h})$  para  $h > 2$

Tem-se

- $E(e_t) = 0$
- $V(e_t) = 1$
- $\text{Cov}(e_t, e_s) = 0$  para  $t \neq s$  devido à independência

# Exercício

a)

- $E(x_t) = E(e_t) - 0.5E(e_{t-1}) + 0.5E(e_{t-2}) = 0$
- $V(x_t) = V(e_t) - 0.5^2V(e_{t-1}) + 0.5^2V(e_{t-2}) = 1 + 0.25 + 0.25 = 1.5$

Não dependem de t

b)

- $$\text{Cor}(x_t, x_{t+1}) = \frac{\text{Cov}(x_t, x_{t+1})}{\sqrt{V(x_t)}\sqrt{V(x_{t+1})}} = \frac{\text{Cov}(x_t, x_{t+1})}{\sqrt{1.5}\sqrt{1.5}} = \frac{-0.75}{1.5} = -0.25$$

$$\begin{aligned}\text{Cov}(x_t, x_{t+1}) &= E(x_t x_{t+1}) = E((e_t - 0.5e_{t-1} + 0.5e_{t-2})(e_{t+1} - 0.5e_t + 0.5e_{t-1})) = \\ &= -0.5E(e_t^2) - 0.25E(e_{t-1}^2) = -0.5 - 0.25 = -0.75\end{aligned}$$

- $$\text{Cor}(x_t, x_{t+2}) = \frac{\text{Cov}(x_t, x_{t+2})}{\sqrt{V(x_t)}\sqrt{V(x_{t+2})}} = \frac{\text{Cov}(x_t, x_{t+2})}{\sqrt{1.5}\sqrt{1.5}} = \frac{0.5}{1.5} = 1/3$$

$$\begin{aligned}\text{Cov}(x_t, x_{t+2}) &= E((e_t - 0.5e_{t-1} + 0.5e_{t-2})(e_{t+2} - 0.5e_{t+1} + 0.5e_t)) = 0.5E(e_t^2) \\ &= 0.5\end{aligned}$$

# Exercício

c)

$$\text{Cor}(x_t, x_{t+h}) = \frac{\text{Cov}(x_t, x_{t+h})}{\sqrt{V(x_t)}\sqrt{V(x_{t+h})}} = \frac{\text{Cov}(x_t, x_{t+h})}{\sqrt{1.5}\sqrt{1.5}} = \frac{0}{1.5} = 0$$
$$\text{Cov}(x_t, x_{t+3}) = E((e_t - 0.5e_{t-1} + 0.5e_{t-2})(e_{t+3} - 0.5e_{t+2} + 0.5e_{t+1})) = 0$$

Para  $h=4, 5, \dots$  também se terá 0

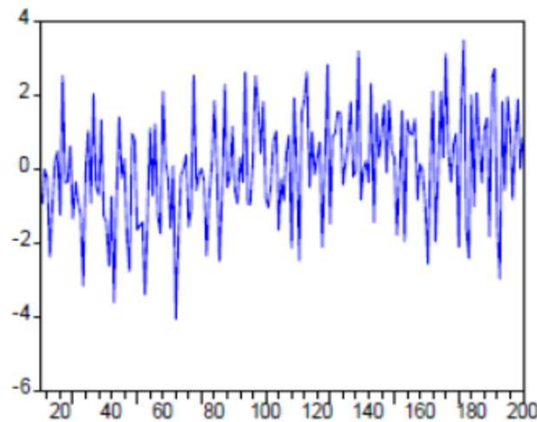
- $\text{Cor}(x_t, x_{t+h}) = 0$  para  $h > 2$

# Modelos ARMA sazonais - SARMA

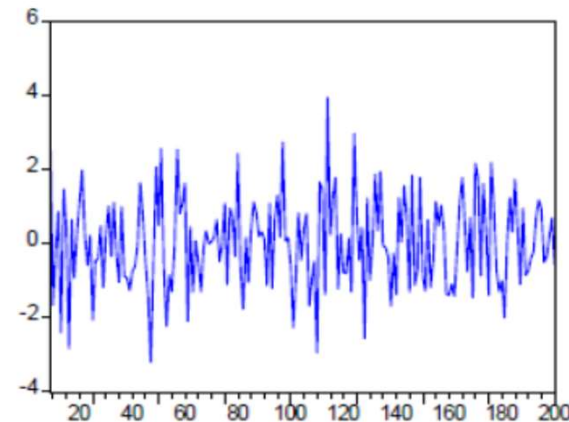
Modelo SARMA( $P, Q$ ) <sub>$s$</sub>

$$y_t = P_1 y_{t-s} + \dots + P_p y_{t-P_s} + \varepsilon_t + \Theta_1 \varepsilon_{t-s} + \dots + \Theta_q \varepsilon_{t-Q_s}$$
$$y_t (1 - P_1 B - \dots - P_p B^{P_s}) = \varepsilon_t (1 + \Theta_1 B + \dots + \Theta_q B^{Q_s})$$

(i)  $(1 - 0.65B^{12})Y_t = (1 + 0.25B^{12})\varepsilon_t$



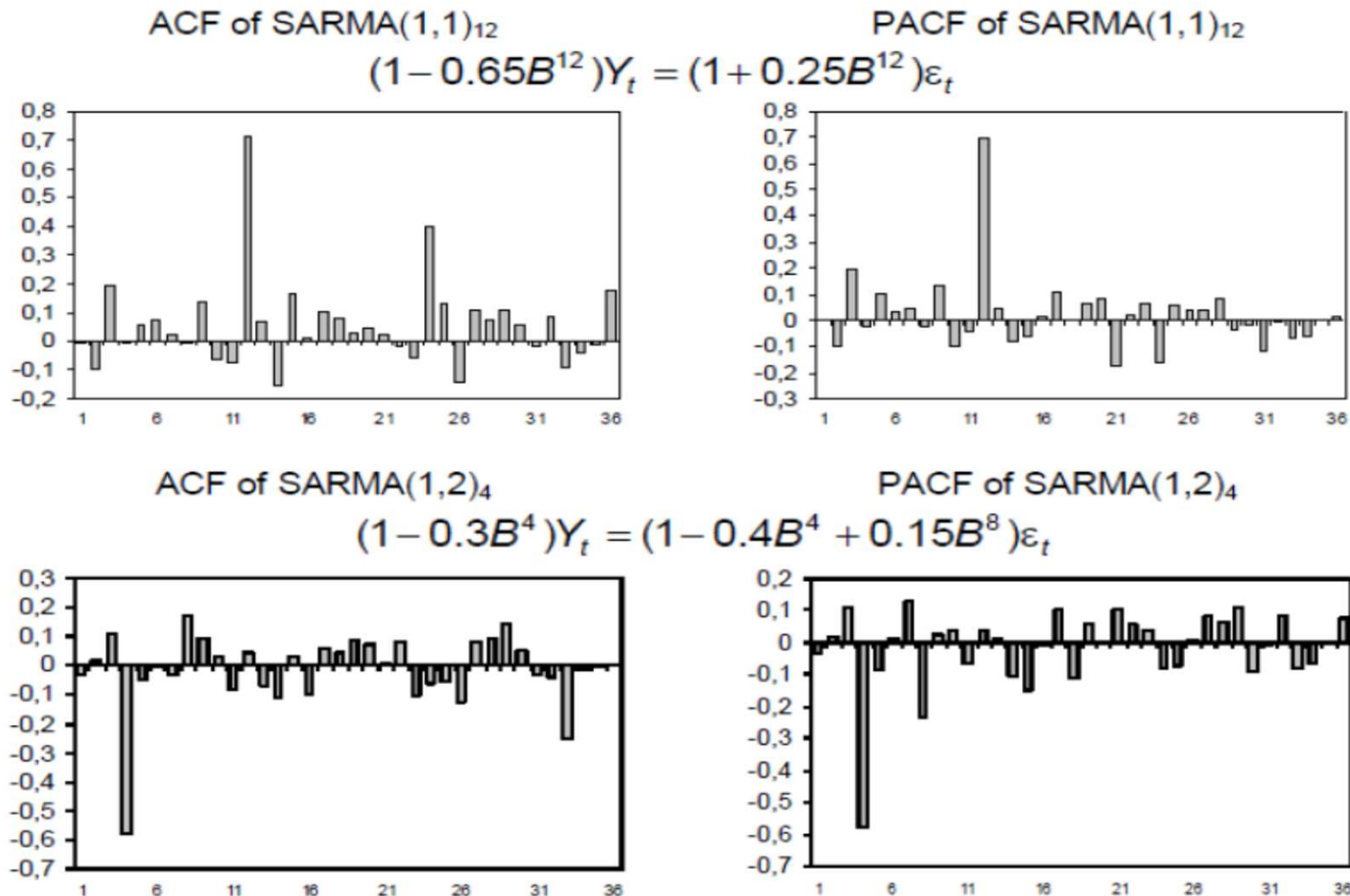
(ii)  $(1 - 0.3B^4)Y_t = (1 - 0.4B^4 + 0.15B^8)\varepsilon_t$



