# Capítulo 4

## Problema 01.

**(a)**

| | Grau de Instrução | | | |
|---|---|---|---|---|
| Procedência | 1º grau | 2º grau | Superior | Total |
| Interior | 3 (0,083) | 7 (0,194) | 2 (0,056) | 12 (0,33) |
| Capital | 4 (0,111) | 5 (0,139) | 2 (0,056) | 11 (0,31) |
| Outra | 5 (0,139) | 6 (0,167) | 2 (0,056) | 13 (0,36) |
| Total | 12 (0,33) | 18 (0,50) | 6 (0,17) | 36 (1,00) |

**(b)** Dos funcionários dessa empresa, 50% têm o segundo grau.

**(c)** Dos funcionários dessa empresa, 19,4% têm o segundo grau e são oriundos do interior.

**(d)** Dentre os funcionários do interior, 7/12 (58,3%) têm o segundo grau.

## Problema 02.

**(a)** No sorteio de um indivíduo dentre os 36, é maior a probabilidade de o mesmo ter o segundo grau.

**(b)** Quanto à região de procedência, a maior probabilidade está associada com a região identificada por "Outra".

**(c)** A probabilidade de um indivíduo sorteado aleatoriamente ter grau superior de instrução é 0,17.

**(d)** A probabilidade pedida é $\frac{0{,}056}{0{,}330} = 0{,}17$.

**(e)** Nesse caso, temos $P(Superior / Capital) = \frac{0{,}056}{0{,}310} = 0{,}18$

## Problema 03.

**(a)** Temos que $md(X) = 2{,}0$ e $md(Y) = 2{,}5$. Assim,

| | Y | | |
|---|---|---|---|
| X | Baixo | Alto | Total |
| Baixo | 1 (0,025) | 7 (0,175) | 8 (0,20) |
| Alto | 19 (0,475) | 13 (0,325) | 32 (0,80) |
| Total | 20 (0,50) | 20 (0,50) | 40 (1,00) |

**(b)** Da tabela, tem-se que 2,5% dos indivíduos encontram-se nessas condições.

**(c)** 50%.

**(d)** Dentre as pessoas com baixa rotatividade, 12,5% ganham pouco.

**(e)** A probabilidade em (c) foi bastante modificada. Isto indica que a maioria das pessoas que ganham pouco têm rotatividade.

**Problema 04.**

**(a)**

| Região de Procedência | Grau de Instrução | | |
|---|---|---|---|
| | 1º grau | 2º grau | Superior |
| Interior | 0,250 | 0,583 | 0,167 |
| Capital | 0,364 | 0,455 | 0,182 |
| Outra | 0,385 | 0,462 | 0,154 |

**(b)** Em caso de independência entre a região de procedência e grau de escolaridade, em cada tabela deveria existir 33% com 1º grau, 50% com 2º grau e 17% com grau Superior.

**Problema 05.**

Tabela do total de linhas

| | Y | | |
|---|---|---|---|
| X | Baixo | Alto | Total |
| Baixo | 1 (12,5%) | 7 (87,5%) | 8 (100,0%) |
| Alto | 19 (59,4%) | 13 (40,6%) | 32 (100,0%) |
| Total | 20 (50,0%) | 20 (50,0%) | 40 (100,0%) |

Tabela do total de colunas.

| | Y | | |
|---|---|---|---|
| X | Baixo | Alto | Total |
| Baixo | 1 (5,0%) | 7 (35,0%) | 8 (20,0%) |
| Alto | 19 (95,0%) | 13 (65,0%) | 32 (80,0%) |
| Total | 20 (100,0%) | 20 (100,0%) | 40 (100,0%) |

As tabelas acima indicam existência de relação entre as variáveis rotatividade e salário, pois as proporções marginais não se repetem no interior da tabela.

**Problema 06.**

**(a)** A proporção de homens entre os indivíduos que usaram o hospital é: $^{100}\!/_{250} = 0,4$

**(b)** A proporção de homens entre os indivíduos que não usaramo hospital é: $^{900}\!/_{1750} = 0,514$

**(c)** Tabela do total de colunas.

| | | | |
|---|---|---|---|
| Usaram o hospital | 100 (0,10) | 150 (0,15) | 0,25 |
| Não usaram o hospital | 900 (0,90) | 850 (0,85) | 0,75 |
| | 1,00 | 1,00 | 1,00 |

Independentemente do sexo, 25% das pessoas usam e 75% não usam o hospital. Essas porcentagens deveriam ser iguais nas duas colunas e não são. Portanto, o uso do hospital depende do sexo do segurado.

**Problema 07.**

Veja a tabela a seguir. Entre parênteses, encontram-se os valores esperados em caso de independência das variáveis.

| Grau de Instrução | | | | |
|---|---|---|---|---|
| Procedência | 1º grau | 2º grau | Superior | Total |
| Interior | 3 (4,00) | 7 (6,00) | 2 (2,00) | 12 |
| Capital | 4 (3,67) | 5 (5,50) | 2 (1,83) | 11 |
| Outra | 5 (4,33) | 6 (6,50) | 2 (2,17) | 13 |
| Total | 12 | 18 | 6 | 36 |

Com isso, os cálculos ficam assim:

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} = 0{,}25 + 0{,}17 + 0 + 0{,}03 + 0{,}05 + 0{,}02 + 0{,}10 + 0{,}04 + 0{,}01 = 0{,}67$$

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}} = \sqrt{\frac{0{,}67}{0{,}67 + 36}} = 0{,}81$$

**Problema 08.**

Para os dados do problema 3, tem-se:

| | Y | | |
|---|---|---|---|
| X | Baixo | Alto | Total |
| Baixo | 1 (4) | 7 (4) | 8 |
| Alto | 19 (16) | 13 (16) | 32 |
| Total | 20 | 20 | 40 |

De modo que,

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} = 2{,}25 + 2{,}25 + 0{,}5625 + 0{,}5625 = 5{,}625$$

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}} = \sqrt{\frac{5{,}625}{5{,}625 + 40}} = 0{,}351$$

$$T = \sqrt{\frac{\chi^2 / n}{(r-1)\times(s-1)}} = \sqrt{\frac{5{,}625/40}{1\times 1}} = 0{,}375$$

Para os dados do problema 6, tem-se:

| | Homens | Mulheres | Total |
|---|---|---|---|
| Usaram o hospital | 100 (125) | 150 (125) | 250 |
| Não usaram o hospital | 900 (875) | 850 (875) | 1750 |
| Total | 1000 | 1000 | 2000 |

De modo que,

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} = 5{,}00 + 5{,}00 + 0{,}71 + 0{,}71 = 11{,}42$$

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}} = \sqrt{\frac{11{,}42}{11{,}42 + 2000}} = 0{,}075$$

$$T = \sqrt{\frac{\chi^2 / n}{(r-1)\times(s-1)}} = \sqrt{\frac{11{,}42/2000}{1\times 1}} = 0{,}076$$

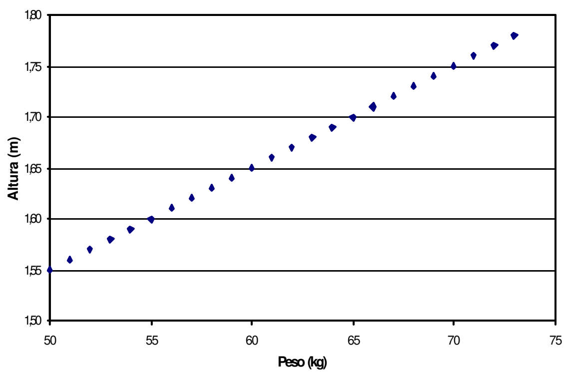
**Problema 09.**

Os dados podem ser assim representados:

