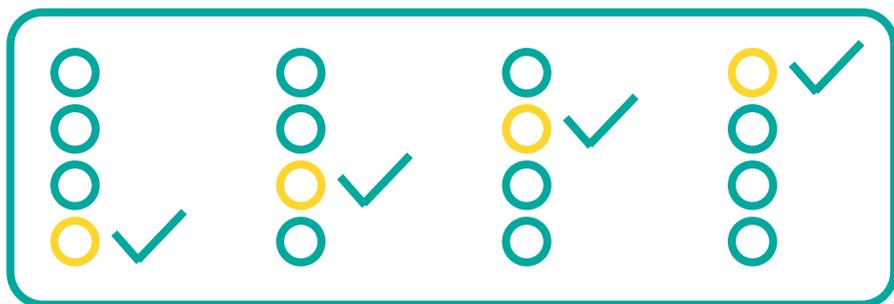


# GABARITO DAS AUTOATIVIDADES



## MÉTODOS QUANTITATIVOS

Prof.<sup>a</sup> Débora Cristina Brandt

2014

## UNIDADE 1

### TÓPICO 1

**Questão única: Analise as sentenças a seguir e classifique V para as verdadeiras ou F para as falsas.**

- a) (F) Quando falamos em pesquisa estatística, estamos nos referindo necessariamente à aplicação de questionários para pessoas responderem.
- b) (V) A estatística descritiva é o cálculo de medidas que permite descrever, com detalhes, o fenômeno que está sendo analisado (CASTANHEIRA, 2008).
- c) (F) A definição do problema pode ser redefinida após a coleta e apuração dos dados, caso as informações obtidas caminhem em outra direção do que a inicialmente delimitada.
- d) (F) Sempre que possível, devemos trabalhar com dados primários, isto é, obtidos por meio de coleta direta: eles se adequam melhor ao objetivo da pesquisa.
- e) (V) Dados são informações que podem ser coletadas de diversas maneiras diferentes como, por exemplo, fichas médicas, registros oficiais, questionários e exames laboratoriais.

---

### TÓPICO 2

**1 Analise as situações a seguir e indique quem compõe a população e, quando for o caso, a amostra.**

- a) **A Polícia Rodoviária Federal quer divulgar o total de acidentes com vítimas nas rodovias federais no último feriado. Para isso, coleta o número de chamados atendidos pelas suas guaritas para este tipo de ocorrência.**

R.: População: chamados atendidos. Amostra: não tem.

- b) **Uma fábrica de refrigerantes quer medir a variação na quantidade de líquido em suas garrafas de 300 ml. Para isso, coleta uma garrafa de cada caixa produzida em um dia de trabalho.**

## MÉTODOS QUANTITATIVOS

R.: População: garrafas de 300 ml. Amostra: uma garrafa por caixa produzida no dia escolhido.

**c) O Estado de Santa Catarina quer saber a eficácia da última campanha de vacinação contra a paralisia. Para isso realiza uma pesquisa com as mães de crianças que procuram o posto de saúde em determinado período.**

R.: População: crianças em idade de vacinação. Amostra: crianças em idade de vacinação que frequentaram o posto de saúde em determinado período.

**d) Um administrador quer estudar o rendimento das ações de certa companhia no último mês. Para isso, considera os dados de fechamento destas ações no período.**

R.: População: valor de fechamento das ações da companhia durante o último mês. Amostra: não tem.

**2 Analise as situações a seguir e indique qual é o tipo de amostragem considerado.**

**a) Astrônomos querem estudar a composição do solo de Marte. Para isso, coletam uma amostra do solo por meio da sonda espacial.**

R.: Amostragem por conveniência.

**b) Professores de um cursinho querem saber quais os cursos que serão mais procurados pelos estudantes de terceiro ano do ensino médio de sua cidade. Para isso, sorteiam algumas escolas, escolhem uma turma de cada uma destas escolas por turno e para realizam um questionário com todos os alunos.**

R.: Amostragem por grupos.