Simulated SARMA(1,1)<sub>12</sub> and SARMA(1,2)<sub>4</sub> models

# Modelos ARIMA sazonais - SARMA

Modelo SARMA( $p, q$ )<sub>s</sub>



ACF and PACF of simulated SARMA(1,1)<sub>12</sub> and SARMA(1,2)<sub>4</sub> models

# Processos não estacionários

Tipos principais de modelos para processos não estacionários:

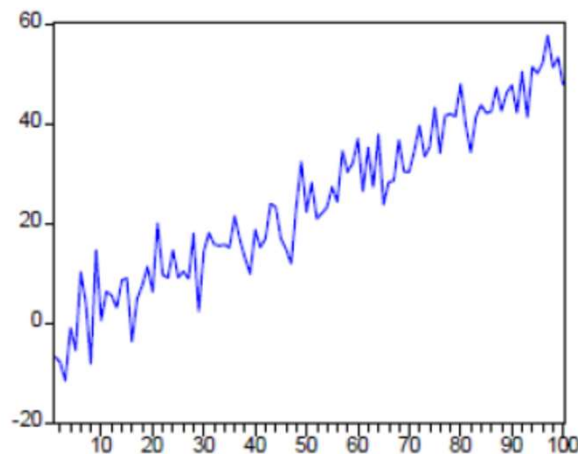
- Modelos determinísticos não estacionários em média: modelos de tendência determinística
- Modelos de tendência estocástica (passeio aleatório, passeio aleatório com “drift”)

# Processos não estacionários

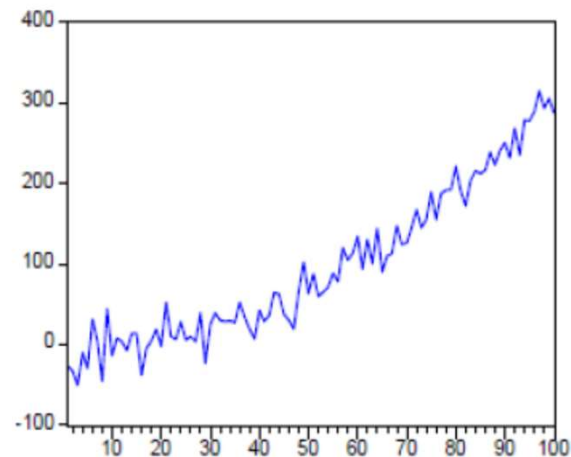
## Modelos de tendência determinística

- A tendência,  $t$ ,  $t = 1, 2, \dots, T$ , representa a ordem da observação na série
- Tendência linear:  $y_t = \beta_0 + \beta_1 t + u_t$
- Tendência quadrática:  $y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + u_t$
- Tendência exponencial:  $\ln(Y_t) = \beta_0 + \beta_1 t + \dots + u_t$

Linear trend model



Quadratic trend model

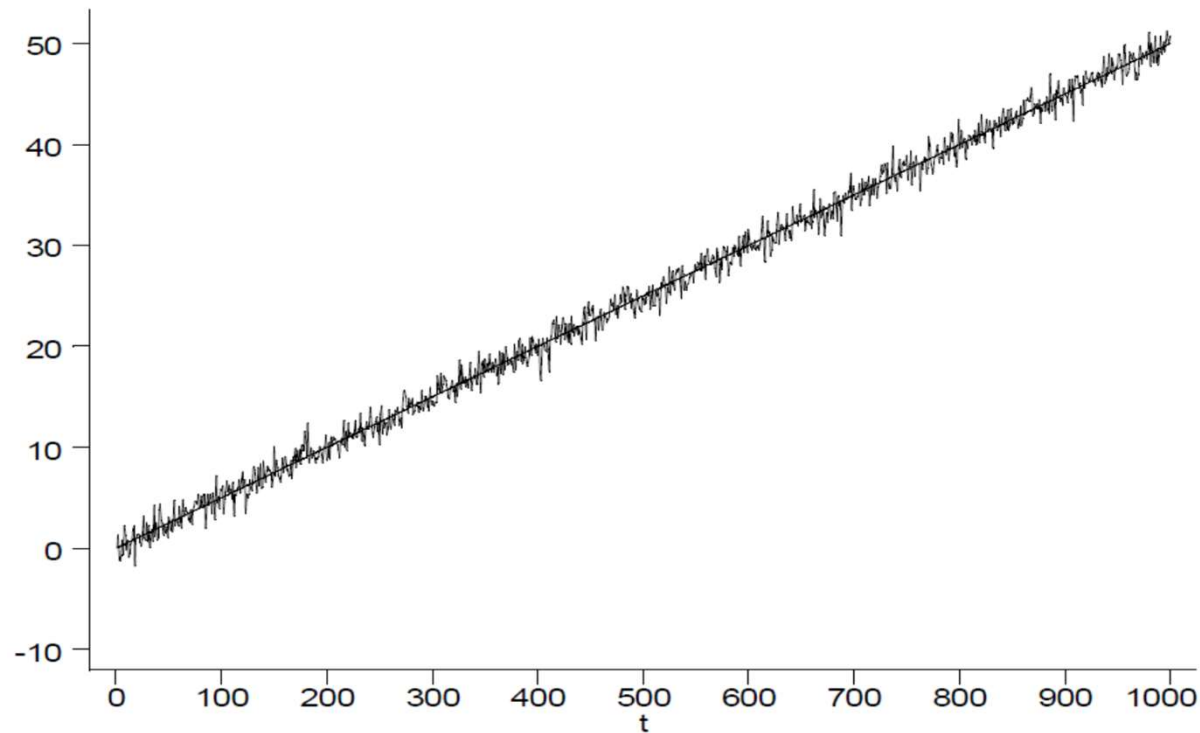




# Processos não estacionários

Exemplo de modelos de tendência determinística

$$y_t = 0.05t + \varepsilon_t$$



# Processos não estacionários

Exemplo de modelos de tendência determinística

LAG	AC	PAC	Q	Prob>Q	-1 [Autocorrelation]	0	1 -1 [Partial Autocor]	0	1
1	0.9923	0.9954	987.57	0.0000		-----		-----	
2	0.9893	0.4900	1970.3	0.0000		-----		---	
3	0.9863	0.3528	2948	0.0000		-----		--	
4	0.9826	0.1933	3919.3	0.0000		-----		-	
5	0.9798	0.2406	4886.1	0.0000		-----		-	
6	0.9771	0.2109	5848.4	0.0000		-----		-	
7	0.9733	0.0723	6804.4	0.0000		-----			
8	0.9710	0.1411	7756.7	0.0000		-----		-	
9	0.9680	0.0999	8704.2	0.0000		-----			
10	0.9648	0.0884	9646.3	0.0000		-----			
11	0.9619	0.0629	10584	0.0000		-----			
12	0.9590	0.0898	11516	0.0000		-----			
13	0.9557	0.0468	12444	0.0000		-----			
14	0.9531	0.0922	13367	0.0000		-----			
15	0.9499	0.0793	14285	0.0000		-----			
16	0.9469	0.0415	15198	0.0000		-----			
17	0.9442	0.0190	16106	0.0000		-----			
18	0.9409	0.0406	17010	0.0000		-----			
19	0.9380	0.0930	17908	0.0000		-----			
20	0.9352	0.0649	18803	0.0000		-----			

# Processos não estacionários

## Modelos de tendência estocástica

### 1) Passeio Aleatório

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

Substituindo sucessivamente:

$$y_t = y_{t-2} + \varepsilon_{t-1} + \varepsilon_t = y_{t-3} + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t$$

