

Modulo II - Introdução à Estatística Básica

Umberto Mignozzetti

6/1/2020

Modulo II

Análise de dados

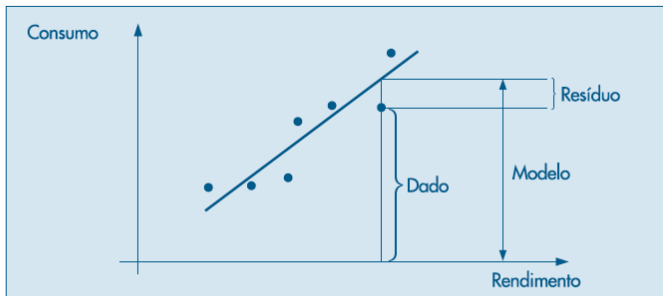
Análise de dados

- ▶ Objetivo Stat é analisar dados.
- ▶ Três etapas:
 1. Entender os dados: análise descritiva
 2. Modelar os dados: probabilidade
 3. Formular hipóteses: inferencia estatística

Análise de dados

- Modelagem: propor uma representação que explique a maior parte da variabilidade dos dados

Figura 1.1: Relação entre consumo e rendimento.



Podemos, então, escrever de modo esquemático:

$$\text{Dados} = \text{Modelo} + \text{Resíduos}$$

Análise de dados

- ▶ Gráficos: visualizar os dados que temos.
- ▶ Objetivos:
 - ▶ Buscar padrões
 - ▶ Checar expectativas
 - ▶ Descobrir fenômenos
 - ▶ Confirmar suposições
 - ▶ Apresentar resultados
- ▶ Altamente recomendável!

Análise de dados

- ▶ Softwares estatísticos:
 - ▶ R / S+
 - ▶ SPSS / PSPP
 - ▶ Excel / Calc
 - ▶ SAS
 - ▶ Stata
- ▶ Qual usar? Qual vc preferir. (Esse tipo de pergunta importa?!)
- ▶ Eu uso R. Motivo: de graça e bom!

Medidas Resumo

Tipos de variáveis

- ▶ Qualitativas: descrevem atributos dos casos:
 - ▶ Pessoa casada
 - ▶ Votou no Bolsonaro
 - ▶ Cidade com mais Corona no Brasil
 - ▶ Superior completo. . .
- ▶ Quantitativas: realizações de uma contagem / mensuração
 - ▶ Idade
 - ▶ Renda
 - ▶ Numero de ligações

Tipos de variáveis

- ▶ Qualitativas:
 - ▶ Nominais: sexo
 - ▶ Ordinais: escolaridade
- ▶ Quantitativas:
 - ▶ Discretas: numero de filhos
 - ▶ Contínuas: salário

Tipos de variáveis

Classifique o banco:

```
dat <- read.csv('https://raw.githubusercontent.com/umberton  
head(dat)
```

```
##      N Estado.Civil  Grau.de.Instrução N.de.Filhos Salario  
## 1 1      solteiro  ensino fundamental          NA  
## 2 2      casado    ensino fundamental           1  
## 3 3      casado    ensino fundamental           2  
## 4 4      solteiro          ensino médio          NA  
## 5 5      solteiro  ensino fundamental          NA  
## 6 6      casado    ensino fundamental           0  
##      Região.de.Procedência  
## 1              interior  
## 2              capital  
## 3              capital  
## 4              outra  
## 5              outra  
## 6              interior
```

Tabela de frequência

- ▶ Contagem de valores para cada um dos níveis pré-definidos
- ▶ E.g., Grau de Instrução:

Tabela 2.2: Frequências e porcentagens dos 36 empregados da seção de orçamentos da Companhia MB segundo o grau de instrução.

