

# CE063 - Lista de exercícios da aula 4

Paulo Justiniano Ribeiro Jr e Elias T. Krainski

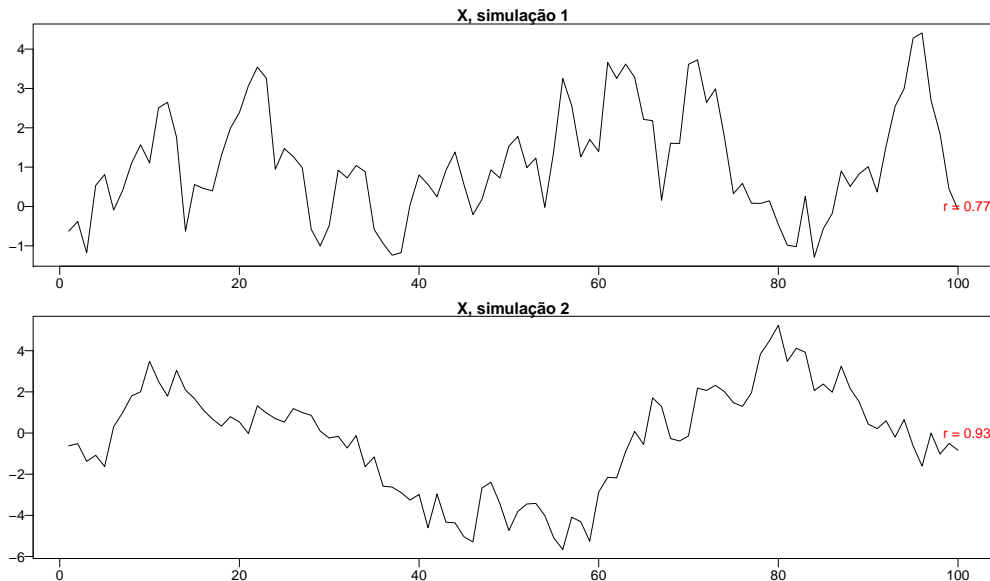
Março, 2019

1. Considere que  $X_1, X_2, \dots, X_n$  é uma série temporal com

- $X_1 \sim N(0, \sigma^2)$ , isto é,  $X_1 = 0 + e_1$ ; e
- $X_i \sim N(\rho X_{i-1}, \sigma^2)$ , para  $i = 2, \dots, n$ , isto é,  $X_i = \rho X_{i-1} + e_i$ .
- $e_i \sim N(0, \sigma^2)$ ,  $i = 1, 2, \dots, n$ .

Simulando desse modelo duas vezes e visualizando:

```
sigma <- 1; rho <- 0.9; n <- 100; x <- numeric(n)
par(mfrow=c(2,1), mar=c(2,2,1,0.5), mgp=c(1.5,0.5,0), las=1)
for (sim in 1:2) {
  x[1] <- rnorm(1, 0, sigma)
  for (i in 2:n)
    x[i] <- rnorm(1, rho*x[i-1], sigma)
  plot.ts(x, main=paste('X, simulação', sim), xlab='', ylab='')
  text(n+1, 0, paste('r =', format(cor(x[2:n], x[1:(n-1)]), dig=2)), col=2, lwd=2)
}
```



Observamos que o coeficiente de correlação empírica,  $r$ , entre valores de  $x_i$  e  $x_{i-1}$  em cada uma das séries simuladas é próximo de  $\rho=0.9$ , mas não necessariamente igual.

- a. (Autocorrelação). Se  $\text{Cov}(aX, bY) = ab\text{Cov}(X, Y)$ , temos que  $\text{Cov}(X_{i-1}, X_i) = \text{Cov}(X_{i-1}, \rho X_{i-1} + e_i) = \rho \text{Var}(X)$ . Também  $\text{Cor}(X, Y) = \text{Cov}(X, Y) / (\sqrt{\text{Var}(X)} \sqrt{\text{Var}(Y)})$ , temos que  $\text{Cor}(X_{i-1}, X_i) = \text{Cor}(X_{i-1}, \rho X_{i-1} + e_i) = \text{Cor}(X_{i-1}, \rho X_{i-1}) = \rho \text{Var}(X) / (\sqrt{\text{Var}(X)} \sqrt{\text{Var}(X)}) = \rho \text{Var}(X) / \text{Var}(X) = \rho$ . Usando esses resultados, calcule a correlação entre  $X_{i-2}$  e  $X_i$ .
- b. (Distribuição teórica  $t$ , sob  $H_0 : \rho = 0$ , e distribuições de  $t$  observadas). Se a correlação entre  $X$  e  $Y$ , duas variáveis aleatórias com distribuição Normal, é zero, então a estatística