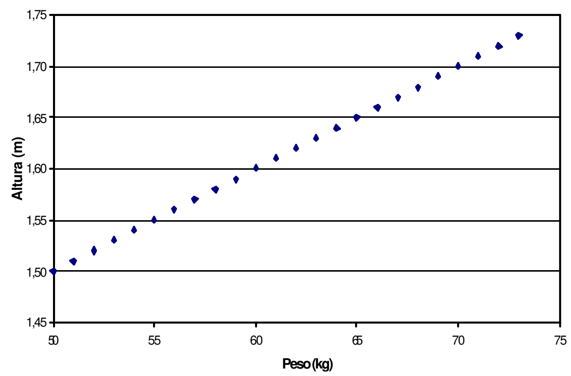
| Companhia | Duração de efeito de dedetização | | |
|---|---|---|---|
| | Menos de 4 meses | De 4 a 8 meses | Mais de 8 meses |
| X | 0,32 | 0,60 | 0,08 |
| Y | 0,35 | 0,58 | 0,07 |
| Z | 0,34 | 0,60 | 0,06 |

Essas proporções indicam que não há diferenças da duração de efeito de dedetização entre as três empresas.

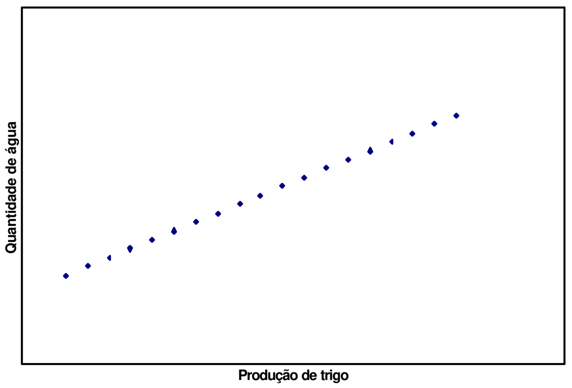
## Problema 10.



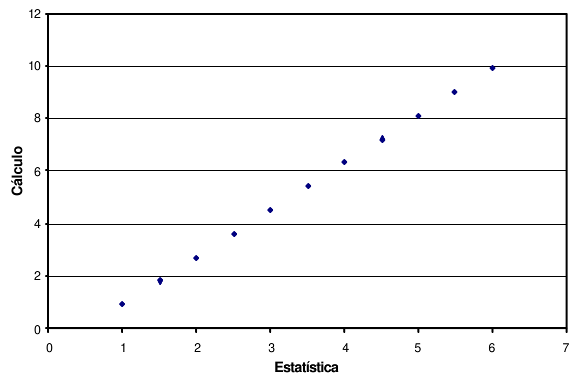

**(a)**



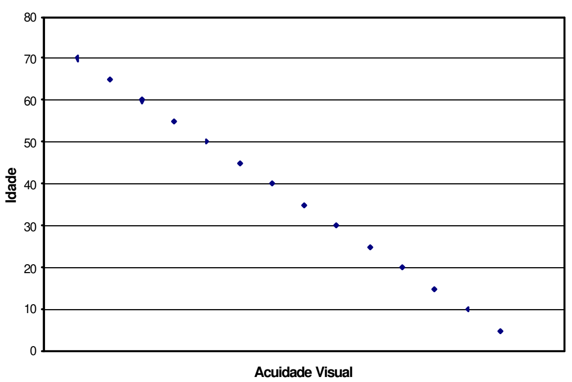

**(b)**



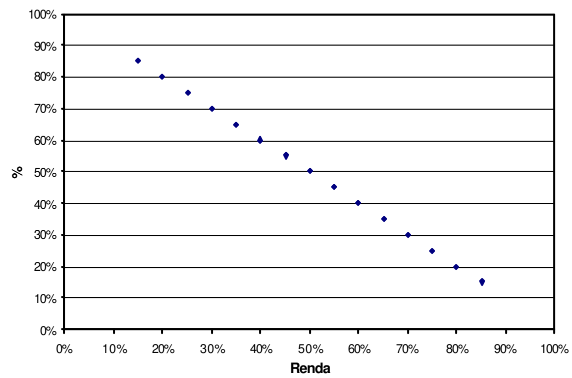

**(c)**



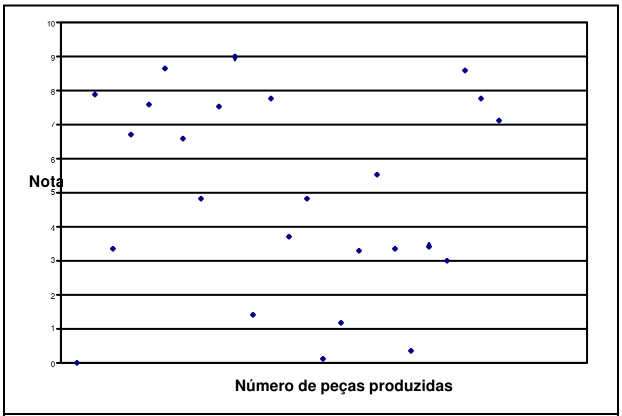

**(d)**




**(e)**



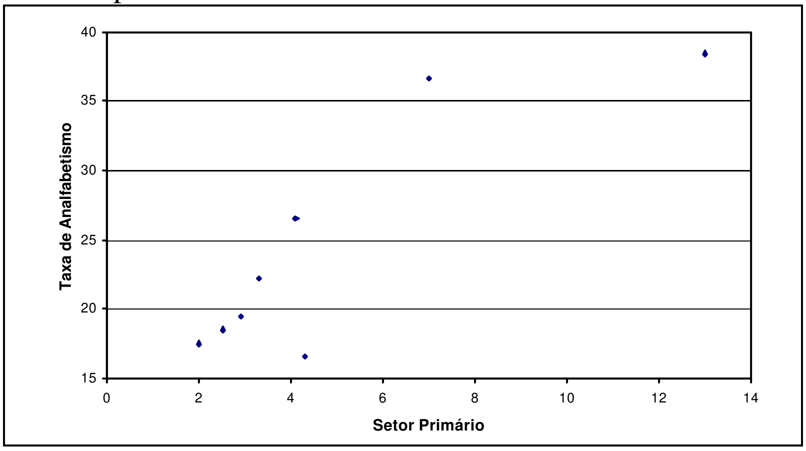

**(f)**




**(g)**

## Problema 11.

**(a)** Diagrama de dispersão




**(b)** O gráfico do item (a) indica dependência linear entre as variáveis.

**(c)** $$Corr(X,Y)=\frac{1}{8}\sum_{i=1}^{8}\left[\left(\frac{x_i-4{,}887}{3{,}62}\right)\times\left(\frac{y_i-24{,}48}{8{,}63}\right)\right]=0{,}86$$

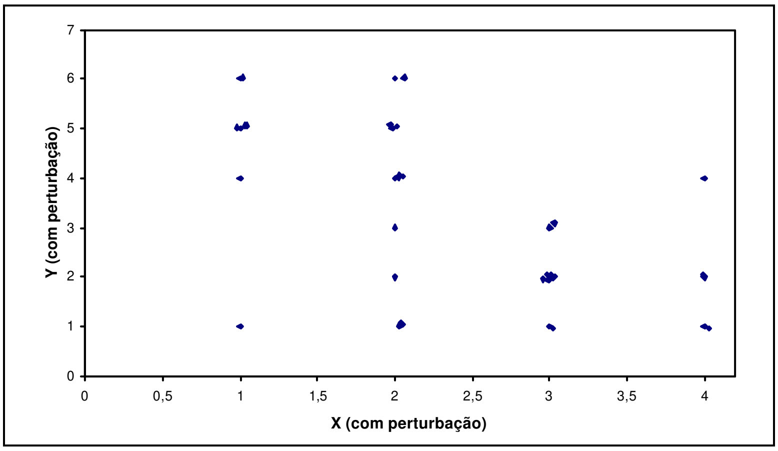
**(d)** As regiões de Porto Alegre e Fortaleza apresentam comportamento diferente das demais. Retirando-se esses elementos do cálculo resulta $Corr(X,Y)=0{,}91$.