Grau de instrução	Frequência n_i	Proporção f_i	Porcentagem $100 f_i$
Fundamental	12	0,3333	33,33
Médio	18	0,5000	50,00
Superior	6	0,1667	16,67
Total	36	1,0000	100,00

Figure 2: f2

Tabela de frequência

Tabela 2.2: Freqüências e porcentagens dos 36 empregados da seção de orçamentos da Companhia MB segundo o grau de instrução.

Grau de instrução	Freqüência n_i	Proporção f_i	Porcentagem $100 f_i$
Fundamental	12	0,3333	33,33
Médio	18	0,5000	50,00
Superior	6	0,1667	16,67
Total	36	1,0000	100,00

Figure 3: f2

► Stats:

► Contagem

► Frequencia (relativa): $f_i = \frac{n_i}{n}$

► Porcentagem: $prop_i = 100 \times \frac{n_i}{n}$

Tabela de frequência

- Para uma variável quanti, temos o seguinte:
1. Criamos intervalos
 2. Contamos valores nos intervalos

Tabela 2.4: Frequências e porcentagens dos 36 empregados da seção de orçamentos da Companhia MB por faixa de salário.

Classe de salários	Frequência n_i	Porcentagem $100 f_i$
4,00 ─ 8,00	10	27,78
8,00 ─ 12,00	12	33,33
12,00 ─ 16,00	8	22,22
16,00 ─ 20,00	5	13,89
20,00 ─ 24,00	1	2,78
Total	36	100,00

Fonte: Tabela 2.1.

Gráficos

- Basta colocar as tabelas que montamos em figuras!?

Figura 2.2: Gráfico em barras para a variável Y : grau de instrução.

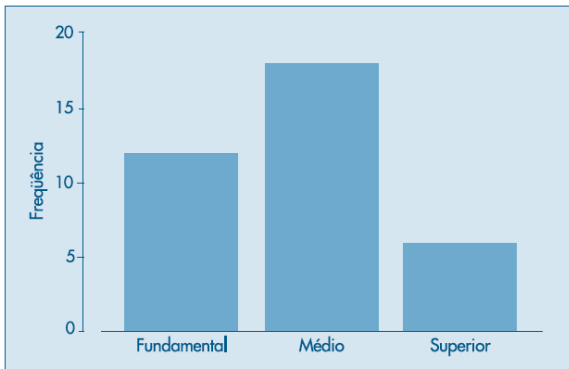
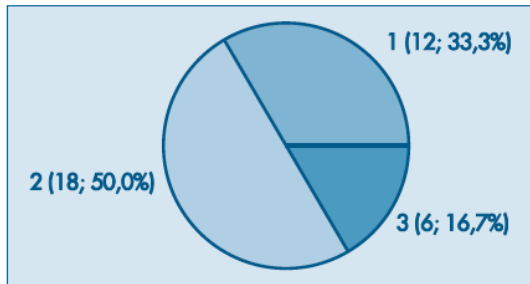


Figure 5: f4

Gráficos

- Basta colocar as tabelas que montamos em figuras!?

Figura 2.3: Gráfico em setores para a variável Y: grau de instrução.



1 = Fundamental, 2 = Médio e 3 = Superior

Figure 6: f5

Gráficos

- Basta colocar as tabelas que montamos em figuras!?

Figura 2.7: Histograma da variável S : salários.

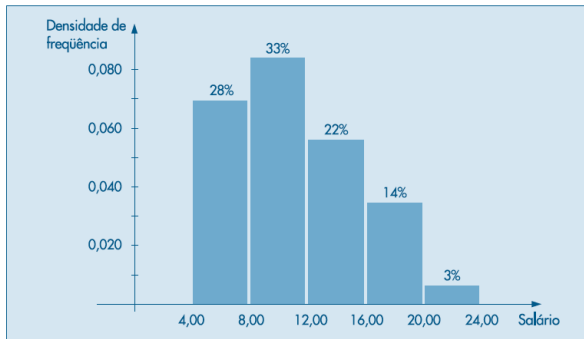


Figure 7: f6

Exercício

6. As taxas médias geométricas de incremento anual (por 100 habitantes) dos 30 maiores municípios do Brasil estão dadas abaixo.