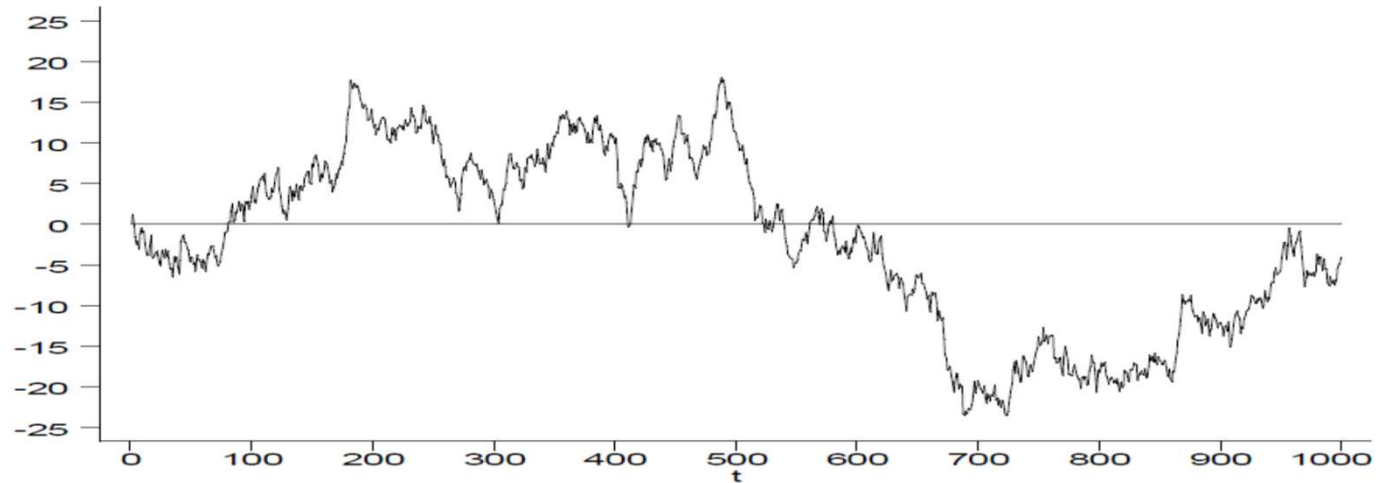
$$y_t = y_0 + \varepsilon_1 + \dots + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t$$

Obtendo-se

- $E(y_t) = y_0$
- $V(y_t) = \sigma_\varepsilon^2 t$
  
- Série não estacionária: variância depende de t

# Processos não estacionários

Exemplo de passeio aleatório



LAG	AC	PAC	Q	Prob>Q	-1 [Autocorrelation]	0	1 -1 [Partial Autocor]	0	1
1	0.9958	0.9958	994.53	0.0000					
2	0.9914	-0.0150	1981.3	0.0000					
3	0.9871	0.0028	2960.5	0.0000					
4	0.9825	-0.0340	3931.7	0.0000					
5	0.9784	0.0561	4895.7	0.0000					
6	0.9740	-0.0395	5852.1	0.0000					
7	0.9691	-0.0594	6799.8	0.0000					
8	0.9647	0.0606	7739.9	0.0000					
9	0.9602	-0.0203	8672.1	0.0000					
10	0.9557	0.0062	9596.6	0.0000					
11	0.9514	0.0093	10514	0.0000					
12	0.9473	0.0295	11424	0.0000					
13	0.9432	-0.0031	12327	0.0000					
14	0.9395	0.0359	13224	0.0000					
15	0.9357	-0.0140	14115	0.0000					
16	0.9318	-0.0084	14999	0.0000					
17	0.9282	0.0256	15877	0.0000					
18	0.9250	0.0486	16750	0.0000					
19	0.9221	0.0281	17619	0.0000					
20	0.9190	-0.0289	18482	0.0000					

# Processos não estacionários

## 2) Passeio aleatório com drift

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

Substituindo sucessivamente:

$$y_t = \alpha + \alpha + y_{t-2} + \varepsilon_{t-1} + \varepsilon_t = \alpha + \alpha + \alpha + y_{t-3} + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t$$

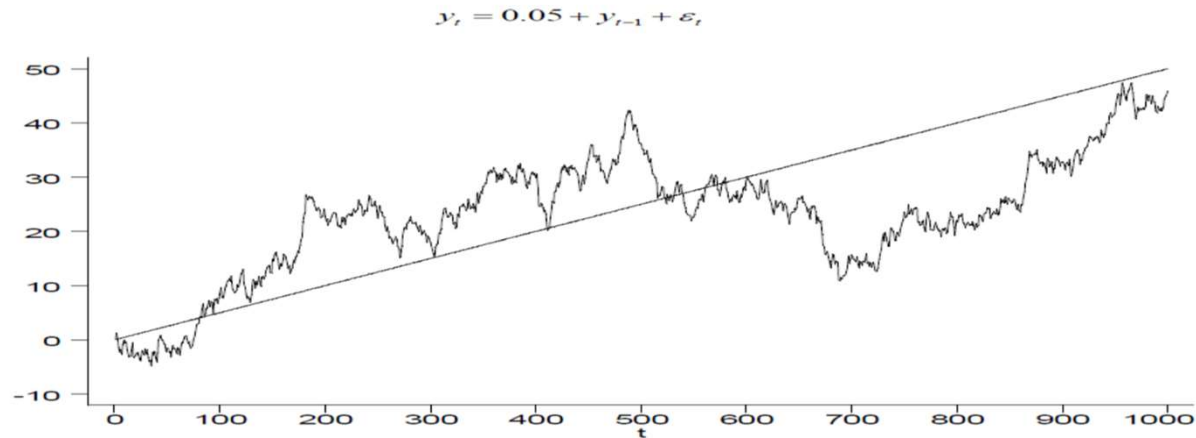
$$y_t = \alpha t + y_0 + \varepsilon_1 + \dots + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t$$

Obtendo-se

- $E(y_t) = \alpha t + y_0$
- $V(y_t) = \sigma_\varepsilon^2 t$
  
- Série não estacionária: tanto a média como a variância dependem de t

# Processos não estacionários

Exemplo de passeio aleatório com drift

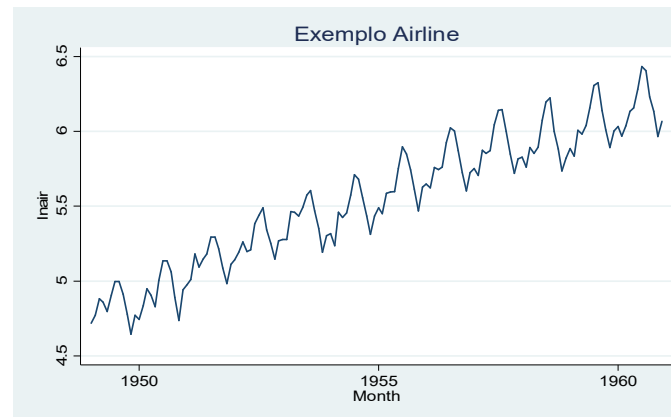
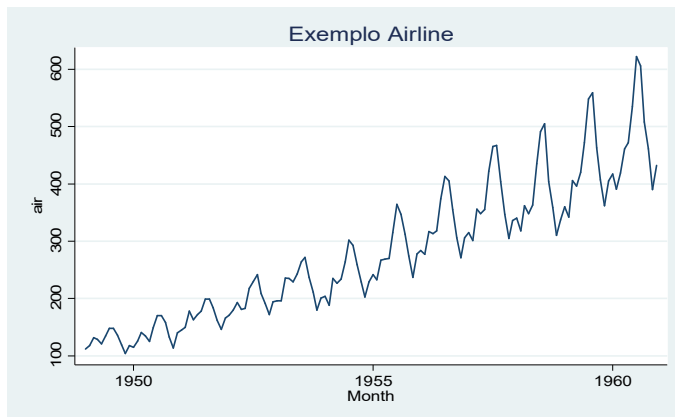


LAG	AC	PAC	Q	Prob>Q	-1	0	1	-1	0	1
					[Autocorrelation]			[Partial Autocor]		
1	0.9916	0.9959	986.24	0.0000						
2	0.9835	-0.0148	1957.4	0.0000						
3	0.9752	0.0029	2913.1	0.0000						
4	0.9664	-0.0340	3852.7	0.0000						
5	0.9582	0.0562	4777.3	0.0000						
6	0.9499	-0.0393	5686.8	0.0000						
7	0.9408	-0.0593	6580	0.0000						
8	0.9327	0.0608	7458.7	0.0000						
9	0.9246	-0.0200	8323	0.0000						
10	0.9162	0.0063	9172.7	0.0000						
11	0.9083	0.0096	10008	0.0000						
12	0.9002	0.0296	10830	0.0000						
13	0.8915	-0.0033	11637	0.0000						
14	0.8832	0.0356	12430	0.0000						
15	0.8745	-0.0144	13208	0.0000						
16	0.8663	-0.0085	13972	0.0000						
17	0.8584	0.0255	14723	0.0000						
18	0.8502	0.0481	15461	0.0000						
19	0.8422	0.0277	16185	0.0000						
20	0.8343	-0.0292	16897	0.0000						

# Processos não estacionários

## Transformação da série para estabilização da variância

Quando o processo não é estacionário em variância, a literatura recomenda uma transformação para estabilização da variância. Tipicamente, em lugar da série em níveis, modela-se o seu logaritmo. Veja-se o exemplo:



Note-se que a previsão para a escala original em níveis requer processos complexos de retransformação. Em alternativa, pode-se modelar  $\Delta \ln(y_t) = \ln(y_t) - \ln(y_{t-1})$  e prever a taxa de crescimento.