## Problema 12.

**(a)**

| | Y | | | | | | |
|---|---|---|---|---|---|---|---|
| X | 1 | 2 | 3 | 4 | 5 | 6 | Total |
| 1 | 1 | 0 | 0 | 1 | 4 | 2 | 8 |
| 2 | 3 | 2 | 1 | 4 | 3 | 2 | 15 |
| 3 | 2 | 7 | 2 | 0 | 0 | 0 | 11 |
| 4 | 3 | 2 | 0 | 1 | 0 | 0 | 6 |
| Total | 9 | 11 | 3 | 6 | 7 | 4 | 40 |

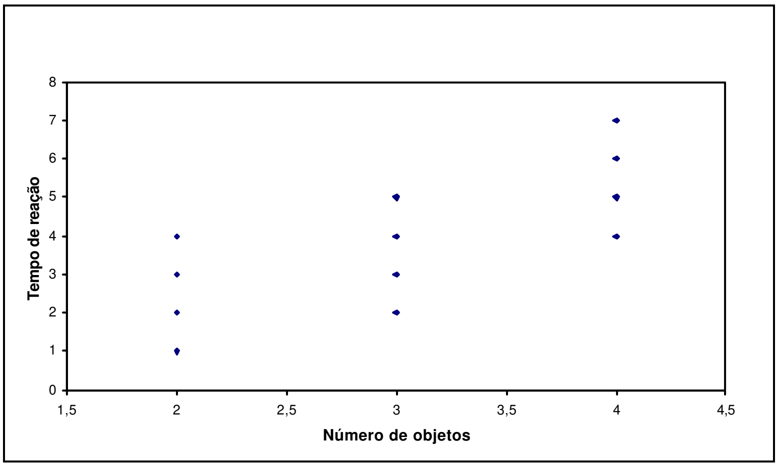
**(b)** Como existem pontos que coincidiriam no caso de um diagrama de dispersão, pode-se representar os pontos coincidentes no gráfico com número de repetições. Outra alternativa, válida do ponto de vista descritivo é adicionar uma perturbação aos pontos. Soma-se uma quantidade pequena às coordenadas, de modo a não haver mais coincidências. A seguir, o gráfico com a perturbação:




**(c)** O coeficiente de correlação entre X e Y é 0,59, indicando uma dependência linear moderada entre as variáveis.

### Problema 13.

**(a)** Gráfico de dispersão




**(b)** O coeficiente de correlação entre as variáveis é 0,74.

### Problema 14.

X: idade

| Estado Civil | n | $\bar{x}$ | dp(X) | var(X) | $x_{(1)}$ | $q_1$ | $q_2$ | $q_3$ | $x_n$ |
|---|---|---|---|---|---|---|---|---|---|
| solteiro | 16 | 34,33 | 7,69 | 59,11 | 20,83 | 27,50 | 35,75 | 40,68 | 46,58 |
| casado | 20 | 35,63 | 5,95 | 35,36 | 26,08 | 31,37 | 34,91 | 39,81 | 48,92 |
| Total | 36 | 34,58 | 6,74 | 45,39 | 20,00 | 30,00 | 34,50 | 40,00 | 48,92 |

$$\overline{\text{var}(X)} = \frac{16 \times 59{,}11 + 20 \times 35{,}36}{36} \cong 45{,}39$$

$$R^2 = \frac{\overline{\text{var}(X)}}{\text{var}(X)} = 1 - \frac{45{,}39}{45{,}39} = 0$$

**Problema 15.**

X: Nota em Estatística.

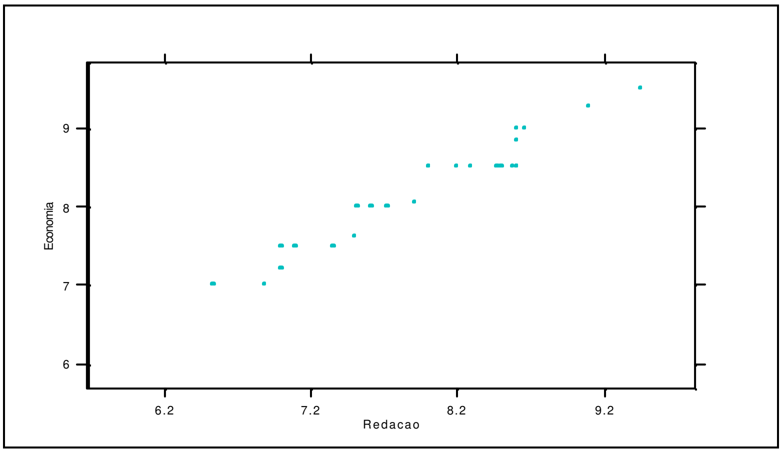
| Seção | N | $\bar{x}$ | dp(X) | var(X) | $x_{(1)}$ | $q_1$ | $q_2$ | $q_3$ | $x_n$ |
|---|---|---|---|---|---|---|---|---|---|
| P | 7 | 8,71 | 0,75 | 0,57 | 8 | 8 | 9 | 9 | 10 |
| T | 7 | 8,29 | 1,11 | 1,24 | 7 | 7,5 | 8 | 9 | 10 |
| V | 11 | 7,91 | 1,64 | 2,69 | 4 | 7 | 8 | 9 | 10 |
| Total | 25 | 8,24 | 1,30 | 1,69 | 4 | 8 | 8 | 9 | 10 |

$$\overline{\text{var}(X)} = \frac{7\times 0{,}57 + 7\times 1{,}24 + 11\times 2{,}69}{25} = \frac{3{,}99 + 8{,}68 + 29{,}59}{25} = \frac{42{,}26}{25} = 1{,}69$$

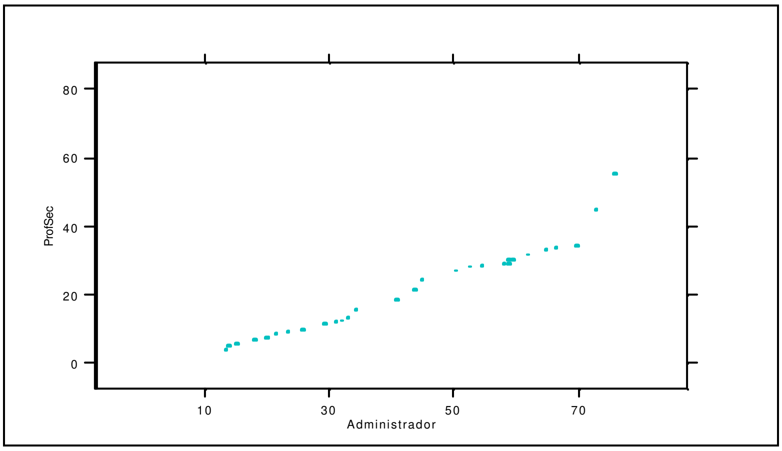
$$R^2 = \frac{\overline{\text{var}(X)}}{\text{var}(X)} = 1 - \frac{45{,}39}{45{,}39} = 0$$

Logo, Seção não serve para explicar nota.

