

# Séries temporais - 2011

Prof. Wagner Hugo Bonat

Laboratório de Estatística e Geoinformação  
Universidade Federal do Paraná

20 de maio de 2011

# Sumário

1 Introdução

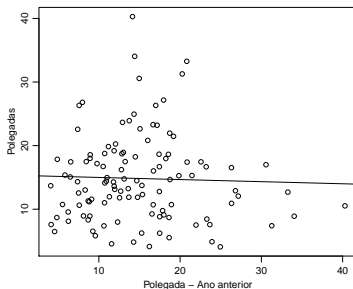
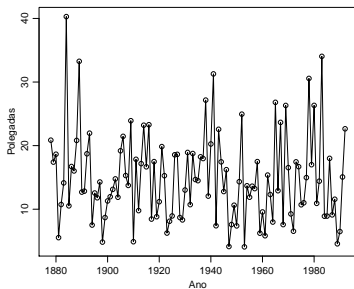
2 Referências

# Introdução

- Dados sequenciais no tempo são comuns em diversas áreas;
- Objetivos em séries temporais são geralmente dois:
  - 1 Entender ou modelar o mecanismo estocástico gerador da série;
  - 2 Predizer (predict or forecast) valores futuros da série.
- A principal diferença na análise de séries temporais é que usualmente não podemos assumir independência entre as observações;
- A idéia é propor modelos que incorporem a dependência.

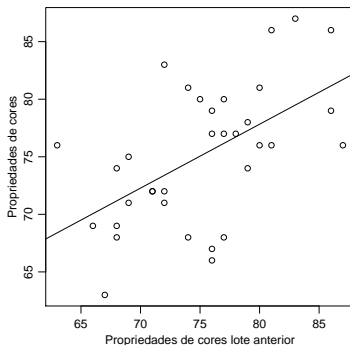
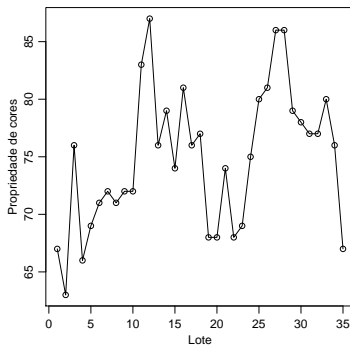
# Exemplo 1 - Precipitação anual em Los Angeles

- Objetivo - Identificar se anos consecutivos estão correlacionados de alguma forma.



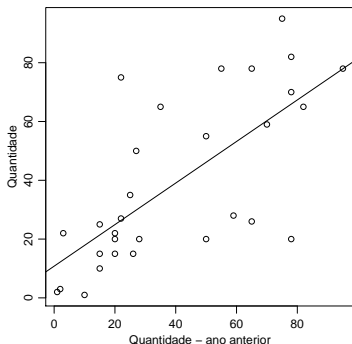
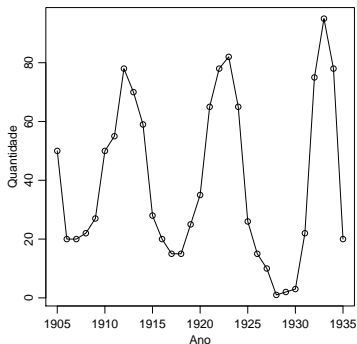
## Exemplo 2 - Processo químico industrial

- Propriedades de cores em lotes consecutivos de determinado material.



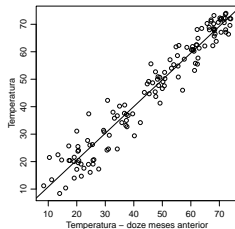
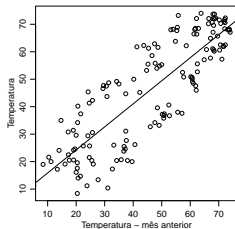
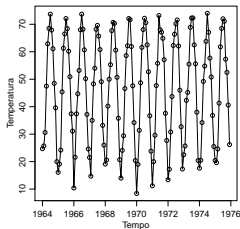
# Exemplo 3 - Abundância anual de lebres Canadenses

- Série anual com uma forte dependência.