3,67	1,82	3,73	4,10	4,30
1,28	8,14	2,43	4,17	5,36
3,96	6,54	5,84	7,35	3,63
2,93	2,82	8,45	5,28	5,41
7,77	4,65	1,88	2,12	4,26
2,78	5,54	0,90	5,09	4,07

- (a) Construa um histograma.

Figure 8: f7

Medidas-Resumo

Medidas Resumo

- ▶ Dois tipos mais importantes:
 - ▶ Posição
 - ▶ Dispersão
- ▶ Além dessas, temos algumas outras que são boas para analisar os dados.

Medidas de posição

- ▶ Média:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

- ▶ Média (com frequências relativas):

$$\bar{x} = \frac{\sum_{i=1}^k f_i x_i}{n}$$

- ▶ Exercício: calcule a média dos dados: 1,5,2,3,2,4,10

Medidas de posição

- Posição e medidas de ordem: em que lugar está o dado se ordenarmos?

Consideremos, agora, as observações ordenadas em ordem crescente. Vamos denotar a menor observação por $x_{(1)}$, a segunda por $x_{(2)}$, e assim por diante, obtendo-se

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n-1)} \leq x_{(n)}. \quad (3.4)$$

Por exemplo, se $x_1 = 3$, $x_2 = -2$, $x_3 = 6$, $x_4 = 1$, $x_5 = 3$, então $-2 \leq 1 \leq 3 \leq 3 \leq 6$, de modo que $x_{(1)} = -2$, $x_{(2)} = 1$, $x_{(3)} = 3$, $x_{(4)} = 3$ e $x_{(5)} = 6$.

Figure 9: f8

- Ex.:

```
x <- c(3, -2, 6, 1, 3)
```

```
x
```

```
## [1]  3 -2  6  1  3
```

```
sort(x)
```

Medidas de posição

► Mediana:

```
x <- c(3,-2,6,1,3)
```

```
x
```

```
## [1] 3 -2 6 1 3
```

```
sort(x)
```

```
## [1] -2 1 3 3 6
```

```
median(x)
```

```
## [1] 3
```

$$\text{md}(X) = \begin{cases} X_{\left(\frac{n+1}{2}\right)}, & \text{se } n \text{ ímpar;} \\ \frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2}, & \text{se } n \text{ par.} \end{cases}$$

Medidas de dispersão

- Suponha as notas dos alunos em cinco grupos:

grupo A (variável X): 3, 4, 5, 6, 7

grupo B (variável Y): 1, 3, 5, 7, 9

grupo C (variável Z): 5, 5, 5, 5, 5

grupo D (variável W): 3, 5, 5, 7

grupo E (variável V): 3, 5, 5, 6, 6

Figure 11: f10

- Exercício: quais são as médias? Elas ajudam a diferenciar esses dados?

Medidas de dispersão

- ▶ Não ajudam nesses casos: os dados acima eram claramente diferentes!
- ▶ Duas medidas mais usadas: desvio-médio absoluto e variância.

$$\text{dm}(X) = \frac{\sum_{i=1}^n |X_i - \bar{X}|}{n},$$
$$\text{var}(X) = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n},$$

Figure 12: f11

- ▶ Exercício: vamos fazer no R? Considere os dados do exercício acima.

Medidas de dispersão: exercício

1. Quer se estudar o número de erros de impressão de um livro. Para isso escolheu-se uma amostra de 50 páginas, encontrando-se o número de erros por página da tabela abaixo.
- (a) Qual o número médio de erros por página?
 - (b) E o número mediano?
 - (c) Qual é o desvio padrão?
 - (d) Faça uma representação gráfica para a distribuição.
 - (e) Se o livro tem 500 páginas, qual o número total de erros esperado no livro?