**Problema 16.**




**Problema 17.**




Pode-se perceber que os pontos estão razoavelmente dispersos abaixo em relação a reta (x=y). Logo, parece que os salários dos professores secundários é menor que o dos administradores.

**Problema 18.**

**(a)**

| Estado Civil | Salário: Menos de 10 SM | Entre 10 e 20 SM | Mais de 20 SM | Total |
|---|---|---|---|---|
| Solteiro | 0,12 | 0,19 | 0,09 | 0,40 |
| Casado | 0,08 | 0,31 | 0,21 | 0,60 |
| Total | 0,20 | 0,50 | 0,30 | 1,00 |

**(b)** Considere-se a tabela do total de colunas:

| Estado Civil | Salário: Menos de 10 SM | Entre 10 e 20 SM | Mais de 20 SM | Total |
|---|---|---|---|---|
| Solteiro | 0,60 | 0,38 | 0,30 | 0,40 |
| Casado | 0,40 | 0,62 | 0,70 | 0,60 |
| Total | 1,00 | 1,00 | 1,00 | 1,00 |

Pelas diferenças entre as proporções marginais e as do interior da tabela, parece haver relação entre as variáveis.

## Problema 19.

**(a)**

| Opinião | Local de residência: Urbano | Suburbano | Rural | Total |
|---|---|---|---|---|
| A favor | 0,33 | 0,58 | 0,70 | 0,50 |
| Contra | 0,67 | 0,42 | 0,30 | 0,50 |

**(b)** A opinião parece depender do local de residência do indivíduo.

| Opinião | Local de residência: Urbano | Suburbano | Rural | Total |
|---|---|---|---|---|
| A favor | 30 (45) | 35 (30) | 35 (25) | 100 |
| Contra | 60 (45) | 25 (30) | 15 (25) | 100 |

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} = 5,00 + 5,00 + 0,83 + 0,83 + 4,00 + 4,00 = 19,66$$

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}} = \sqrt{\frac{19,66}{19,66 + 200}} = 0,30$$

## Problema 20.

Considere a tabela com os valores observados e os esperados:

| Propriedade | Atividade: Costeira | Fluvial | Internacional | Total |
|---|---|---|---|---|
| Estatal | 5 (33,64) | 141 (129,02) | 51 (34,34) | 197 |
| Particular | 92 (63,64) | 231 (242,98) | 48 (64,66) | 371 |

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} = 24,38 + 1,11 + 8,08 + 12,64 + 0,59 + 4,29 = 51,09$$

Parece existir associação entre o tipo de atividade e propriedade das embarcações.

## Problema 21.

Considere a tabela com os valores observados e esperados :

| Participaram | Cidade | | | |
|---|---|---|---|---|
| | São Paulo | Campinas | Rib. Preto | Santos |
| Sim | 50 (64,76) | 65 (80,95) | 105 (97,14) | 120 (97,14) |
| Não | 150 (135,24) | 185 (169,05) | 195 (202,86) | 180 (202,86) |

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} = 3{,}36 + 3{,}14 + 0{,}64 + 5{,}38 + 1{,}61 + 1{,}50 + 0{,}30 + 2{,}58 = 18{,}51$$

Os dados da tabela indicam que a participação em atividades esportivas depende da cidade.

## Problema 22.

**(a)** Tabela dos totais de colunas.

| Pretende continuar? | Classe social | | | |
|---|---|---|---|---|
| | Alta | Média | Baixa | Total |
| Sim | 0,50 | 0,44 | 0,38 | 0,40 |
| Não | 0,50 | 0,56 | 0,72 | 0,60 |

Há evidências de que a distribuição das respostas afirmativas e negativas não coincidem.

**(b)** Tabela dos valores observados e esperados:

| Pretende continuar? | Classe social | | | |
|---|---|---|---|---|
| | Alta | Média | Baixa | Total |
| Sim | 200 (160) | 220 (200) | 380 (440) | 800 |
| Não | 200 (240) | 280 (300) | 720 (660) | 1200 |

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} = 10{,}00 + 2{,}00 + 8{,}18 + 6{,}67 + 1{,}33 + 5{,}45 = 33{,}63$$

Existe dependência entre as variáveis.

**(c)** Se houvesse tal modificação, a dependência entre as variáveis seria apenas menor ($\chi^2 = 7{,}01$).

## Problema 23.

$$\frac{n_{11}}{n_{\cdot 1}} = \frac{30}{90} = 0{,}33 \quad \text{e} \quad \frac{n_{21}}{n_{\cdot 1}} = \frac{60}{90} = 0{,}67$$

$$\frac{n_{12}}{n_{\cdot 2}} = \frac{35}{60} = 0{,}58 \quad \text{e} \quad \frac{n_{22}}{n_{\cdot 2}} = \frac{25}{60} = 0{,}42$$

$$\frac{n_{13}}{n_{\cdot 3}} = \frac{35}{50} = 0{,}70 \quad \text{e} \quad \frac{n_{23}}{n_{\cdot 3}} = \frac{15}{50} = 0{,}30$$

## Problema 24.

$$Corr(X,Y) = \frac{1}{n}\sum_i \left[\left(\frac{x_i - \bar{x}}{dp(X)}\right)\left(\frac{y_i - \bar{y}}{dp(Y)}\right)\right] = \frac{1}{n}\sum_i \left[\frac{x_i y_i - x_i \bar{y} - y_i \bar{x} + \overline{xy}}{\sqrt{\left(\sum_i x_i^2 - n\bar{x}^2\right)\left(\sum_i y_i^2 - n\bar{y}^2\right) \Big/ n^2}}\right] =$$

$$= \frac{\sum_i x_i y_i - \bar{y}\sum_i x_i - \bar{x}\sum_i y_i + n\overline{xy}}{\sqrt{\left(\sum_i x_i^2 - n\bar{x}^2\right)\left(\sum_i y_i^2 - n\bar{y}^2\right)}} = \frac{\sum_i x_i y_i - n\overline{xy}}{\sqrt{\left(\sum_i x_i^2 - n\bar{x}^2\right)\left(\sum_i y_i^2 - n\bar{y}^2\right)^2}}$$

**Problema 25.**

O coeficiente de correlação linear entre X e Y é -0,92, indicando forte correlação linear entre as variáveis.

$$Corr(X,Y) = \frac{53 - 5\times(3,2)\times(4,4)}{\sqrt{\left[62 - 5\times(3,2)^2\right]\times\left[130 - 5\times(4,4)^2\right]}} = -\frac{17,4}{18,93} = -0,92$$

**Problema 26.**

Pode-se calcular, com os dados fornecidos, $Corr(X,Y) = 0,95$ e $Corr(X,Z) = 0,71$. Como o valor mais alto encontrado é 0,95, a variável Y é a mais indicada para explicar a variação de X.