# Testes de estacionaridade

## 1) Teste de Dickey Fuller Aumentado (ADF)

$H_0 : \theta = 0$  (existe raiz unitária - série não estacionária)

$H_0 : \theta < 0$  (não existe raiz unitária - série estacionária)

Num dos modelos, fazer o teste t para  $\theta$ :

$$\Delta Y_t = \theta Y_{t-1} + \beta_1 \Delta Y_{t-1} + \dots + \beta_q \Delta Y_{t-q} + u_t$$

$$\Delta Y_t = \alpha + \theta Y_{t-1} + \beta_1 \Delta Y_{t-1} + \dots + \beta_q \Delta Y_{t-q} + u_t$$

$$\Delta Y_t = \alpha + \theta Y_{t-1} + \delta t + \beta_1 \Delta Y_{t-1} + \dots + \beta_q \Delta Y_{t-q} + u_t$$

Correspondendo a um teste t unilateral, onde o valor crítico deverá ser recolhido da tabela DF

## 2) Teste de Kwiatkowski Phillips Schmidt Shin (KPSS)

$H_0$ : série estacionária e não sazonal

## 3) Outros: Phillips Perron (PP), testes sazonais



# Testes de estacionaridade

## Valores críticos assintóticos dos testes de raízes unitárias DF e ADF

Tipo de modelo	Nível de significância		
	1%	5%	10%
passeio aleatório	-2.56	-1.94	-1.62
passeio aleatório com deslocação	-3.43	-2.86	-2.57
passeio aleatório com deslocação e tendência	-3.96	-3.41	-3.13

## Valores críticos (5%) assintóticos do teste de raízes unitárias KPSS

Tipo de modelo	Valor crítico
modelo sem tendência	0.463
modelo com tendência	0.146

# Testes de estacionaridade

Com base nos testes de raízes unitárias, podem ser distinguidas dois tipos de variáveis:

- ▶ Variáveis  $I(0)$ : variáveis integradas de ordem 0 (sem raízes unitárias, estacionárias)
- ▶ Variáveis  $I(1)$ : variáveis integradas de ordem 1 (com uma raiz unitária, não estacionárias)

Consequências para a estimação de um modelo de regressão:

- ▶ Todas as variáveis do modelo são  $I(0)$  - o método dos MQ pode ser aplicado da forma habitual
- ▶ Pelo menos uma das variáveis do modelo é  $I(1)$  - o método dos MQ pode conduzir a uma *relação espúria*: os resultados da estimação e dos testes sugerem que existe uma relação entre a variável dependente e as variáveis explicativas mesmo que na verdade não exista qualquer relação

# Estacionarização de séries

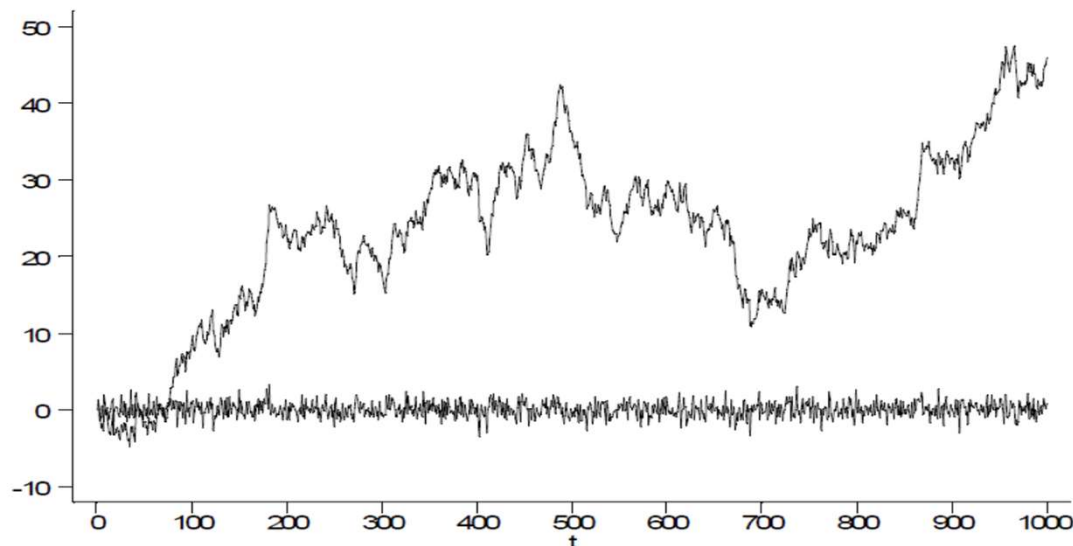
Caso a série não seja estacionária, poderá estacionarizar-se. De facto, se a série  $y_t$  for  $I(1)$ , a sua primeira diferença será  $I(0)$ . Veja-se o exemplo do passeio aleatório:

$$Y_t = \alpha + Y_{t-1} + u_t \rightarrow \text{passeio aleatório}$$

$$Y_t - Y_{t-1} = \alpha + u_t$$

$$\Delta Y_t = \alpha + u_t \rightarrow \text{série estacionária}$$

$$y_t = 0.05 + y_{t-1} + \varepsilon_t \text{ vs. } \Delta y_t = y_t - y_{t-1}$$



# Estacionarização de séries

Exemplo da primeira diferença de um passeio aleatório

$$\Delta y_t = y_t - y_{t-1}$$

LAG	AC	PAC	Q	Prob>Q	-1 [Autocorrelation]	0	1	-1 [Partial Autocor]	0	1
1	0.0129	0.0129	.16646	0.6833						
2	-0.0048	-0.0049	.18915	0.9098						
3	0.0316	0.0318	1.1948	0.7542						
4	-0.0573	-0.0583	4.493	0.3434						
5	0.0353	0.0374	5.7472	0.3316						
6	0.0592	0.0570	9.2748	0.1587						
7	-0.0645	-0.0629	13.474	0.0614						
8	0.0209	0.0180	13.913	0.0841						
9	-0.0083	-0.0084	13.982	0.1230						
10	-0.0203	-0.0115	14.398	0.1556						
11	-0.0192	-0.0316	14.77	0.1933						
12	-0.0043	0.0010	14.789	0.2532						
13	-0.0425	-0.0379	16.625	0.2170						
14	0.0164	0.0121	16.898	0.2617						
15	0.0047	0.0064	16.92	0.3237						
16	-0.0309	-0.0275	17.888	0.3305						
17	-0.0468	-0.0504	20.121	0.2681						
18	-0.0315	-0.0298	21.13	0.2729						
19	0.0194	0.0272	21.514	0.3091						
20	0.0131	0.0051	21.689	0.3576						

# Estacionarização de séries

## Operador de diferença

- A  $d^{\text{a}}$  diferença de  $y_t$  é dada por

$$\nabla^d y_t = (1 - B)^d y_t, \text{ d inteiro } \geq 1$$

- $d=1$ :  $\nabla y_t = (1 - B)y_t = y_t - y_{t-1}$
- $d=2$ :  $\nabla^2 y_t = (1 - B)^2 y_t = \nabla y_t - \nabla y_{t-1}$
- $d$ :  $\nabla^d y_t = (1 - B)^d y_t = \nabla^{d-1} y_t - \nabla^{d-1} y_{t-1}$