Erros	Frequência
0	25
1	20
2	3
3	1
4	1

Figure 13: f12

Quantís empíricos

- ▶ Apenas com média e desvio-padrão não temos ideia do que está acontecendo nos dados:
 - ▶ Valores extremos?
 - ▶ Assimetria?
- ▶ Quantís: boas medidas de resumo dos dados
- ▶ Posição e medidas de ordem: em que lugar está o dado se ordenarmos?

Consideremos, agora, as observações ordenadas em ordem crescente. Vamos denotar a menor observação por $x_{(1)}$, a segunda por $x_{(2)}$, e assim por diante, obtendo-se

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n-1)} \leq x_{(n)}. \quad (3.4)$$

Por exemplo, se $x_1 = 3$, $x_2 = -2$, $x_3 = 6$, $x_4 = 1$, $x_5 = 3$, então $-2 \leq 1 \leq 3 \leq 3 \leq 6$, de modo que $x_{(1)} = -2$, $x_{(2)} = 1$, $x_{(3)} = 3$, $x_{(4)} = 3$ e $x_{(5)} = 6$.

Figure 14: f8

Quantís empíricos

- ▶ Quantís: medidas de posição, para uma dada ordem nos dados.
- ▶ E.g.: mediana: $q(0.5)$: valor que divide os dados pela metade.
- ▶ E.g.: percentil 0.95: $q(0.95)$: valor que divide os dados com 95% dos casos abaixo e 5% acima desse valor.

Quantís empíricos

```
x <- c(15, 5, 3, 8, 10, 2, 7, 11, 12)
sort(x)
```

```
## [1]  2  3  5  7  8 10 11 12 15
```

```
quantile(x)
```

```
##    0%   25%   50%   75%  100%
##     2     5     8    11    15
```

```
quantile(x, probs = 0.95)
```

```
## 95%
## 13.8
```

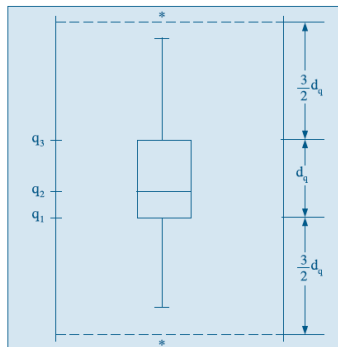
```
summary(x)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      2.000   5.000   8.000   8.111  11.000  15.000
```

Quantís empíricos

- ▶ Box-plot: jeito de apresentar os quantís que dá uma noção da distribuição e dispersão dos dados.
- ▶ $LS = MD + 1.5 \times IIQ$
- ▶ $LI = MD - 1.5 \times IIQ$
- ▶ $IIQ = q(0.75) - q(0.25)$

Figura 3.4: *Box Plot.*



Quantís empírics

► Motivo estatístic

Figura 3.8: Àrea sob a curva normal entre LI e LS.

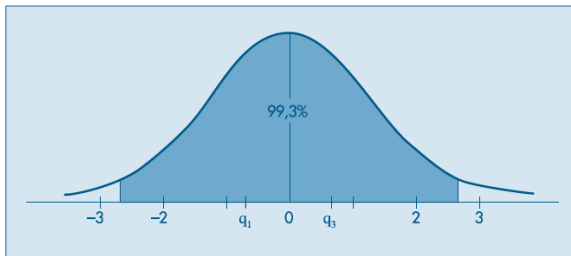


Figure 16: f14

Exercício

- ▶ Faça uma análise dos dados da empresa MB.

Análise bidimensional

Análise Bidimensional

- ▶ Três tipos:
 - ▶ Quali x Quali
 - ▶ Quali x Quant
 - ▶ Quant x Quant

Associação Quali-Quali

Tabela 4.2: Distribuição conjunta das frequências das variáveis grau de instrução (Y) e região de procedência (V).