**Problema 27.**

**(a)**

| Idade | Salário [0,15) | Salário [15,30) | Total |
|---|---|---|---|
| [0,30) | 4 | 4 | 8 |
| [30,40) | 6 | 12 | 18 |
| [40,50) | 3 | 7 | 10 |
| Total | 13 | 23 | 36 |

**(b)** O cálculo do coeficiente de correlação neste caso, poderia ser feito utilizando-se os pontos médios de cada categoria.

**(c)** Com a idéia que foi descrita no item anterior, o cálculo do coeficiente de correlação agrupados poderia ser feito com a fórmula usual, onde haveria 4 pares (15;7,5) repetidos, 6 pares (35;7,5) repetidos, etc. Assim a fórmula seria:

$$Corr(X,Y) = \frac{1}{n}\sum_{i=1}^{6} \frac{\left[n_i\left(x_i - \bar{x}\right)\right]}{dp(X)} \frac{\left[n_i\left(y_i - \bar{y}\right)\right]}{dp(Y)}$$

onde $x_i$, $y_i$ são os pontos médios, $n_1 = n_2 = 4$, $n_3 = 6$, $n_4 = 12$, $n_5 = 3$, $n_6 = 7$

**Problema 28.**

**(a)** Tabela dos valores observados e dos observados:

| | Cara | Coroa | Total |
|---|---|---|---|
| Cara | 24 (23,92) | 22 (22,08) | 46 |
| Coroa | 28 (28,08) | 26 (25,92) | 54 |
| Total | 52 | 48 | 100 |

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} = 0,0002 + 0,0002 + 0,0002 + 0,0002 = 0,0008$$

Logo, não há associação entre os resultados das moedas de um real e de um quarto de dólar.

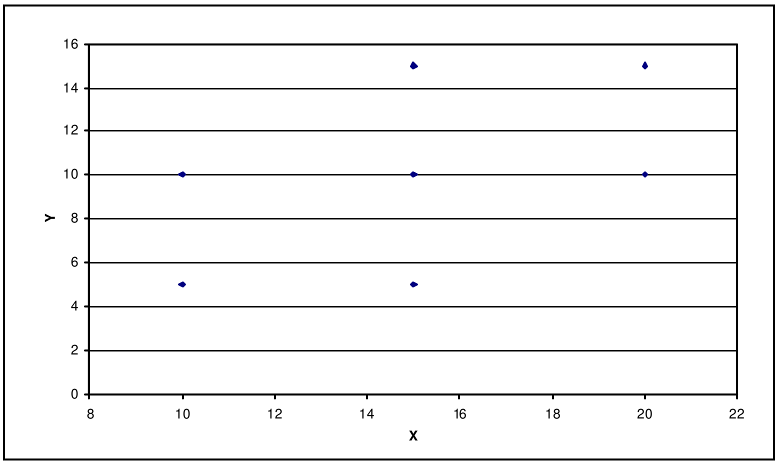
**(b)** O coeficiente de correlação linear entre as variáveis $X_1$ e $X_2$ é 0, pois $X_1$ e $X_2$ são independentes. Esse resultado está de acordo com o resultado do item anterior.

## Problema 29.

**(a)** O salário anual médio dos homens é 15 e o desvio-padrão 3,87.

**(b)** O salário anual médio das mulheres é 10 e o desvio-padrão 3,16.

**(c)**



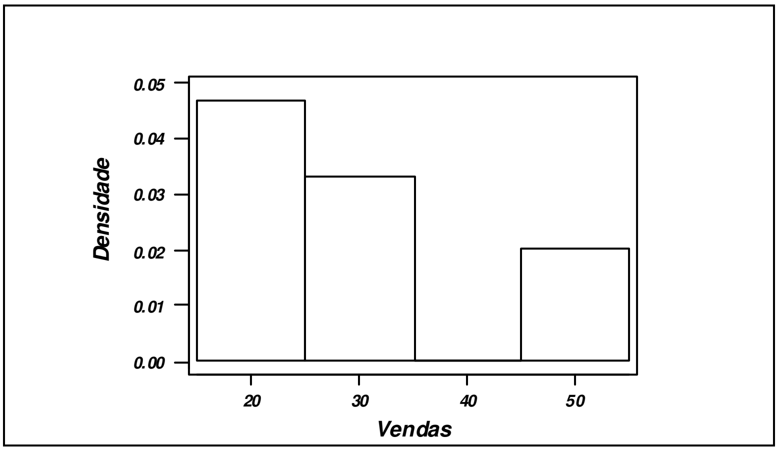

**(d)** $$Corr(X,Y) = \frac{1550-1500}{\sqrt{[2400-2250]\times[1100-1000]}} = 0{,}41$$

**(e)** O salário médio familiar é 25. A variância do salário familiar é 35.

**(f)** Descontando 8% dos salários de todos os homens da amostra e 6% do salário de todas as mulheres, o salário médio familiar cai para 23,2 e a variância vai a 30,18.

## Problema 30.

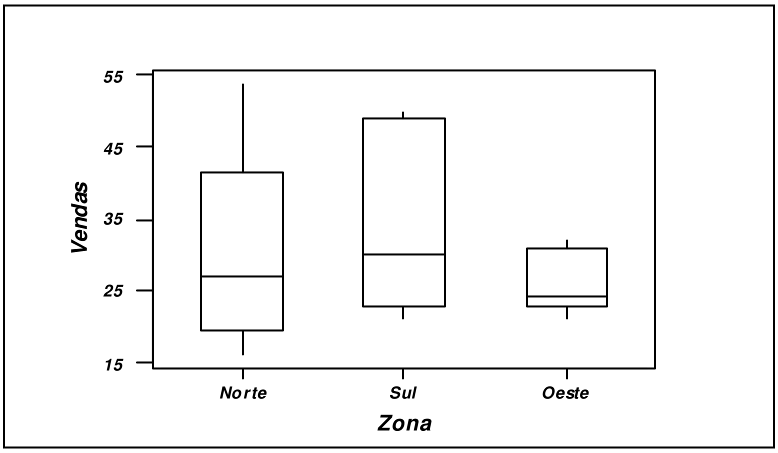
**(a)** Histograma



**(b)** A média da variável V é 30,2 e a variância 130,6. Como dp(V)=11,43, $\bar{v} + 2dp(V) = 53{,}05$ é o limite para se considerar um vendedor excepcional. Acima desse valor, há apenas 1 dentre os 15 indivíduos analisados.

**(c)** O primeiro quartil da distribuição de V é 23,5.

**(d)** Os box-plots a seguirindicam que existe alguma diferença entre a distribuição das vendas nas três diferentes zonas. Assim, não é justo aplicar um mesmo critério para todas as zonas.




**(e)** $Corr(T,V) = 0{,}71$, $Corr(E,V) = 0{,}26$, logo a variável teste parece ser a mais importante na contratação de um empregado.

**(f)**

| Conceito do gerente | Zona | | | Total |
|---|---|---|---|---|
| | Norte | Sul | Leste | |
| Bom | 4 (2,7) | 3 (2,7) | 1 (2,7) | 8 |
| Mau | 1 (2,3) | 2 (2,3) | 4 (2,3) | 7 |
| Total | 5 | 5 | 5 | 15 |

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} = 3{,}76$$

Logo, existe uma baixa associação entre o Conceito do gerente e a Zona.

