

Coeficientes de correlação e testes de independência

Linguagem R

2023

```
# Separador decimal: ","  
options(OutDec = ",")
```

Neste exemplo são usados os dados de Anscombe (1973) do pacote `datasets` em R. São 11 observações de quatro pares de variáveis.

```
dim(anscombe)
```

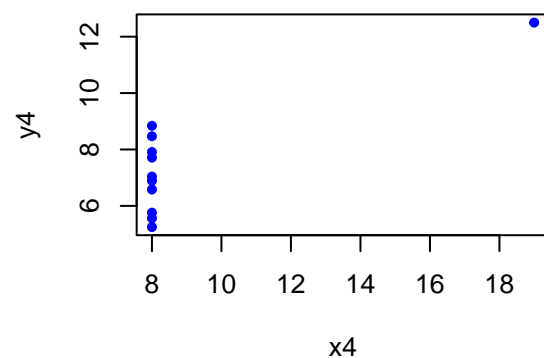
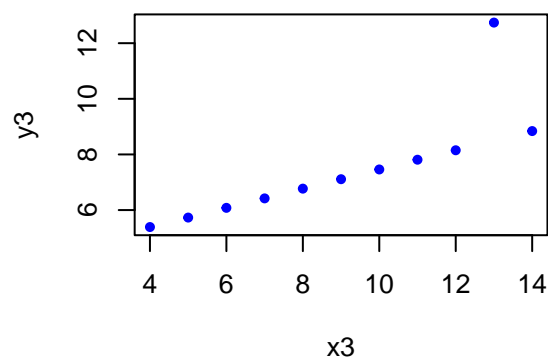
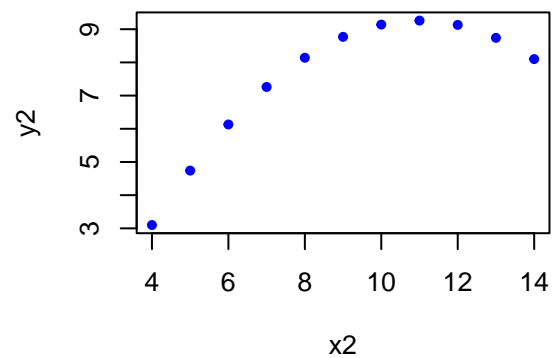
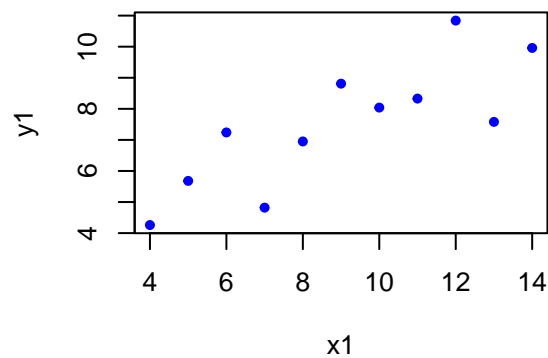
```
## [1] 11  8
```

```
names(anscombe)
```

```
## [1] "x1" "x2" "x3" "x4" "y1" "y2" "y3" "y4"
```

```
# Gráficos de dispersão
```