$V \backslash Y$	Ensino Fundamental	Ensino Médio	Superior	Total
Capital	4	5	2	11
Interior	3	7	2	12
Outra	5	6	2	13
Total	12	18	6	36

Fonte: Tabela 2.1.

Figure 17: f15

Associação Quali-Quali

Tabela 4.3: Distribuição conjunta das proporções (em porcentagem) em relação ao total geral das variáveis Y e V definidas no texto.

$V \backslash Y$	Fundamental	Médio	Superior	Total
Capital	11%	14%	6%	31%
Interior	8%	19%	6%	33%
Outra	14%	17%	5%	36%
Total	33%	50%	17%	100%

Fonte: Tabela 4.2.

Figure 18: f16

Associação Quali-Quali

Tabela 4.4: Distribuição conjunta das proporções (em porcentagem) em relação aos totais de cada coluna das variáveis Y e V definidas no texto.

$V \backslash Y$	Fundamental	Médio	Superior	Total
Capital	33%	28%	33%	31%
Interior	25%	39%	33%	33%
Outra	42%	33%	34%	36%
Total	100%	100%	100%	100%

Fonte: Tabela 4.2.

Figure 19: f17

Associação Quali-Quali

Tabela 4.7: Distribuição conjunta das frequências e proporções (em porcentagem), segundo o sexo (X) e o curso escolhido (Y).

$Y \backslash X$	Masculino	Feminino	Total
Física	100 (71%)	20 (33%)	120 (60%)
Ciências Sociais	40 (29%)	40 (67%)	80 (40%)
Total	140 (100%)	60 (100%)	200 (100%)

Fonte: Dados hipotéticos.

Figure 20: f18

Associação Quali-Quali

Tabela 4.8: Cooperativas autorizadas a funcionar por tipo e estado, junho de 1974.

Estado	Tipo de Cooperativa				Total
	Consumidor	Produtor	Escola	Outras	
São Paulo	214 (33%)	237 (37%)	78 (12%)	119 (18%)	648 (100%)
Paraná	51 (17%)	102 (34%)	126 (42%)	22 (7%)	301 (100%)
Rio G. do Sul	111 (18%)	304 (51%)	139 (23%)	48 (8%)	602 (100%)
Total	376 (24%)	643 (42%)	343 (22%)	189 (12%)	1.551 (100%)

Fonte: Sinopse Estatística do Brasil — IBGE, 1977.

Figure 21: f19

Associação Quali-Quali

Tabela 4.9: Valores esperados na Tabela 4.8 assumindo a independência entre as duas variáveis.

Estado	Tipo de Cooperativa				Total
	Consumidor	Produtor	Escola	Outras	
São Paulo	157 (24%)	269 (42%)	143 (22%)	79 (12%)	648 (100%)
Paraná	73 (24%)	124 (42%)	67 (22%)	37 (12%)	301 (100%)
Rio G. do Sul	146 (24%)	250 (42%)	133 (22%)	73 (12%)	602 (100%)
Total	376 (24%)	643 (42%)	343 (22%)	189 (12%)	1.551 (100%)

Fonte: Tabela 4.8.

Figure 22: f20

Associação Quali-Quali

Tabela 4.10: Desvios entre observados e esperados.

Estado	Tipo de Cooperativa			
	Consumidor	Produtor	Escola	Outras
São Paulo	57 (20,69)	-32 (3,81)	-65 (29,55)	40 (20,25)
Paraná	-22 (6,63)	-22 (3,90)	59 (51,96)	-15 (6,08)
Rio G. do Sul	-35 (8,39)	54 (11,66)	6 (0,27)	-25 (8,56)

Fonte: Tabelas 4.8 e 4.9.

Figure 23: f21

Associação Quali-Quali

6. Uma companhia de seguros analisou a frequência com que 2.000 segurados (1.000 homens e 1.000 mulheres) usaram o hospital. Os resultados foram:

	Homens	Mulheres
Usaram o hospital	100	150
Não usaram o hospital	900	850

- (a) Calcule a proporção de homens entre os indivíduos que usaram o hospital.
- (b) Calcule a proporção de homens entre os indivíduos que não usaram o hospital.
- (c) O uso do hospital independe do sexo do segurado?