**(g)** Considere X: resultado do teste.

| Conceito do gerente | n | média | dp | var |
|---|---|---|---|---|
| Bom | 8 | 6,00 | 2,14 | 4,57 |
| Mau | 7 | 6,14 | 1,68 | 2,81 |
| Total | 15 | 6,07 | 1,87 | 3,50 |

$$\overline{\text{var}(X)} = \frac{8 \times 4{,}57 + 7 \times 2{,}81}{15} \cong 3{,}50$$

$$R^2 = \frac{\overline{\text{var}(X)}}{\text{var}(X)} = 1 - \frac{3{,}50}{3{,}50} = 0$$

Considere agora X: vendas:

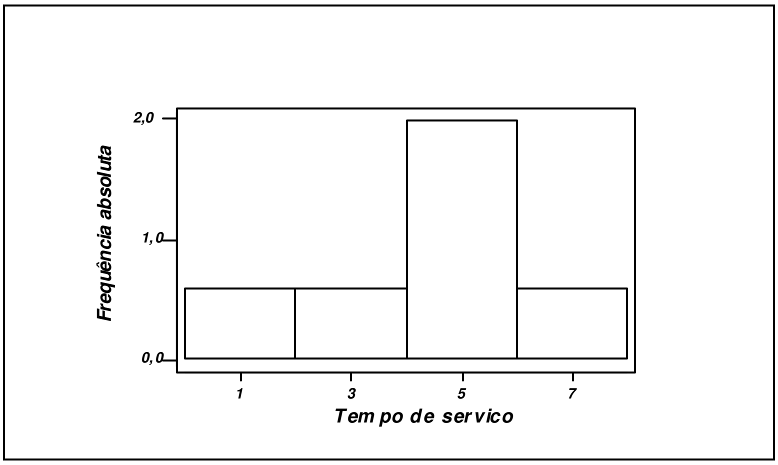
| Zona | n | média | dp | var |
|---|---|---|---|---|
| Norte | 5 | 29,8 | 14,4 | 207,7 |
| Sul | 5 | 34,6 | 13,56 | 183,8 |
| Oeste | 5 | 26,2 | 4,6 | 21,2 |
| Total | 15 | 30,2 | 11,43 | 130,6 |

$$\overline{\mathrm{var}(X)} = \frac{5\times 207,7 + 5\times 183,8 + 5\times 21,2}{15} \cong 130,5$$

$$R^2 = \frac{\overline{\mathrm{var}(X)}}{\mathrm{var}(X)} = 1 - \frac{130,5}{130,6} = 0,0008$$

## Problema 31.

**(a)**

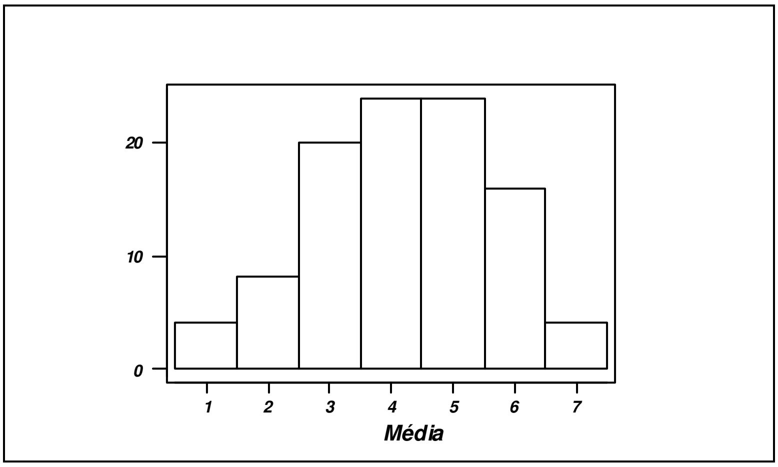



**(b)** $me(X) = 4,2$; $md(X) = 5,0$; $\mathrm{var}(X) = 5,2$

**(c)** (A,A),..., (A,E), (B,A),..., (B,E), (C,A),..., (C,E), (D,A),..., (D,E), (E,A),...,(E,E)

**(d)**

| $\overline{X}$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Freq. | 0,04 | 0,08 | 0,20 | 0,24 | 0,24 | 0,16 | 0,04 |

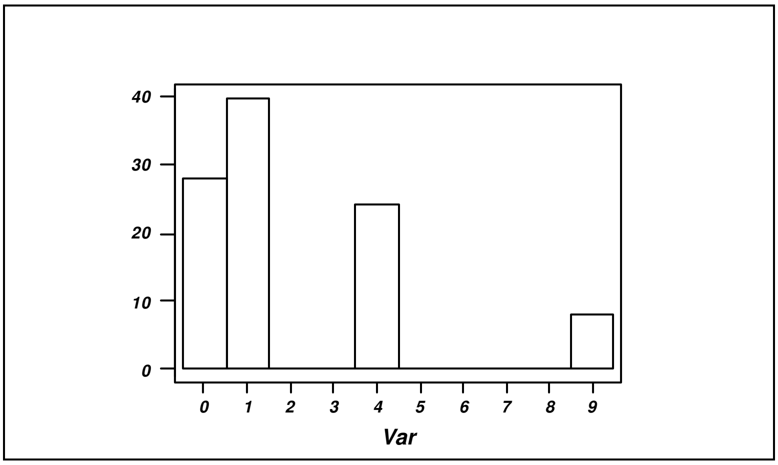


**(e)** $me(\overline{X}) = 4{,}2$; $md(\overline{X}) = 4{,}0$; $\text{var}(\overline{X}) = 2{,}6$

Vemos que $me(\overline{X}) = me(X)$ e $\text{var}(\overline{X}) = \dfrac{\text{var}(X)}{2}$

**(f)**

| $S^2$ | 0 | 1 | 4 | 9 |
|---|---|---|---|---|
| Freq. | $\frac{7}{25}$ | $\frac{10}{25}$ | $\frac{6}{25}$ | $\frac{2}{25}$ |




**(g)** $me(S^2) = 2{,}08$; $\text{var}(S^2) = 6{,}39$.

**(h)**

| | $X_2$ | | | | |
|---|---|---|---|---|---|
| $X_1$ | 1 | 3 | 5 | 7 | Total |
| 1 | 0,04 | 0,04 | 0,08 | 0,04 | 0,20 |
| 3 | 0,04 | 0,04 | 0,08 | 0,04 | 0,20 |
| 5 | 0,08 | 0,08 | 0,16 | 0,08 | 0,40 |
| 7 | 0,04 | 0,04 | 0,08 | 0,04 | 0,20 |
| Total | 0,20 | 0,20 | 0,40 | 0,20 | 1,00 |

**(i)** As variáveis são independentes, pois $P(X_1 = i, X_2 = j) = P(X_1 = i) \times P(X_2 = j)$

**(j)** São iguais entre si e à distribuição de X.

**(k)** Não tem esse item.

**(l)** Teremos $5^3$=125 triplas.

**(m)** Histograma mais próximo de uma normal; $me(\overline{X}) = me(X)$, $\text{var}(\overline{X}) = \text{var}(X)$

**(n)** Histograma com assimetria à direita.

**(o)** Distribuições marginais iguais à distribuição de X.

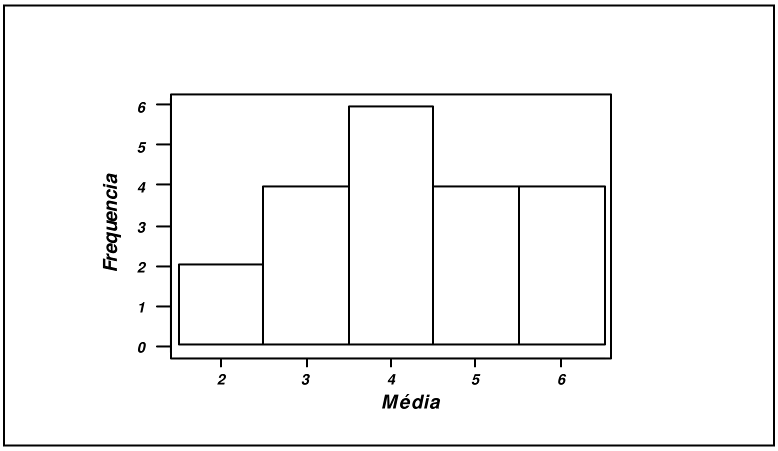
## Problema 32.