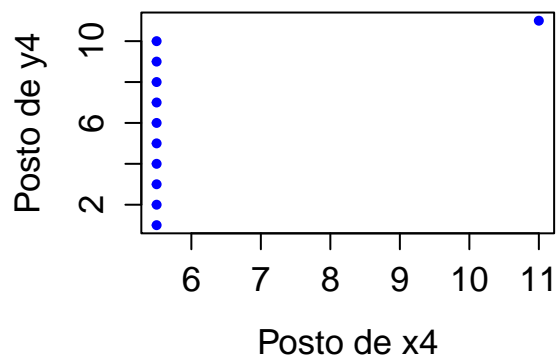
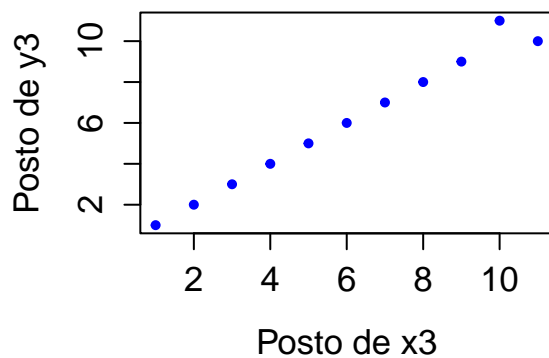
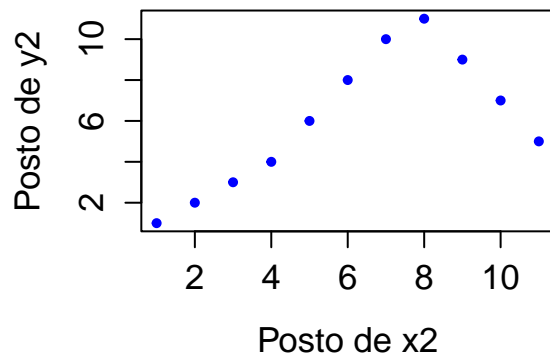
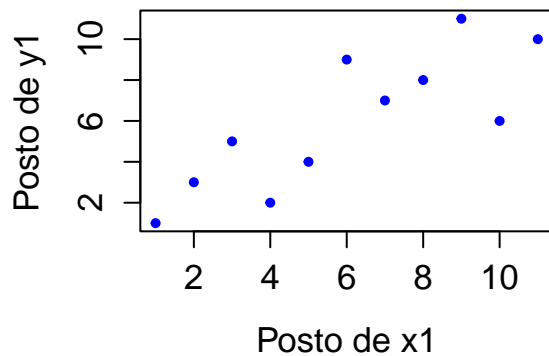
```
par(mfrow = c(2, 2))  
par(mai = c(1, 1, 0.1, 0.1))  
for (i in 1:4) {  
  plot(anscombe[, i], anscombe[, i + 4], pch = 20, col = "blue",  
        xlab = names(anscombe[i]), ylab = names(anscombe[i + 4]))  
}
```



```

# Gráficos de dispersão dos postos
par(mfrow = c(2, 2))
par(mai = c(1, 1, 0.1, 0.1))
for (i in 1:4) {
  plot(rank(anscombe[, i]), rank(anscombe[, i + 4]), pch = 20,
       xlab = paste("Posto de", names(anscombe[i])),
       ylab = paste("Posto de", names(anscombe[i + 4])),
       cex.lab = 1.3, cex.axis = 1.3, col = "blue")
}

```



```

# Correlações (default: Pearson)
for (i in 1:4) {
  cat("\n\n Exemplo", i)
  cat("\n Pearson =", cor(anscombe[, i],
                          anscombe[, i + 4]))
  cat("\n Kendall =", cor(anscombe[, i],
                          anscombe[, i + 4], method = "kendall"))
  cat("\n Spearman =", cor(anscombe[, i],
                           anscombe[, i + 4], method = "spearman"))
}

```

```

##
##
## Exemplo 1
## Pearson = 0,8164205
## Kendall = 0,6363636
## Spearman = 0,8181818
##
## Exemplo 2

```

```
## Pearson = 0,8162365
## Kendall = 0,5636364
## Spearman = 0,6909091
##
## Exemplo 3
## Pearson = 0,8162867
## Kendall = 0,9636364
## Spearman = 0,9909091
##
## Exemplo 4
## Pearson = 0,8165214
## Kendall = 0,4264014
## Spearman = 0,5
```