Figure 24: f22b

Quanti x Quanti

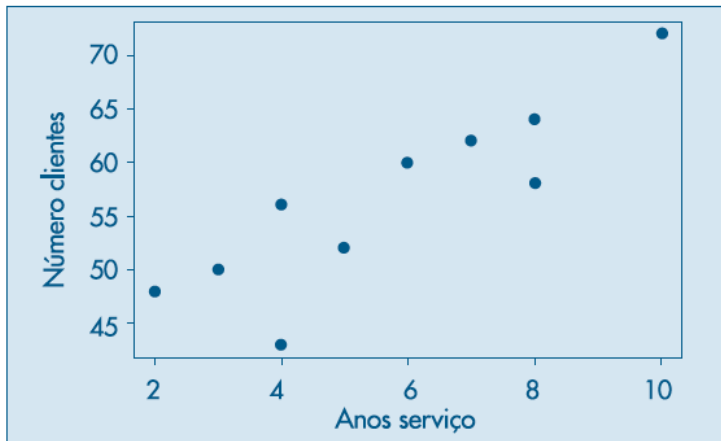
Quanti x Quanti: Correlação

- ▶ A medida principal de associação entre duas variáveis quanti é o coeficiente de correlação.
- ▶ O coeficiente de correlação é uma medida que varia entre -1 e 1 onde:
 - ▶ Mais próximo de -1 significa correlação negativa
 - ▶ Mais próximo de 0 significa ausência de correlação
 - ▶ Mais próximo de 1 significa correlação positiva
- ▶ Mas como funciona a correlação?

Quanti x Quanti: Correlação

- ▶ Diagrama de dispersão: ajuda a observar os dados.
 - ▶ Como fazer: colocar os dados em dois eixos coordenados.

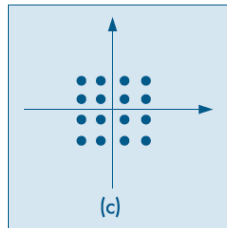
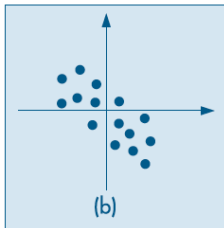
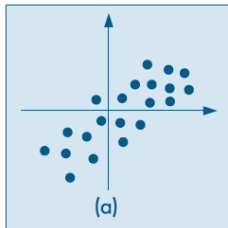
Figura 4.2: Gráfico de dispersão para as variáveis X : anos de serviço e Y : número de clientes.



Quanti x Quanti: Correlação

- ▶ Como correlação aparece no diagrama de dispersão:
 - ▶ Na esquerda temos uma correlação positiva.
 - ▶ No centro, correlação negativa.
 - ▶ Na direita, correlação zero.
- ▶ Para calcular correlação precisamos medir a concentração dos dados nos quadrantes.

Figura 4.6: Tipos de associações entre duas variáveis.



Quanti x Quanti: Correlação

- Para isso, fazemos os seguintes passos:
 1. Subtraímos a média do valor de cada variável. Isso centraliza a variável no zero (fig. esquerda).
 2. Dividimos pelo desvio-padrão: isso faz com que a unidade de variação da variável desapareça (fig. direita).

Figura 4.7: Mudança de escalas para o cálculo do coeficiente de correlação.