**(a)** Não tem.

**(b)** Não tem.

**(c)** (A,B),..., (A,E), (B,A),..., (B,E), (C,A),..., (C,E), (D,A),..., (D,E), (E,A),...,(E,D)

| $\overline{X}$ | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Freq. | 0,10 | 0,20 | 0,30 | 0,20 | 0,20 |

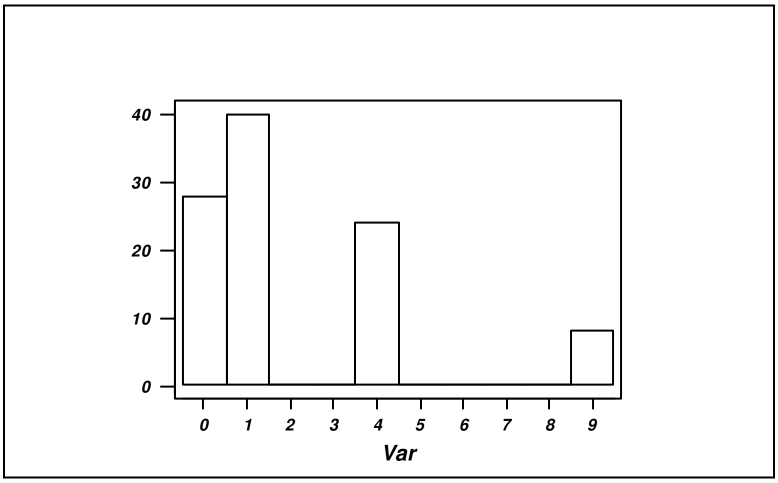



**(d)** $me(\overline{X}) = 4{,}2$; $md(\overline{X}) = 4{,}0$; $\mathrm{var}(\overline{X}) = 1{,}6$

Vemos que $me(\overline{X}) = me(X)$

**(e)**

| $S^2$ | 0 | 1 | 4 | 9 |
|---|---|---|---|---|
| Freq. | $\frac{2}{20}$ | $\frac{10}{20}$ | $\frac{6}{20}$ | $\frac{2}{20}$ |




**(f)** $me(S^2) = 2{,}60$; $\mathrm{var}(S^2) = 6{,}64$.

| $X_1$ | $X_2$ 1 | 3 | 5 | 7 | Total |
|---|---|---|---|---|---|
| 1 | 0,04 | 0,04 | 0,08 | 0,04 | 0,20 |
| 3 | 0,04 | 0,04 | 0,08 | 0,04 | 0,20 |
| 5 | 0,08 | 0,08 | 0,16 | 0,08 | 0,40 |
| 7 | 0,04 | 0,04 | 0,08 | 0,04 | 0,20 |
| Total | 0,20 | 0,20 | 0,40 | 0,20 | 1,00 |

**(g)** As variáveis são independentes, pois $P(X_1 = i, X_2 = j) = P(X_1 = i) \times P(X_2 = j)$

**(h)** São iguais entre si e à distribuição de X.

**(i)** Não tem esse item.

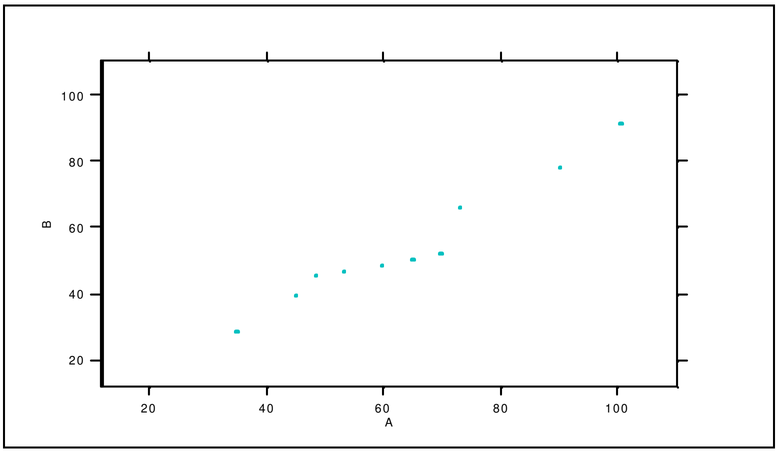
**(j)** Teremos 60 triplas.

**(k)** Histograma mais próximo de uma normal; $me(\overline{X}) = me(X)$, $\mathrm{var}(\overline{X}) = \mathrm{var}(X)$

**(l)** Histograma com assimetria à direita.

**(m)** Distribuições marginais iguais à distribuição de X.

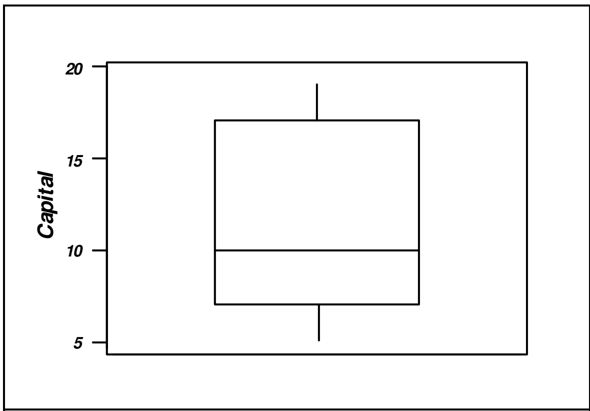
## Problema 34.



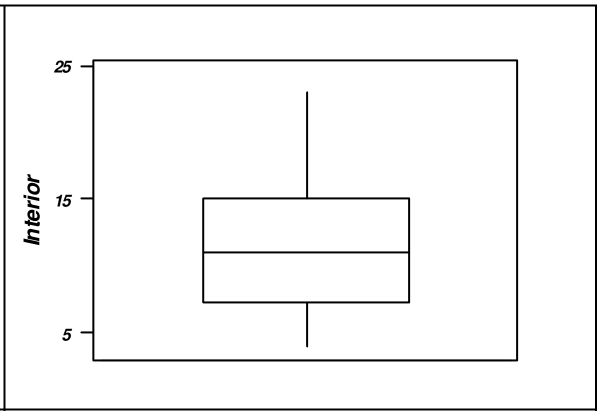

## Problema 35.

Dotplot para as regiões de procedência:

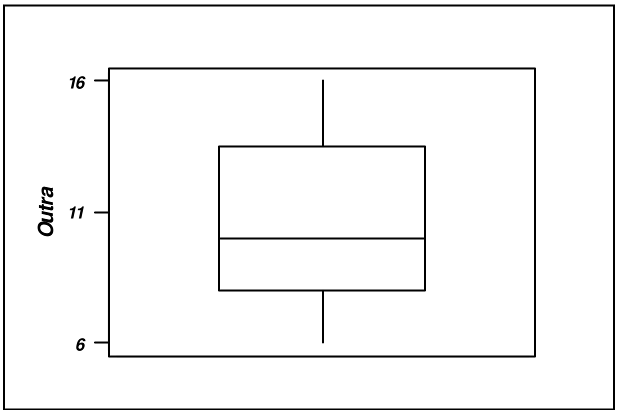
```
     .   :     .      .            .   .     :
  +---------+---------+---------+---------+---------+-Capital
6.0       9.0      12.0      15.0      18.0      21.0

     :   .  .      :     .      :      .
  +---------+---------+---------+---------+---------+-Interior
6.0       9.0      12.0      15.0      18.0      21.0


  :  .      :  .   .  .  .   :     .
  +---------+---------+---------+---------+---------+-Outra
6.0       9.0      12.0      15.0      18.0      21.0
```




BoxPlot - Capital



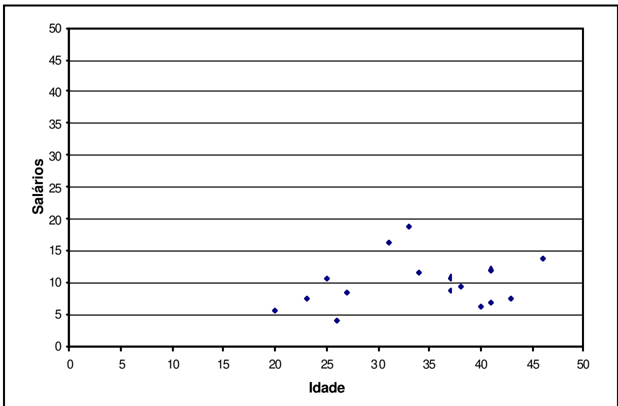

BoxPlot – Interior



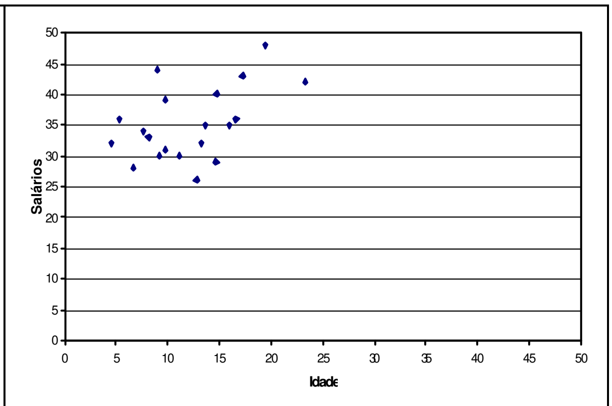

BoxPlot - Outra

Pode-se observar que os salários da Capital têm variabilidade maior e distribuição mais assimétrica. As médias e medianas são similares.

## Problema 36.



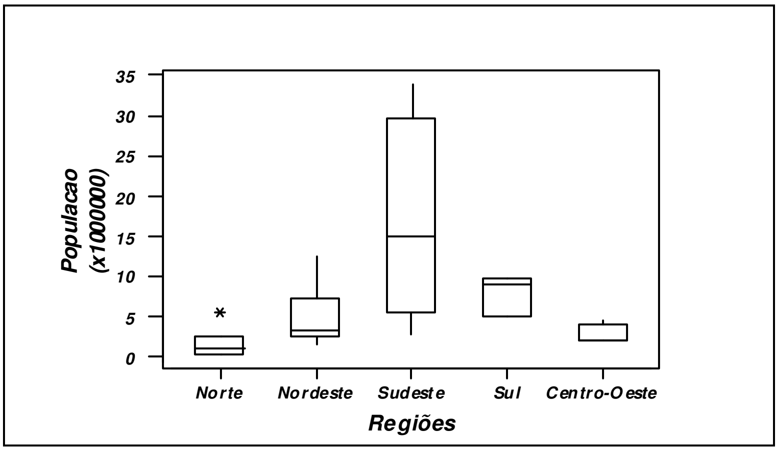

Solteiros




Casados

Os gráficos de dispersão não mostram tendências particulares.

## Problema 37.




Os boxplots acima mostram que todas as distribuições são assimétricas, sendo que a região Sul se destaca pelo seu aspecto peculiar. A região Sudeste tem variabilidade maior, pela inclusão do estado de São Paulo, que é bastante populoso.

## Problema 38.

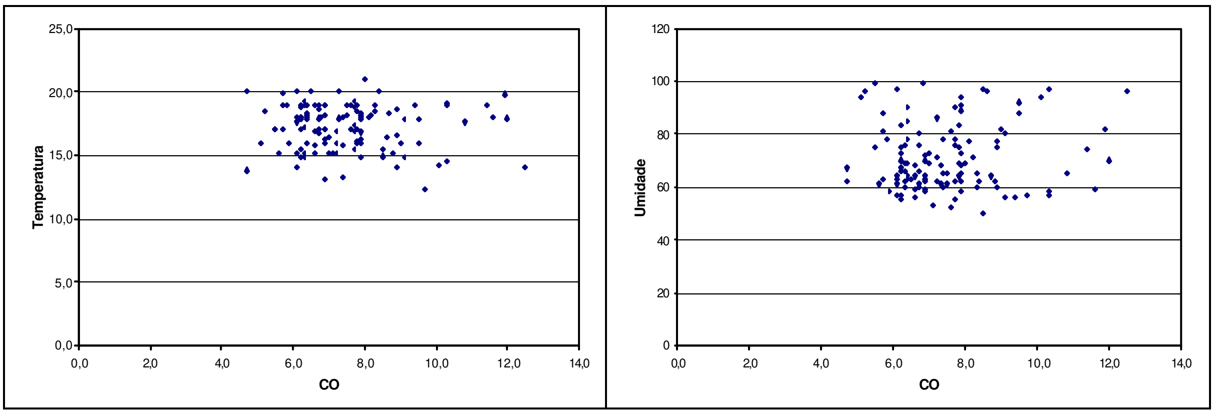
| Telebrás | Ibovespa | | Total |
|---|---|---|---|
| | Baixa | Alta | |
| Baixa | 14 (5,4) | 0 (8,6) | 14 |
| Alta | 1 (9,6) | 24 (15,4) | 25 |
| Total | 15 | 24 | 39 |

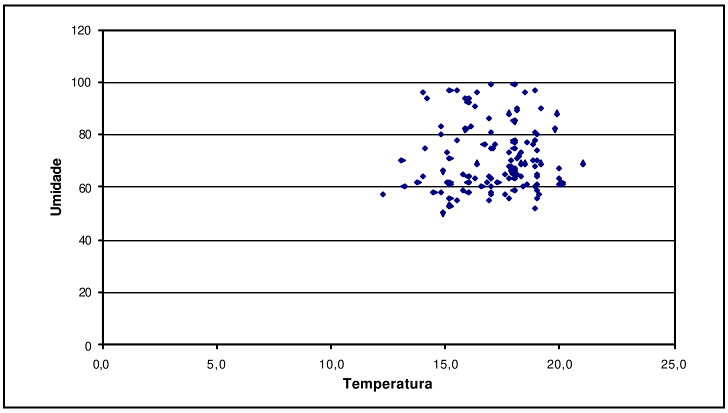
$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} = 34{,}83$$

$$T = \sqrt{\frac{\chi^2 / n}{(r-1)\times(s-1)}} = 0{,}945$$

Logo, percebe-se grande associação entre os preços das ações da Telebrás e Ibovespa.

## Problema 39.




120
100
80
60
40
20
0
Umidade
0,0
5,0
10,0
15,0
20,0
25,0
Temperatura