Em geral  $y_t \sim I(d)$ ,  $\nabla^d y_t \sim I(0)$

- Para séries sazonais usa-se a diferença sazonal  $s$

$$\nabla^s y_t = (1 - B^s)y_t = y_t - y_{t-s}$$

- A  $D^{\text{a}}$  diferença de uma série sazonal é

$$(\nabla^s)^D y_t = (1 - B^s)^D y_t,$$

geralmente  $D=1, 2$  é suficiente para obter estacionaridade sazonal

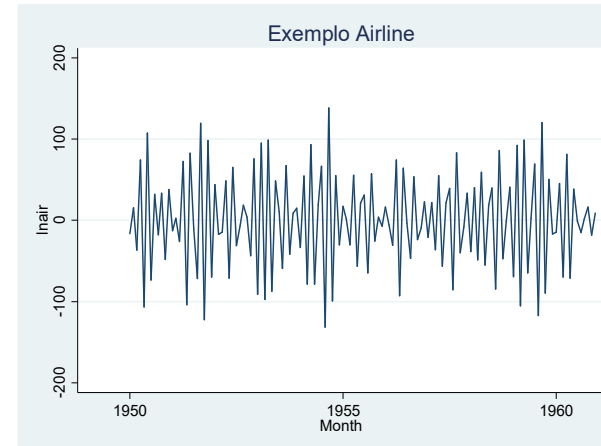
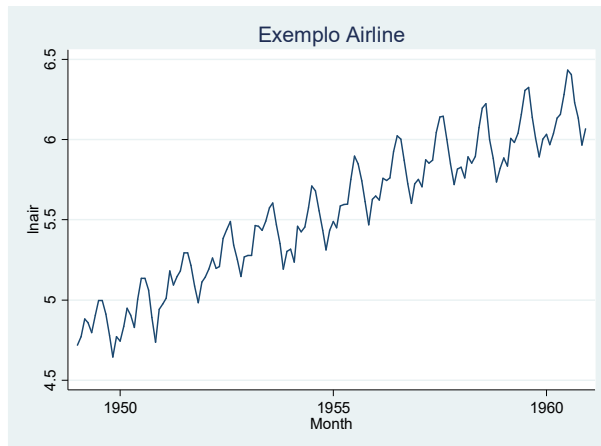
# Estacionarização de séries

## Operador de diferença com sazonalidade

Para séries sazonais utiliza-se a diferença sazonal  $s$

$$\nabla^s y_t = (1 - B^s)y_t = y_t - y_{t-s}$$

para obter estacionaridade sazonal



## 2 tipos de operadores de diferença (tendência e sazonalidade)

$$\nabla \nabla^s y_t = (y_t - y_{t-s}) - (y_{t-1} - y_{t-s-1}) = (y_t - y_{t-1}) - (y_{t-s} - y_{t-s-1})$$

Neste caso perdem-se  $s+1$  observações iniciais

# Modelo ARIMA

Uma série não estacionária e não sazonal pode ser descrita por um modelo ARMA integrado, ARIMA(p,d,q). O processo ARIMA(p,d,q), não sendo estacionário em níveis, torna-se estacionário com d diferenças

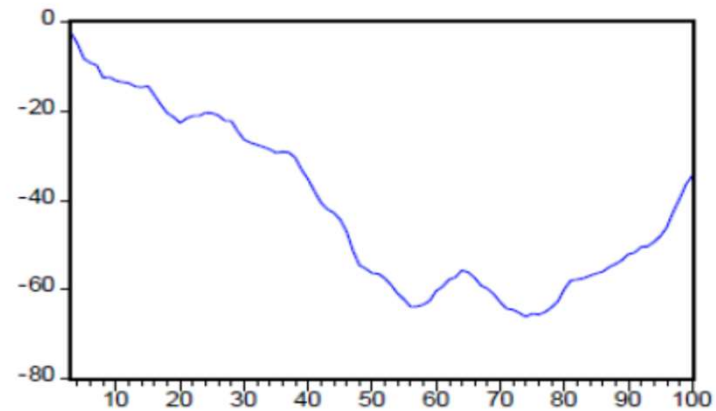
$$(1 - \rho_1 B - \rho_2 B^2 - \dots - \rho_p B^p)(1 - B)^d y_t = (1 + \theta_1 B + \dots + \theta_q B^q) \varepsilon_t$$
$$\rho_p(B)(1 - B)^d y_t = \theta_q(B) \varepsilon_t$$

Casos particulares:

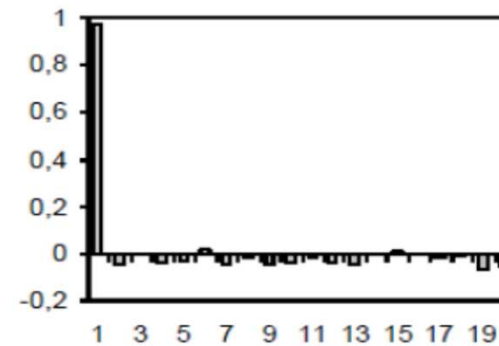
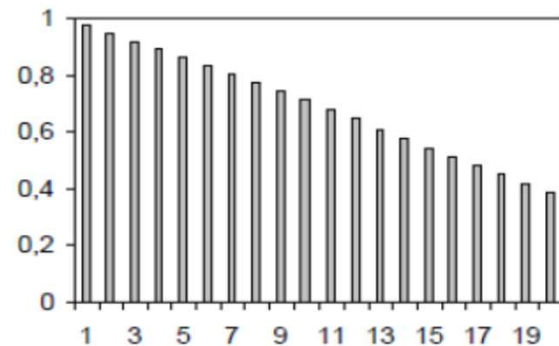
- ARIMA(1,1,1):  $(1 - \rho_1 B)(1 - B) y_t = (1 + \theta_1 B) \varepsilon_t$
- ARIMA(0,1,0):  $(1 - B)^d y_t = \varepsilon_t$  = Passeio aleatório
- ARIMA(1,0,1)=ARMA(1,1)
- ARIMA(1,1,0)=AR(1) integrado
- ARIMA(0,1,1)=MA(1) integrado

Note-se que  $\nabla^d y_t$  segue um processo ARMA (p, q). Por exemplo, ARIMA (2, 1, 3):  
 $\nabla y_t = \rho_1 \nabla y_{t-1} - \rho_2 \nabla y_{t-2} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3}$

# Modelo ARIMA



A simulated series from ARIMA(1,1,0) model:  $(1 - 0,75B)(1 - B)Y_t = \varepsilon_t$



ACF and PACF of the ARIMA(1,1,0) model:  $(1 - 0,75B)(1 - B)Y_t = \varepsilon_t$



# Modelo SARIMA

SARIMA(p,d,q)(P,D,Q)<sub>s</sub>

$$\begin{aligned} & (1 - \rho_1 B - \rho_2 B^2 - \dots - \rho_p B^p)(1 - P_1 B^{12} - \dots - P_p B^{P_s})(1 - B)^d y_t \\ & = (1 + \theta_1 B + \dots + \theta_q B^q)(1 + \Theta_1 B^{12} + \dots + \Theta_Q B^{Q_s}) \varepsilon_t \\ & \rho_p(B) P_P(B^s)(1 - B)^d (1 - B^s)^D y_t = \theta_q(B) \Theta_Q(B^s) \varepsilon_t \end{aligned}$$

- Modelo multiplicativos

Exemplos:

- SARIMA(1,0,1)(1,0,1)<sub>12</sub>

$$\begin{aligned} & (1 - \rho_1 B)(1 - P_1 B^{12}) y_t = (1 + \theta_1 B)(1 + \Theta_1 B^{12}) \varepsilon_t \\ & (1 - \rho_1 B - P_1 B^{12} + \rho_1 P_1 B^{13}) y_t = (1 + \theta_1 B + \Theta_1 B^{12} + \theta_1 \Theta_1 B^{13}) \varepsilon_t \\ & y_t - \rho_1 y_{t-1} - P_1 y_{t-12} + \rho_1 P_1 y_{t-13} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \Theta_1 \varepsilon_{t-12} + \theta_1 \Theta_1 \varepsilon_{t-13} \\ & y_t = \rho_1 y_{t-1} + P_1 y_{t-12} - \rho_1 P_1 y_{t-13} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \Theta_1 \varepsilon_{t-12} + \theta_1 \Theta_1 \varepsilon_{t-13} \end{aligned}$$