**c) Um jornal percebeu uma queda no número de assinantes no último ano. Para verificar a causa, considerou a lista de antigos clientes em ordem alfabética e escolheu aleatoriamente 50 nomes nesta lista, e realizou ligações telefônicas perguntando o motivo da não renovação da assinatura.**

R.: Amostragem aleatória simples.

## MÉTODOS QUANTITATIVOS

**d) Uma academia de ginástica quer oferecer uma nova modalidade de atividade de física, mas não sabe bem ao certo qual. Então, escolheu aleatoriamente alguns de seus clientes. Como 70% dos frequentadores são mulheres, levou esta proporção em consideração na hora de compor sua amostra.**

R.: Amostragem ponderada.

**e) O governo quer saber qual é a renda *per capita* média da população brasileira para saber em quais setores deve oferecer subsídios. Como suspeita de que há diferenças consideráveis em relação aonde a população mora, resolve tomar amostras contemplando todos os estados, e levando este fato em consideração.**

R.: Amostra estratificada.

**3 Analise e indique a que tipo de erro as situações a seguir estão propensas.**

**a) Uma montadora de automóveis quer saber quais são os itens, que não são de série, mais valorizados na hora da compra. Para isso, escolhe os clientes que adquiram seus automóveis nos meses de maio e junho e realiza via telefone um questionário.**

R.: Erro de amostragem, pois naqueles meses pode ter havido menor volume de vendas em relação aos outros, erro de delineamento, pois o comportamento de venda nestes meses pode ser bem diferente em relação aos outros, erro de resposta e erro de falta de resposta.

**b) A mesma montadora quer saber o nível de satisfação dos clientes com o pós venda de suas lojas. Para isso, seleciona clientes que vêm às concessionárias para a revisão do veículo durante um ano e aplica um questionário.**

R.: Erro de resposta e erro de falta de resposta.

**c) Um laboratório quer testar uma nova vacina em animais. Para isso considera uma amostra de 16 ratos doentes, aplica a vacina contendo o medicamento em metade deles, aplica uma solução de soro fisiológico na outra metade e observa a evolução da doença.**

R.: Erro de resposta, porque alguns ratos podem não reagir ao medicamento, por causa de outro fator.

- d) Uma rede de supermercado quer dimensionar o nível de satisfação dos clientes que fazem uso do serviço de entrega oferecido pela rede. Para isso, escolhe no seu cadastro 30 clientes e realiza um questionário por telefone.**

R.: Erro de resposta, erro de falta de resposta, erro de amostragem, visto que está considerando o número de clientes da rede.



### TÓPICO 4

- 1 Classifique as variáveis a seguir como variáveis qualitativas (nominais ou ordinais) ou quantitativas (discretas ou contínuas), lembrando que esta classificação depende do contexto do problema.**

- a) Em uma pesquisa, pede-se para o entrevistado escrever o algarismo 1 em um campo, caso seja do sexo masculino, e 2, caso seja do sexo feminino.**

R.: A variável em questão é qualitativa nominal.

- b) Um laboratório está testando um novo material para sacolas plásticas biodegradáveis, e está medindo a quantidade de tempo que ele leva para se desintegrar completamente no meio ambiente.**

R.: A variável em questão é quantitativa, e pode ser discreta ou contínua, dependendo da maneira como o tempo é medido.

- c) Após inúmeras denúncias, a fiscalização municipal está medindo o tempo médio de espera entre um ônibus e outro de determinada linha.**

R.: A variável em questão pode ser tanto quantitativa discreta como quantitativa contínua, dependendo da maneira como o tempo for contado.

- d) Uma pesquisa sobre saúde pública pergunta se a pessoa é fumante ou não e, se for, quantos cigarros fuma por dia.**

R.: Duas variáveis envolvidas: a primeira, qualitativa nominal, a segunda, quantitativa discreta.

## MÉTODOS QUANTITATIVOS

e) O Procon está de olho na alta de preços do material escolar. Para isso, está fazendo uma pesquisa na cidade, comparando os preços de lápis, caneta, borracha e cadernos nos estabelecimentos comerciais de uma cidade.