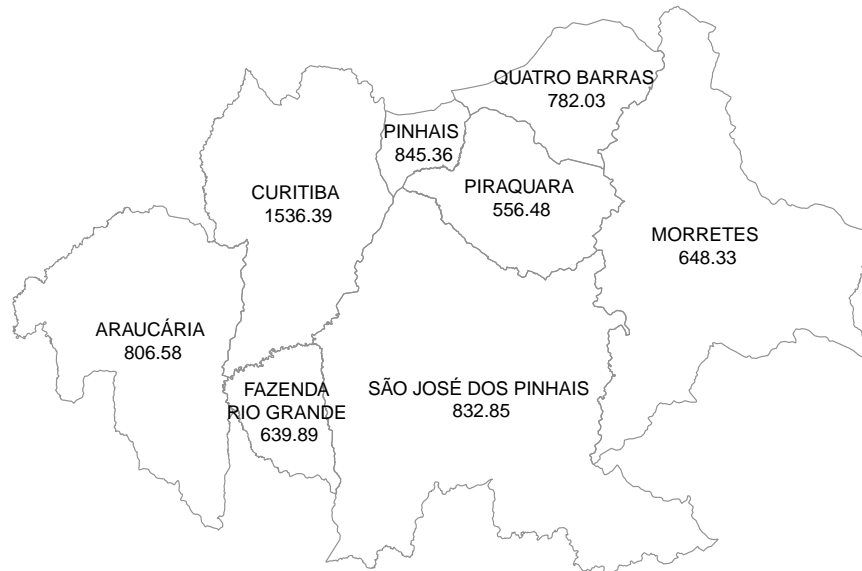
$$t = r \sqrt{\frac{n-2}{1-r^2}}$$

tem distribuição t-Student com  $n - 2$  graus de liberdade. Considere alguns valores de  $\rho$  (use

$\rho = \{0, 0.1, 0.2, 0.3\}$ ). Para cada um, simule várias séries temporais, calcule a estatística  $t$  para cada série temporal simulada, faça o histograma dessas estatísticas e compare com a distribuição t-Student com  $n - 2$  graus de liberdade. Interprete esses quatro histogramas.

- c. (Distribuição empírica de  $r$  sob  $H_0 : \rho = 0$ ). Considere uma das séries temporais simuladas na primeira figura. Mesmo que a correlação entre dois valores em tempos vizinhos seja aproximadamente 0.9, é razoável considerar que a correlação entre tempos razoavelmente distantes é próxima de zero. Assim, o coeficiente de correlação empírico  $r$  entre os valores da série (de 1 até  $n$ ) e entre outros  $n$  valores dessa série selecionados em ordem aleatória (ou simplesmente embaralhados), é, provavelmente, próximo de zero. Faça este procedimento várias vezes, visualize o histograma dos valores de  $r$  obtidos e compare com o valor de  $r$  serial observado, isto é, entre  $x_{2:n}$  e  $x_{1:(n-1)}$ . Interprete.

2. Na figura a seguir temos a renda média *per capita* nos municípios identificados, no ano de 2010.



- a. Calcule o índice de autocorrelação espacial de Moran, dado por

$$\hat{I} = \frac{n}{\sum_{i,j} \mathbf{W}_{i,j}} \frac{\sum_{i,j} \mathbf{W}_{i,j} z_i z_j}{\sum_i z_i^2}$$

- $z_i = x_i - \bar{x}$ , com  $x_i$  sendo a renda média per-capita no município  $i$ .
- $\mathbf{W}$  é a matriz de vizinhança ponderada, quadrada de ordem  $n$  tal que

$$\mathbf{W}_{i,j} = \begin{cases} 0 & \text{se } i = j \\ 1/n_i & \text{se a área } i \text{ é vizinha da área } j \\ 0 & \text{c.c.} \end{cases}$$

onde  $n_i$  é o número de vizinhos da área  $i$ .

- b. Aleatorize os 8 valores e calcule novamente  $\hat{I}$ . Repita este procedimento várias vezes, faça um histograma e compare com o valor de  $\hat{I}$  calculado para os dados originais. Interprete.
- c. Considere a renda média *per capita* de FAZENDA RIO GRANDE como 1000 e refaça o teste anterior