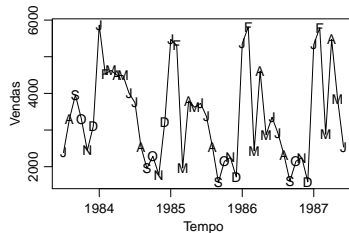
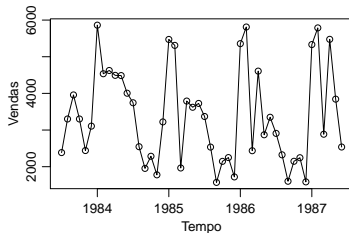
# Exemplo 4 - Temperatura média mensal

- Série mensal com forte sazonalidade.



# Exemplo 5 - Vendas mensais de filtro de óleo

- Sazonalidade não explícita.

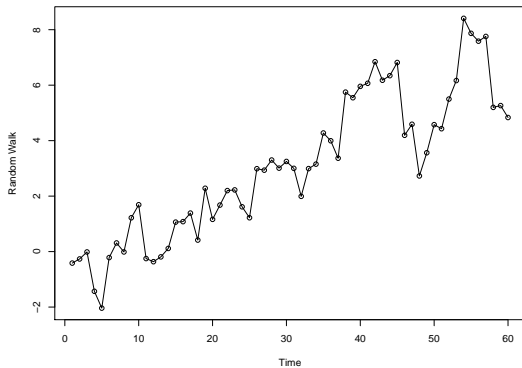




# Construindo modelos

- Passos para a construção de um modelo:
  - ① Especificação do modelo (olhar gráficos, estatísticas descritivas, conhecimento subjetivo da série, parcimônia, ...);
  - ② Ajuste do modelo (Minimos quadrados, máxima verossimilhança, momentos, Bayes, ...) ;
  - ③ Diagnóstico do modelo ( Testar se as suposições do modelo estão satisfeitas, quanto bom é o ajuste, ...).
- Se tudo der certo... acabou !!! Se não começa de novo ...
- Podemos usar o modelo para entender o comportamento da série e prever valores futuros.

# Simulando um Random Walk



# Simulando um random walk

```
> random.walk <- function(amostra, sigma) {  
+   Y <- c()  
+   Y[1] <- rnorm(1, mean = 0, sd = sigma)  
+   for (i in 2:amostra) {  
+     Y[i] <- Y[i - 1] + rnorm(1, mean = 0, sd = sigma)  
+   }  
+   return(Y)  
+ }  
> y <- random.walk(100, sigma = 0.5)  
> plot(y, type = "o", ylab = "Random Walk")
```

# Simulando - Média Móvel

```
> media.movel <- function(amostra, sigma) {  
+   Y <- c()  
+   e <- rnorm(amostra, 0, sd = sigma)  
+   Y[1] <- e[1]  
+   for (t in 2:amostra) {  
+     Y[t] <- (e[t] + e[t - 1])/2  
+   }  
+   return(Y)  
+ }
```

# Minimos quadrados

```
> data(rwalk)
> model1 <- lm(rwalk ~ time(rwalk))
> summary(model1)
```

```
Call:
lm(formula = rwalk ~ time(rwalk))
```

Residuals:

Min	1Q	Median	3Q	Max
-2.70045	-0.79782	0.06391	0.63064	2.22128

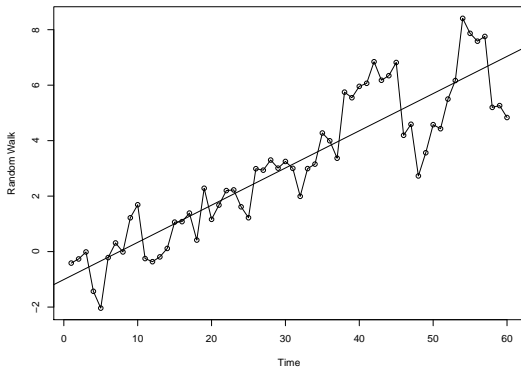
Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-1.007888	0.297245	-3.391	0.00126 **
time(rwalk)	0.134087	0.008475	15.822	< 2e-16 ***