# Modelo SARIMA

- SARIMA(0,1,1)(0,1,1)<sub>12</sub>

$$\begin{aligned}\nabla\nabla_{12}y_t &= (1 + \theta_1 B)(1 + \Theta_1 B^{12})\varepsilon_t \\ \nabla\nabla_{12}y_t &= (1 + \theta_1 B + \Theta_1 B^{12} + \theta_1 \Theta_1 B^{13})\varepsilon_t \\ \nabla\nabla_{12}y_t &= \varepsilon_t + \theta_1 \varepsilon_{t-1} + \Theta_1 \varepsilon_{t-12} + \theta_1 \Theta_1 \varepsilon_{t-13}\end{aligned}$$

# Identificação, estimação, validação e selecção de modelos

- Identificação do modelo com base no correlograma, testes de estacionaridade, análise da ACF e PACF
- Estimação dos parâmetros do modelo
  - Método da máxima verosimilhança
  - OLS
    - AR(p)'s já estacionários
    - Para Ma's e ARMA's é necessário re-escrever como AR( $\infty$ )
- Validação: significancia individual e conjunta (OLS: testes t e F), teste de Ljung-Box para verificar se o residuo é ruído branco
- Selecção do modelo: escolhe-se o modelo que minimize AIC (Akaike Information Criterion) e/ou BIC (Schwarz Bayesian Information Criterion)

# Identificação do modelo

Em geral, com base no correlograma, verifica-se se a série tem tendência, sazonalidade, outliers, variância constante

1. Estabilizar a variância, via transformação logarítmica
2. Testar a estacionaridade, apurar qual o grau de diferenciação  $d$  necessário para a série ficar estacionária (tipicamente  $d=0, 1$  ou  $2$ ) e estacionarizar a série
3. Com base na ACF e na PACF da série  $I(0)$  identificar  $(p, q)$  e  $(P, Q)$ , nos lags não sazonais e sazonais, respectivamente. Tipicamente as ordens são menores ou iguais a 3

# Identificação do modelo

Figurinos teóricos das FAC e FACP dos modelos ARMA (Caiado (2016),

Modelo	FAC	FACP
AR( $p$ )	Decaimento exponencial ou sinusoidal amortecido para zero	Queda brusca para zero a partir do $lag\ p + 1$
MA( $q$ )	Queda brusca para zero a partir do $lag\ q + 1$	Decaimento exponencial ou sinusoidal amortecido para zero
ARMA( $p, q$ )	Decaimento exponencial ou sinusoidal amortecido para zero	Decaimento exponencial ou sinusoidal amortecido para zero
SAR( $P$ )	Decaimento exponencial ou sinusoidal amortecido para zero sobre os $lags\ s, 2s, \dots$	Queda brusca para zero a partir do $lag\ (P + 1)s$
SMA( $Q$ )	Queda brusca para zero a partir do $lag\ (Q + 1)s$	Decaimento exponencial ou sinusoidal amortecido para zero sobre os $lags\ s, 2s, \dots$
SARMA( $P, Q$ )	Decaimento exponencial ou sinusoidal amortecido para zero sobre os $lags\ s, 2s, \dots$	Decaimento exponencial ou sinusoidal amortecido para zero sobre os $lags\ s, 2s, \dots$
SARMA( $p, q$ )( $P, Q$ ) $_s$	Decaimento exponencial ou sinusoidal amortecido para zero	Decaimento exponencial ou sinusoidal amortecido para zero

# Seleccção do modelo

Minimiza-se

- AIC (1974):  $-2\ln L + 2m$
- BIC (1978):  $-2\ln L + m\ln N$

onde  $L$  é a função de verosimilhança

# Previsão

No momento  $T$ , com base nas observações  $y_T, y_{T-1}, y_{T-2}, \dots, y_1$  pretende-se prever  $y_{T+h}$

A previsão optima para  $y_{T+h}$  será a que minimiza o erro quadrático médio  $E(e_{T+h}^2)$ , correspondendo ao valor esperado de  $y_{T+h}$  condicional na informação disponível:

$$\hat{y}_{T+h} = E(y_{T+h} | y_T, y_{T-1}, \dots, y_1)$$

onde  $T$  é a origem da previsão e  $h$  é o horizonte de previsão

A previsão pode ser feita  $m$  passos à frente ou passo a passo

# Previsão: AR(1)

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

A previsão optima é

$$\hat{y}_{T+h} = E(\rho y_{T+h-1} + \varepsilon_{T+h} | y_T, y_{T-1}, \dots, y_1) = \rho y_{T+h-1}$$

## h passos à frente

- Um passo à frente (h=1):  $\hat{y}_{T+1} = \rho y_T$
- Para  $h > 1$ , tem-se recursivamente:
  - h=2:  $\hat{y}_{T+2} = \rho \hat{y}_{T+1} = \rho^2 y_T$
  - h=3:  $\hat{y}_{T+3} = \rho \hat{y}_{T+2} = \rho^3 y_T$
  - Genericamente:  $\hat{y}_{T+h} = \rho^h y_T$ 
    - Dado que  $|\rho| < 1$ , quando  $h \rightarrow \infty$ ,  $\hat{y}_{T+h}$  converge para 0 (de facto a previsão converge para a media de  $y_t$ , que se provou anteriormente ser 0)
    - Para um passeio aleatório, a melhor previsão é  $\hat{y}_{T+h} = y_T$ , o ultimo valor disponivel para a série



# Previsão: AR(1)

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

## Passo a passo

Dado que em  $t+h$  está disponível a observação para  $t+h-1$ , não é necessário considerar valores estimados na forma recursiva

- $\hat{y}_{T+1} = \rho y_T$
- $\hat{y}_{T+2} = \rho y_{T+1}$
- $\hat{y}_{T+3} = \rho y_{T+2}$
- ...
- $\hat{y}_{T+h} = \rho y_{T+h-1}$

# Previsão: AR(1) com drift

$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t$$

onde  $E(y_t) = \frac{\alpha}{1-\rho}$

A previsão optima é

$$\hat{y}_{T+h} = E(\alpha + \rho y_{T+h-1} + \varepsilon_{T+h} | y_T, y_{T-1}, \dots, y_1) = \alpha + \rho y_{T+h-1}$$

## h passos à frente

- $\hat{y}_{T+1} = \alpha + \rho y_T$
- $\hat{y}_{T+2} = \alpha + \rho \hat{y}_{T+1} = \alpha + \rho \alpha + \rho^2 y_T = \alpha(1 + \rho) + \rho^2 y_T$
- $\hat{y}_{T+3} = \alpha + \rho \hat{y}_{T+2} = \alpha + \rho \alpha + \rho^2 \alpha + \rho^3 y_T = \alpha(1 + \rho + \rho^2) + \rho^3 y_T$
- $\hat{y}_{T+h} = (\rho^{h-1} + \rho^{h-2} \dots + 1)\alpha + \rho^h y_T = \frac{\alpha}{1-\rho^2} + \rho^h y_T$
- Dado que  $|\rho| < 1$ , quando  $h \rightarrow \infty$ ,  $\hat{y}_{T+h}$  converge para a media de  $y_t, \frac{\alpha}{1-\rho^2}$
- Para um passeio aleatório com drift  $\hat{y}_{T+h} = h\alpha + y_T$

# Previsão: MA(1)

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

$$\text{com } E(\varepsilon_{T+h} | y_T, \dots, y_1) = \begin{cases} 0 & \text{se } h > 0 \\ \varepsilon_{T+h} & \text{se } h \leq 0 \end{cases}$$

## h passos à frente

- $\hat{y}_{T+1} = E(\varepsilon_{T+1} + \theta \varepsilon_T | y_T, y_{T-1}, \dots, y_1) = E(\varepsilon_{T+1} + \theta \varepsilon_T | \varepsilon_T, \varepsilon_{T-1}, \dots, \varepsilon_1) = 0 + \theta E(\varepsilon_T | \varepsilon_T, \varepsilon_{T-1}, \dots, \varepsilon_1) = \theta \varepsilon_T$
- $h > 1: \hat{y}_{T+h} = E(\varepsilon_{T+h} + \theta \varepsilon_{T+h-1} | y_T, y_{T-1}, \dots, y_1) = 0$