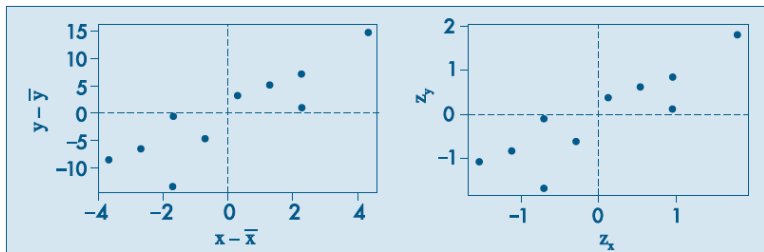


Figure 27: f25

Quanti x Quanti: Correlação

► Olha como fica na tabela:

Tabela 4.15: Cálculo do coeficiente de correlação.

Agente	Anos x	Clientes y	$x - \bar{x}$	$y - \bar{y}$	$\frac{x - \bar{x}}{dp(x)} = z_x$	$\frac{y - \bar{y}}{dp(y)} = z_y$	$z_x \cdot z_y$
A	2	48	-3,7	-8,5	-1,54	-1,05	1,617
B	3	50	-2,7	-6,5	-1,12	-0,80	0,846
C	4	56	-1,7	-0,5	-0,71	-0,06	0,043
D	5	52	-0,7	-4,5	-0,29	-0,55	0,160
E	4	43	-1,7	-13,5	-0,71	-1,66	1,179
F	6	60	0,3	3,5	0,12	0,43	0,052
G	7	62	1,3	5,5	0,54	0,68	0,367
H	8	58	2,3	1,5	0,95	0,19	0,181
I	8	64	2,3	7,5	0,95	0,92	0,874
J	10	72	4,3	15,5	1,78	1,91	3,400
Total	57	565	0	0			8,769

$$\bar{x} = 5,7,$$

$$dp(X) = 2,41,$$

$$\bar{y} = 56,5,$$

$$dp(Y) = 8,11$$

Figure 28: f24

Quanti x Quanti: Correlação

► E essa é a fórmula:

Definição. Dados n pares de valores $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, chamaremos de coeficiente de correlação entre as duas variáveis X e Y a

$$\text{corr}(X, Y) = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{dp(X)} \right) \left(\frac{y_i - \bar{y}}{dp(Y)} \right), \quad (4.7)$$

ou seja, a média dos produtos dos valores padronizados das variáveis.

Não é difícil provar que o coeficiente de correlação satisfaz

$$-1 \leq \text{corr}(X, Y) \leq 1. \quad (4.8)$$

Figure 29: f26

Exercício

11. Abaixo estão os dados referentes à porcentagem da população economicamente ativa empregada no setor primário e o respectivo índice de analfabetismo para algumas regiões metropolitanas brasileiras.

Regiões metropolitanas	Setor primário	Índice de analfabetismo
São Paulo	2,0	17,5
Rio de Janeiro	2,5	18,5
Belém	2,9	19,5
Belo Horizonte	3,3	22,2
Salvador	4,1	26,5
Porto Alegre	4,3	16,6
Recife	7,0	36,6
Fortaleza	13,0	38,4

Fonte: Indicadores Sociais para Áreas Urbanas — IBGE — 1977.

- (a) Faça o diagrama de dispersão.
- (b) Você acha que existe uma dependência linear entre as duas variáveis?
- (c) Calcule o coeficiente de correlação.
- (d) Existe alguma região com comportamento diferente das demais? Se existe, elimine o valor correspondente e recalcule o coeficiente de correlação.

Figure 30: f26b

Quali x Quanti: ANOVA

Quali x Quanti: ANOVA

- Cruzar uma quali versus uma quanti é mais complicado.
- Note esse cruzamento, entre salário e escolaridade:

Figura 4.8: Box plots de salário segundo grau de instrução.

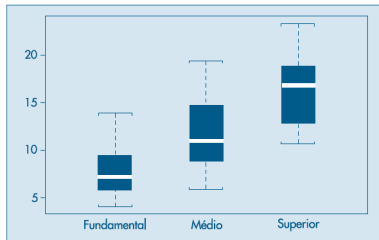


Figure 31: f34

Tabela 4.16: Medidas-resumo para a variável salário, segundo o grau de instrução, na Companhia MB.