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1.137 on 58 degrees of freedom
Multiple R-squared:  0.8119, Adjusted R-squared:  0.8086
F-statistic: 250.3 on 1 and 58 DF,  p-value: < 2.2e-16
```

# Minimos quadrados - Random Walk



# Tendência sazonal

```
> data(tempdub)
> month. = season(tempdub)
> model2 <- lm(tempdub ~ month. - 1)
> summary(model2)
```

Call:

```
lm(formula = tempdub ~ month. - 1)
```

Residuals:

```
    Min       1Q   Median       3Q      Max
-8.2750 -2.2479  0.1125  1.8896  9.8250
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
month.January	16.608	0.987	16.83	<2e-16 ***
month.February	20.650	0.987	20.92	<2e-16 ***
month.March	32.475	0.987	32.90	<2e-16 ***
month.April	46.525	0.987	47.14	<2e-16 ***
month.May	58.092	0.987	58.86	<2e-16 ***
month.June	67.500	0.987	68.39	<2e-16 ***
month.July	71.717	0.987	72.66	<2e-16 ***
month.August	69.333	0.987	70.25	<2e-16 ***
month.September	61.025	0.987	61.83	<2e-16 ***
month.October	50.975	0.987	51.65	<2e-16 ***
month.November	36.650	0.987	37.13	<2e-16 ***
month.December	23.642	0.987	23.95	<2e-16 ***

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 3.419 on 132 degrees of freedom

Multiple R-squared: 0.9957, Adjusted R-squared: 0.9953

# Tendência sazonal modelo via coseno

```
> har. = harmonic(tempdub, 1)
> model4 <- lm(tempdub ~ har.)
> summary(model4)
```

```
Call:
lm(formula = tempdub ~ har.)
```

Residuals:

Min	1Q	Median	3Q	Max
-11.1580	-2.2756	-0.1457	2.3754	11.2671

Coefficients:

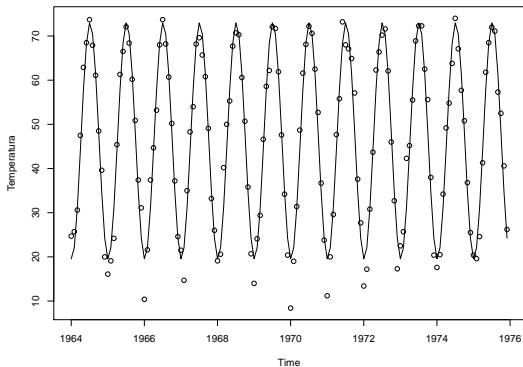
	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	46.2660	0.3088	149.816	< 2e-16 ***
har.cos(2*pi*t)	-26.7079	0.4367	-61.154	< 2e-16 ***
har.sin(2*pi*t)	-2.1697	0.4367	-4.968	1.93e-06 ***

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

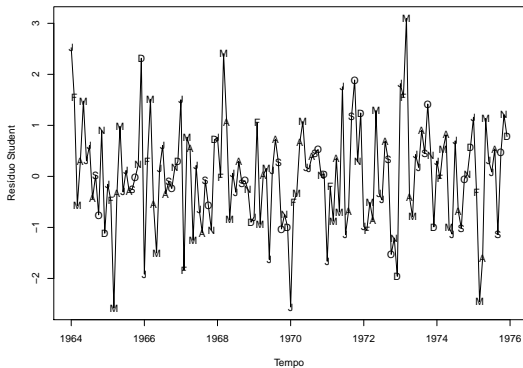
```
Residual standard error: 3.706 on 141 degrees of freedom
Multiple R-squared:  0.9639, Adjusted R-squared:  0.9634
F-statistic: 1882 on 2 and 141 DF,  p-value: < 2.2e-16
```