# Previsão: ARMA(1,1)

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

## h passos à frente

- $\hat{y}_{T+1} = E(\rho y_T + \varepsilon_{T+1} + \theta \varepsilon_T | y_T, y_{T-1}, \dots, y_1) = \rho y_T + \theta \varepsilon_T$
- Para  $k > 1$ , tem-se recursivamente
  - $\hat{y}_{T+2} = E(\rho y_{T+1} + \varepsilon_{T+2} + \theta \varepsilon_{T+1} | y_T, y_{T-1}, \dots, y_1) = E(\rho(\rho y_T + \varepsilon_{T+1} + \theta \varepsilon_T) + \varepsilon_{T+2} + \theta \varepsilon_{T+1} | y_T, y_{T-1}, \dots, y_1) = \rho^2 y_T + \rho \theta \varepsilon_T$
  - ...
  - $\hat{y}_{T+h} = \rho^h y_T + \rho^{h-1} \theta \varepsilon_T$
- Dado que  $|\rho| < 1$ , quando  $h \rightarrow \infty$ ,  $\hat{y}_{T+h}$  converge para a media de  $y_t$ , que é zero

# Previsão: ARIMA(0,1,1)

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

## **h passos à frente**

- $\hat{y}_{T+1} = E(y_T + \varepsilon_{T+1} + \theta \varepsilon_T | y_T, y_{T-1}, \dots, y_1) = y_T + \theta \varepsilon_T$
- Para  $h > 1$ , tem-se recursivamente
  - $\hat{y}_{T+2} = E(y_{T+1} + \varepsilon_{T+2} + \theta \varepsilon_{T+1} | y_T, y_{T-1}, \dots, y_1) = E(y_T + \varepsilon_{T+1} + \theta \varepsilon_T + \varepsilon_{T+2} + \theta \varepsilon_{T+1} | y_T, y_{T-1}, \dots, y_1) = y_T + \theta \varepsilon_T = \hat{y}_{T+1}$
  - ...
  - $\hat{y}_{T+h} = \hat{y}_{T+h-1}$

# Previsão: SARIMA(0,0,0) (0,1,1)12

Previsão no fim do ano T, para um qualquer mês j do ano seguinte:

$$\nabla_{12}y_t = (1 + \Theta_1 B^{12})\varepsilon_t$$

$$\nabla_{12}y_t = \varepsilon_t + \Theta_1\varepsilon_{t-12}$$

$$y_{t+j} = y_{t+j-1} + \varepsilon_{t+j} + \Theta_1\varepsilon_{t+j-12}$$

## h passos à frente

- $j \leq 12$ :  $\hat{y}_{T+j} = E(y_{T+j-1} + \varepsilon_{T+j} + \Theta_1\varepsilon_{T+j-1} | y_T, \dots, y_1) = y_{T+j-1} + \Theta_1\varepsilon_{T+j-1}$
- $\hat{y}_{T+12+j} = \hat{y}_{T+j}$ : as previsões para todos os meses do segundo ano são iguais às do primeiro ano

# Previsão por intervalo

## Estrutura do intervalo

$$\left( \hat{y}_{T+h} - z^{\alpha/2} \sqrt{\hat{\sigma}_{e_{T+h}}^2}; \hat{y}_{T+h} + z^{\alpha/2} \sqrt{\hat{\sigma}_{e_{T+h}}^2} \right)$$

onde  $e_{T+h} = y_{T+h} - \hat{y}_{T+h}$  e  $\hat{\sigma}_{e_{T+h}}^2 = \hat{\sigma}^2 + \hat{\sigma}_{\hat{y}_{T+h}}^2$

## AR(1), h passos à frente

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

- $e_{T+h} = \rho y_{T+h-1} + \varepsilon_{T+h} - \rho^h y_T$  (substituindo repetidamente)
- $e_{T+h} = \rho^h y_T + \varepsilon_{T+h} + \rho \varepsilon_{T+h-1} + \dots + \rho^{h-1} \varepsilon_{T+1} - \rho^h y_T$
- $E(e_{T+h}^2) = (1 + \rho^2 + \rho^4 + \dots + \rho^{2h-2}) \sigma^2$ 
  - Reduz-se a  $\sigma^2$  para 1 passo à frente

# Previsão por intervalo

## ARMA(1,1), h passos à frente

- $e_{T+h} = \rho y_{T+h-1} + \varepsilon_{T+h} - \rho^h y_T$  (substituindo repetidamente)
- $e_{T+h} = \rho^h y_T + \varepsilon_{T+h} + \rho \varepsilon_{T+h-1} + \dots + \rho^{h-1} \varepsilon_{T+1} - \rho^h y_T$
- $E(e_{T+h}^2) = (1 + \rho^2 + \rho^4 + \dots + \rho^{2h-2})\sigma^2$ 
  - Reduz-se a  $\sigma^2$  para 1 passo à frente

ARMA(p,q), ver Harris & Sollis (2005), p.14



# Modelos com regressores

## Modelos autorregressivos de defasamentos distribuídos ( $ADL(p, q)$ )

- Os regressores são  $p$  lags da variável dependente  $y_t$  e  $q$  lags da variável explicativa  $x_t$ :

$$y_t = \delta_0 + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \dots + \rho_p y_{t-p} + \beta_0 x_t + \beta_1 x_{t-1} + \dots + \beta_q x_{t-q} + u_t$$

- Modelo AR( $p$ ):  $\beta_0 = \beta_1 = \dots = \beta_q = 0$
- Modelo sem componente dinâmica:  $\rho_1 = \rho_2 = \dots = \rho_p = 0$

# Modelos com regressores

## Dummies

Sazonalidade: conforme a sazonalidade  $s$  (12, 4, 2, para dados mensais, trimestrais, semestrais, respectivamente), incluem-se  $(s - 1)$  dummies. Por exemplo, para trimestres:

$$D1_t = \begin{cases} 1 & \text{se observação em trim. 1} \\ 0 & \text{observação noutra trim.} \end{cases}, D2_t = \begin{cases} 1 & \text{se observação em trim. 2} \\ 0 & \text{observação noutra trim.} \end{cases}$$
$$D3_t = \begin{cases} 1 & \text{se observação em trim. 3} \\ 0 & \text{observação noutra trim.} \end{cases}$$

Step dummy: capta o efeito de um acontecimento que muda o rumo da serie de forma permanente :

$$D_t = \begin{cases} 1 & \text{a partir do momento do acontecimento} \\ 0 & \text{antes do acontecimento} \end{cases}$$

Impulse dummy – capta o efeito de um acontecimento que muda momentaneamente a série:

$$D_t = \begin{cases} 1 & \text{no periodo de ocorrência} \\ 0 & \text{nos restantes momentos} \end{cases}$$

# Modelos com regressores

## Testes de especificação

### Forma funcional

- A incorrecta especificação do modelo gera estimadores OLS inconsistentes
- Teste RESET
  - $H_0$ : *Forma funcional correcta*

# Modelos com regressores

## Implementação

1. Estimar o modelo de interesse,  $Y_t = \beta_0 + \beta_1 X_{t1} + \dots + \beta_k X_{tk} + u_t$  e obter  $\hat{Y}^2 = (X\hat{\beta})^2$ ,  $\hat{Y}^3 = (X\hat{\beta})^3$
2. Estimar a regressão auxiliar:  $Y_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik} + \gamma_1 \hat{Y}_i^2 + \gamma_2 \hat{Y}_i^3 \dots + v$
3. Fazer teste F para  $H_0: \gamma_1 = \dots = \gamma_k = 0$

Solução para quando a forma funcional é rejeitada: re-especificar o modelo. Nota: a regressão auxiliar, não é candidata a modelo alternativo

# Modelos com regressores

## Testes de especificação

### Heteroscedasticidade

- Variância condicional do erro (skedastic function):  $Var(u|X)$  em lugar de ser constante,  $Var(u|X) = \sigma^2$  (homocedasticidade), depende de  $X$ ,  $Var(u|X) = \sigma^2 h(X)$
- Estimadores OLS consistentes mas variâncias e, conseqüentemente, inferência, invalidas
- Teste de Breusch-Pagan
  - $H_0: Var(u|X) = \sigma^2$  (homoscedasticidade)

# Modelos com regressores

Implementação do teste BP:

1. Estimar o modelo de interesse,  $Y_t = \beta_0 + \beta_1 X_{t1} + \dots + \beta_k X_{tk} + u_t$  e obter os resíduos  $\hat{u}_t$
2. Estimar a regressão auxiliar:  $\hat{u}_t^2 = \gamma_0 + \gamma_1 X_{t1} + \dots + \gamma_k X_{tk} + e$
3. Fazer teste f ou LM ( $LM = NR_{\hat{u}_t^2}^2 \sim X_k^2$ ) à significância conjunta dos regressores