Grau de instrução	<i>n</i>	\bar{x}	$dip(S)$	$var(S)$	s_{00}	q_1	q_2	q_3	s_{00}
Fundamental	12	7,84	2,79	7,77	4,00	6,01	7,13	9,16	13,65
Médio	18	11,54	3,62	13,10	5,73	8,84	10,91	14,48	19,40
Superior	6	16,48	4,11	16,89	10,53	13,65	16,74	18,38	23,30
Todos	36	11,12	4,52	20,46	4,00	7,55	10,17	14,06	23,30

Figure 32: f35

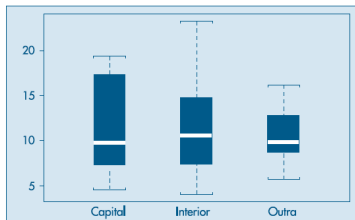
Quali x Quanti: ANOVA

- Compare agora com esse cruzamento, entre salário e região de procedência:

Tabela 4.17: Medidas-resumo para a variável salário segundo a região de procedência, na Companhia MB.

Região de procedência	n	\bar{s}	$dp(S)$	$var(S)$	$s_{(1)}$	q_1	q_2	q_3	$s_{(n)}$
Capital	11	11,46	5,22	27,27	4,56	7,49	9,77	16,63	19,40
Interior	12	11,55	5,07	25,71	4,00	7,81	10,64	14,70	23,30
Outra	13	10,45	3,02	9,13	5,73	8,74	9,80	12,79	16,22
Todos	36	11,12	4,52	20,46	4,00	7,55	10,17	14,06	23,30

Figura 4.9: Box plots de salário segundo região de procedência.



Quali x Quanti: ANOVA

- ▶ Salário x Idade: aparentemente mais relacionadas
- ▶ Salário x região: aparentemente menos relacionadas
- ▶ Como medir a associação? Note as variâncias!

Quali x Quanti: ANOVA

- ▶ Se compararmos as variâncias dentro dos grupos, com a variância total, temos uma medida!
 - ▶ Essa medida é o quanto a nossa quali consegue explicar nossa quanti.
 - ▶ O nome disso é ANOVA: ANalysis Of VAriance.
 - ▶ Como fazer?

Quali x Quanti: ANOVA

- ▶ A variância intra grupo é a média ponderada da variância em cada grupo.
- ▶ A gente calcula assim:
 - ▶ Variância intra-grupo:

$$\overline{var(X)} = \frac{\sum_{i=1}^k n_i var_i(S)}{\sum_{i=1}^k n_i}$$

- ▶ Variância total: variância dos dados: $var(X)$...

Quali x Quanti: ANOVA

- Associação entre as variáveis: R^2 :

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

Quali x Quanti: ANOVA

Exemplo 4.9. Voltando aos dados do Exemplo 4.8, vemos que para a variável S na presença de grau de instrução, tem-se

$$\overline{\text{var}(S)} = \frac{12(7,77) + 18(13,10) + 6(16,89)}{12 + 18 + 6} = 11,96,$$

$$\text{var}(S) = 20,46,$$

de modo que

$$R^2 = 1 - \frac{11,96}{20,46} = 0,415,$$

e dizemos que 41,5% da variação total do salário é *explicada* pela variável grau de instrução.

Para S e região de procedência temos

$$\overline{\text{var}(S)} = \frac{11(27,27) + 12(25,71) + 13(9,13)}{11 + 12 + 13} = 20,20,$$

e, portanto,

$$R^2 = 1 - \frac{20,20}{20,46} = 0,013,$$

Figure 34: f37

Exercício

- ▶ Calcule no banco MB a associação entre estado civil e idade.
- ▶ Faça análises uni e bidimensional de uma das bases de dados que temos na pasta (escolha sua):
 - ▶ MB
 - ▶ voteincome
 - ▶ PErisk

Px aula: probabilidade!