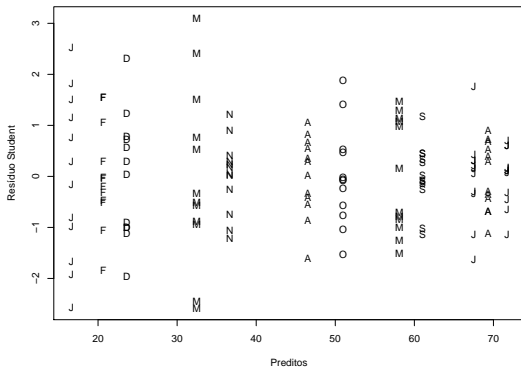
# Tendência sazonal modelo via coseno



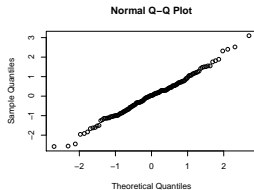
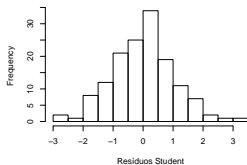
# Análise de resíduos - Resíduos versus Tempo



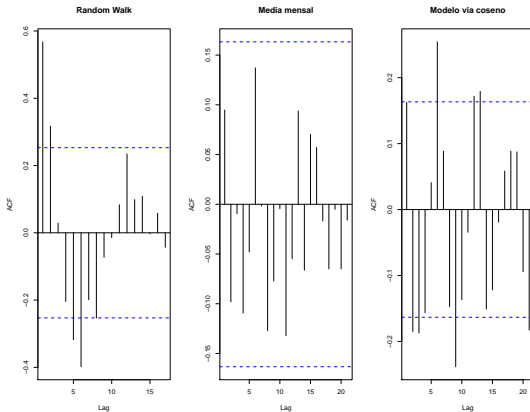
# Análise de resíduos - Resídeos versus Valores Preditos



# Análise de resíduos - Normalidade



# Função de autocorrelação amostral



# Exercícios

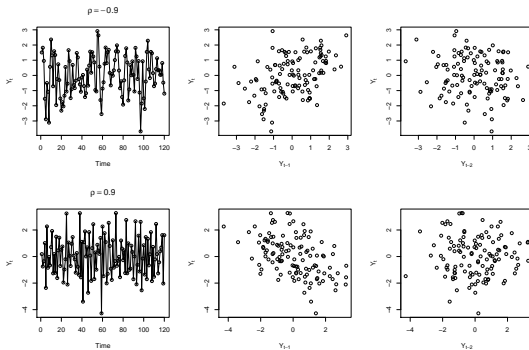
- Baixe os três conjuntos de dados da página do curso;
- Cada conjunto foi simulado de acordo com o seguinte modelo

$$\mu_t = X\beta + e_t$$

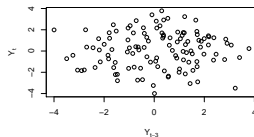
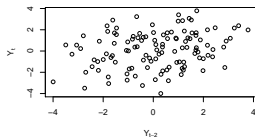
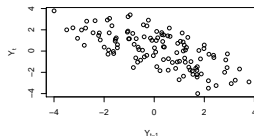
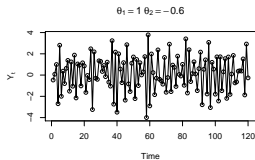
$X\beta = ?$  e  $e_t \sim ?(\sigma^2)$

- Identifique a forma da função de média determinística (caso exista) e a estrutura dos  $e_t$ 's.

# Médias móveis - Primeira ordem



# Médias móveis - Segunda ordem





# Referências

Cryer J. D. , Kung-Sik, C. (2008). Times Series Analysis.**Springer**.

R Development Core Team. (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, url <http://www.R-project.org>.