Solução para quando a heteroscedasticidade existe: utilizar a versão robusta da variância

# Modelos com regressores

## Testes de especificação

### Autocorrelação

- Forma mais comum:

$$y_t = \beta_0 + \beta_1 x_{t1} + \dots + \beta_k x_{tk} + u_t$$

$$u_t = \rho u_{t-1} + \varepsilon_t, \rho \neq 0$$

$$\varepsilon_t \sim \text{ruído branco}$$

- Estimadores OLS consistentes mas variâncias e, conseqüentemente, inferência, inválidas
- Teste de Breusch-Godfrey
  - $H_0: \rho = 0$  (sem autocorrelação)

# Modelos com regressores

Implementação do teste BG:

1. Estimar o modelo de interesse,  $Y_t = \beta_0 + \beta_1 X_{t1} + \dots + \beta_k X_{tk} + u_t$  e obter os resíduos  $\hat{u}_t$
2. Estimar a regressão auxiliar:  $\hat{u}_t = \gamma_0 + \gamma_1 X_{t1} + \dots + \gamma_k X_{tk} + \rho \hat{u}_{t-1} + \varepsilon_t$
3. Fazer teste LM à significância conjunta dos regressores:

$$LM = (N - p)R_{\hat{u}}^2 \sim X_p^2$$

Generalização para autocorrelação AR(p): no passo 3:  $\hat{u}_t = \gamma_0 + \gamma_1 X_{t1} + \dots + \gamma_k X_{tk} + \rho_1 \hat{u}_{t-1} + \dots + \rho_p \hat{u}_{t-p} + \varepsilon_t$

Solução para quando a autocorrelação existe: adicional lags ao modelo até ter um modelo dinamicamente completo (sem autocorrelação). Wooldridge () considera importante para efeitos de previsão que o modelo seja dinamicamente completo



# Modelos com regressores

## Modelos para vectores autoregressivos (VAR)

- $VAR(p)$  é composto por equações  $ADL(p, p)$ , cada uma com uma variável dependente diferente e incluindo como regressores os lags das restantes variáveis do modelo

$$\begin{cases} y_t = \delta_0 + \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + \gamma_1 x_{t-1} + \dots + \gamma_p x_{t-p} + \dots + u_t \\ x_t = \delta_0 + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \varphi_1 x_{t-1} + \dots + \varphi_p x_{t-p} + \dots + v_t \\ \dots \end{cases}$$

- Cada equação é estimada separadamente

# Modelos com regressores

## VAR e causalidade à Granger

- A variável  $x_t$  causa à Granger  $y_t$  se, tendo em conta a informação do passado de  $y_t$ , o passado de  $x_t$  é relevante para prever  $y_t$
- Simples testes à significância conjunta de lags:
  - $H_0: \gamma_1 = \dots = \gamma_p = 0$  ( $x_t$  não causa  $y_t$  à Granger)
  - $H_0: \beta_1 = \dots = \beta_p = 0$  ( $y_t$  não causa  $x_t$  à Granger)

# Modelos com regressores

## VAR e previsão

Previsão h passos à frente: o lag mais recente no VAR é dado por h, de modo a que existam observações disponíveis para prever

$$\cdot \quad h=1 \begin{cases} y_t = \delta_0 + \alpha_1 y_{t-1} + \dots + \gamma_1 x_{t-1} + \dots \\ x_t = \delta_0 + \beta_1 y_{t-1} + \dots + \varphi_1 x_{t-1} + \dots \\ \dots \end{cases} \rightarrow \begin{cases} y_{T+1} = \delta_0 + \alpha_1 y_T + \dots + \gamma_1 x_T + \dots \\ x_{T+1} = \delta_0 + \beta_1 y_T + \dots + \varphi_1 x_T + \dots \\ \dots \end{cases}$$

$$\cdot \quad h=2 \begin{cases} y_t = \delta_0 + \alpha_1 y_{t-2} + \dots + \gamma_1 x_{t-2} + \dots \\ x_t = \delta_0 + \beta_1 y_{t-2} + \dots + \varphi_1 x_{t-2} + \dots \\ \dots \end{cases} \rightarrow \begin{cases} y_{T+2} = \delta_0 + \alpha_1 y_T + \dots + \gamma_1 x_T + \dots \\ x_{T+2} = \delta_0 + \beta_1 y_T + \dots + \varphi_1 x_T + \dots \\ \dots \end{cases}$$

# Modelos com regressores

## VAR e previsão

Previsão passo a passo: independente do lag mais recente no VAR, pois substituem-se previsões quando não existem observações disponíveis para prever

$$\begin{aligned} \cdot \quad h=1 \text{ com } & \begin{cases} y_t = \delta_0 + \alpha_1 y_{t-1} + \dots + \gamma_1 x_{t-1} + \dots \\ x_t = \delta_0 + \beta_1 y_{t-1} + \dots + \varphi_1 x_{t-1} + \dots \\ \dots \end{cases} \rightarrow \begin{cases} y_{T+1} = \delta_0 + \alpha_1 y_T + \dots + \gamma_1 x_T + \dots \\ x_{T+1} = \delta_0 + \beta_1 y_T + \dots + \varphi_1 x_T + \dots \\ \dots \end{cases} \\ \cdot \quad h=2 \text{ com } & \begin{cases} y_t = \delta_0 + \alpha_1 y_{t-1} + \dots + \gamma_1 x_{t-1} + \dots \\ x_t = \delta_0 + \beta_1 y_{t-1} + \dots + \varphi_1 x_{t-1} + \dots \\ \dots \end{cases} \rightarrow \begin{cases} y_{T+2} = \delta_0 + \alpha_1 y_{T+1} + \dots + \gamma_1 x_{T+1} + \dots \\ x_{T+2} = \delta_0 + \beta_1 y_{T+1} + \dots + \varphi_1 x_{T+1} + \dots \\ \dots \end{cases} \\ \text{usando resultados do primeiro passo} & \rightarrow \begin{cases} y_{T+2} = \delta_0 + \alpha_1 \hat{y}_{T+1} + \dots + \gamma_1 \hat{x}_{T+1} + \dots \\ x_{T+2} = \delta_0 + \beta_1 \hat{y}_{T+1} + \dots + \varphi_1 \hat{x}_{T+1} + \dots \\ \dots \end{cases} \end{aligned}$$

# Quebras de estrutura

Ocorrem quando um acontecimento, em data conhecida ou desconhecida, altera o efeito das variáveis explicativas processo

Testes para detecção:

- Data conhecida: Teste Chow
- Data desconhecida: Teste de Quandt LR (QLR) / sup-Wald

# Quebras de estrutura

## Teste de Chow

- Traduz-se o antes e o após a quebra por uma variável dummy

$$D = \begin{cases} 1 & \text{no momento e após a quebra} \\ 0 & \text{antes da quebra)} \end{cases}$$

- Estima-se o modelo original de forma 'duplicada':

$$y_t = \theta_0 + \theta_1 x_{1t} + \dots + \theta_k x_{kt} + \gamma_0 D + \gamma_1 D x_{1t} + \dots + \gamma_k D x_{kt} + v$$

- Aplica-se um teste F para a significância das variáveis envolvendo  $D$ :

$$H_0: \gamma_1 = \dots = \gamma_k = 0 \text{ (sem quebra de estrutura)}$$

$$H_1: \text{Não } H_0 \text{ (com quebra de estrutura)}$$

# Quebras de estrutura

## Teste QLR

- Suponha-se que a data da quebra,  $\tau$ , é desconhecida
- Calculam-se múltiplos testes de Chow no intervalo  $\tau_0 < \tau < \tau_1$ .  
Tipicamente, faz-se um trimming para 70% da amostra, excluindo os primeiros e os últimos 15% de observações
- Escolhe-se o máximo e compara-se com o valor crítico, que é recolhido de tabela relativa ao teste:

Number of Restrictions ( $q$ )	10%	5%	1%
1	7.12	8.68	12.16
2	5.00	5.86	7.78
3	4.09	4.71	6.02
4	3.59	4.09	5.12
5	3.26	3.66	4.53
6	3.02	3.37	4.12
7	2.84	3.15	3.82
8	2.69	2.98	3.57
9	2.58	2.84	3.38
10	2.48	2.71	3.23