Nos quatro exemplos, as três primeiras casas decimais do coeficiente de correlação linear de Pearson coincidem.

```
# Teste de independência com a função cor.test
# Sem empates: valor-p exato (default: exact = NULL)
cat("\n Exemplo 1")
```

```
##
## Exemplo 1
```

```
cor.test(anscombe[, 1], anscombe[, 5], method = "kendall")
```

```
##
## Kendall's rank correlation tau
##
## data: anscombe[, 1] and anscombe[, 5]
## T = 45, p-value = 0,005707
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##      tau
## 0,6363636
```

```
cor.test(anscombe[, 1], anscombe[, 5], method = "spearman")
```

```
##
## Spearman's rank correlation rho
##
## data: anscombe[, 1] and anscombe[, 5]
## S = 40, p-value = 0,003734
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0,8181818
```

```
cat("\n Exemplo 4")
```

```
##
## Exemplo 4
```

```
cor.test(anscombe[, 4], anscombe[, 8], method = "kendall")
```

```
## Warning in cor.test.default(anscombe[, 4], anscombe[, 8], method =
## "kendall"): Cannot compute exact p-value with ties
```

```
##
## Kendall's rank correlation tau
```

```

##
## data:  anscombe[, 4] and anscombe[, 8]
## z = 1,5811, p-value = 0,1138
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##      tau
## 0,4264014
cor.test(anscombe[, 4], anscombe[, 8], method = "spearman")

## Warning in cor.test.default(anscombe[, 4], anscombe[, 8], method =
## "spearman"): Cannot compute exact p-value with ties

##
## Spearman's rank correlation rho
##
## data:  anscombe[, 4] and anscombe[, 8]
## S = 110, p-value = 0,1173
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
## rho
## 0,5
# Teste de independência com a função cor.test
# Com correção de continuidade
cat("\n Exemplo 4")

##
## Exemplo 4
cor.test(anscombe[, 4], anscombe[, 8], method = "kendall",
         continuity = TRUE)

## Warning in cor.test.default(anscombe[, 4], anscombe[, 8], method =
## "kendall", : Cannot compute exact p-value with ties

##
## Kendall's rank correlation tau
##
## data:  anscombe[, 4] and anscombe[, 8]
## z = 1,423, p-value = 0,1547
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##      tau
## 0,4264014
cor.test(anscombe[, 4], anscombe[, 8], method = "spearman",
         continuity = TRUE)

## Warning in cor.test.default(anscombe[, 4], anscombe[, 8], method =
## "spearman", : Cannot compute exact p-value with ties

##
## Spearman's rank correlation rho
##
## data:  anscombe[, 4] and anscombe[, 8]
## S = 110, p-value = 0,1154
## alternative hypothesis: true rho is not equal to 0

```

```
## sample estimates:  
## rho  
## 0,5
```

Nota 1. Compare com os valores- p sem correção de continuidade.

```
# Teste de independência com a função Kendall  
# Sem empates: valor-p exato (default: exact = NULL)  
# Com empates: dist. normal com correção de continuidade
```

```
library(Kendall)
```

```
cat("\n Exemplo 1")
```

```
##  
## Exemplo 1
```

```
Kendall(anscombe[, 1], anscombe[, 5])
```

```
## tau = 0,636, 2-sided pvalue =0,0081234
```

```
cat("\n Exemplo 4")
```

```
##  
## Exemplo 4
```

```
Kendall(anscombe[, 4], anscombe[, 8])
```

```
## tau = 0,426, 2-sided pvalue =0,15473
```

Nota 2. Para o Exemplo 1, compare o valor- p exato acima com o valor- p exato. A função `Kendall` fornece resultados mais acurados.

Para o Exemplo 4, os valores- p com correção de continuidade das funções `cor.test` e `Kendall` coincidem.

```
# Teste de independência com a função spearman.test
```

```
library(pspearman)
```

```
cat("\n Exemplo 1. Cálculo do valor-p")
```

```
##  
## Exemplo 1. Cálculo do valor-p
```

```
spearman.test(anscombe[, 1], anscombe[, 5])$p.value
```

```
## [1] 0,003299262
```

```
spearman.test(anscombe[, 1], anscombe[, 5], approximation = "AS89")$p.value
```

```
## [1] 0,003734471
```

```
spearman.test(anscombe[, 1], anscombe[, 5],  
  approximation = "t-distribution")$p.value
```

```
## [1] 0,002083145
```

```
cor.test(anscombe[, 1], anscombe[, 5], method = "spearman")$p.value
```

```
## [1] 0,003734471
```

Para $n < 1290$ e cálculo do valor- p “exato”, no teste com o coeficiente de Spearman a função `cor.test` utiliza o método de aproximação AS89. Caso contrário, é utilizada a aproximação com a distribuição t de Student.

```
cor.test(anscombe[, 1], anscombe[, 5], method = "spearman", exact = FALSE)$p.value
```

```
## [1] 0,002083145
```

```
cor.test(anscombe[, 1], anscombe[, 5], method = "spearman",  
         exact = FALSE, continuity = TRUE)$p.value
```

```
## [1] 0,002043642
```

Nota 3. A função `spearman.test` fornece resultados mais acurados.

Nota 4. Com as funções `cor.test` e `spearman.test`, por default a hipótese alternativa é bilateral (`alternative = "two.sided"`). Hipóteses alternativas unilaterais podem ser testadas com as opções `"less"` (associação negativa) e `"greater"` (associação positiva).