R.: A variável preço é quantitativa contínua.

f) A organização de uma maratona quer medir a quantidade de peso perdida pelos atletas no decorrer da prova. Para isso, realiza uma medição no início e no fim da prova, e classifica a perda de peso como leve, moderada ou forte ao final.

R.: Duas variáveis envolvidas: a variável peso é quantitativa contínua neste caso, e a variável classificação é qualitativa ordinal.

g) Uma loja de departamento quer saber o grau de satisfação dos seus clientes com o atendimento recebido. Para isso, estabelece um critério de satisfação, que varia de 1 a 5, sendo 1 totalmente insatisfeito e 5 totalmente satisfeito.

R.: Variável qualitativa ordinal.

2 Arredonde os números a seguir para 2 casas decimais depois da vírgula segundo as regras estabelecidas pela ABNT:

a) 203,1

R.: 203,10

b) 444,444

R.: 444,44

c) 592,55

R.: 592,55

d) 5,456

R.: 5,46

e) 78,885

R.: 78,88

f) **85,1150**

R.: 85,12

g) **101,144**

R.: 101,14

h) **54,165**

R.: 54,16

i) **45,1651**

R.: 45,17

j) **56**

R.: 56,00

**3 Repita o exercício anterior, considerando 3 casas decimais depois da vírgula:**

a) **0,0000002**

R.: 0,000

b) **10,000009**

R.: 10,000

c) **10000**

R.: 10000,000

d) **103,3465**

R.: 103,346

e) **45,5555**

R.: 45,556

## MÉTODOS QUANTITATIVOS

f) **45,555**

R.: 45,555

g) **45,55555**

R.: 45,556

h) **12,5551**

R.: 12,555

i) **13,3091**

R.: 13,309

j) **14,0009**

R.: 14,001



### TÓPICO 4

**1 Classifique as séries estatísticas a seguir como série numérica, série territorial, série categórica ou série mista.**

a) NÚMERO DE CONCESSIONÁRIAS COM TRÁFEGO PEDAGIADO POR PROGRAMA - 2013

Programa	Número de concessionárias
Federal	14
São Paulo	19
Paraná	6
Rio Grande do Sul	7
Pernambuco/Bahia/Espirito Santo/Minas Gerais/Rio de Janeiro	8
FONTE: ABCR	

R.: Série Territorial

## MÉTODOS QUANTITATIVOS

### b) ESTADO GERAL DAS RODOVIAS - EXTENSÃO PÚBLICA - 2010

Ano	Estado Geral				
	Ótimo	Bom	Regular	Ruim	Péssimo
2010	7,10%	25,30%	37,60%	20,50%	9,50%
2011	5,60%	28,20%	34,20%	21,50%	10,50%
2012	3,20%	24,60%	37,60%	23,80%	10,80%
2013	2,70%	24,00%	38,40%	25,30%	9,60%

FONTE: ABCR

NOTA: Pesquisa CNT de Rodovias 2010/2011/2012/2013.

R.: Série mista, envolvendo a série temporal e a série categórica.

### c) ÍNDICE DO CUSTO DE VIDA - SEGUNDO SEMESTRE 2003

Mês	Índice (%)
Julho	0,35
Agosto	-0,15
Setembro	1,26
Outubro	0,47
Novembro	0,26
Dezembro	0,32

FONTE: Portal Brasil

NOTA: Base de dados Portal Brasil e Fundação Getúlio Vargas.

R.: Série histórica

### d) NÍVEL DE SATISFAÇÃO COM O GOVERNO - 2001

Categoria	Nível (%)
Péssimo	8,10
Ruim	19,20
Regular	32,40
Bom	22,30
Muito Bom	6,60
Não souberam opinar	5,40
Não quiseram opinar	6,00

FONTE: Dados fictícios

- 2 Observe a série estatística a seguir e faça a correspondência entre os números e os conceitos correspondentes.**

### ÁREA TERRITORIAL DOS ESTADOS DA REGIÃO SUL (1)

Estado (2)	Área (km <sup>2</sup> ) (3)
Paraná (4)	199.307,922 (5)
Santa Catarina (6)	95.736,165 (7)
Rio Grande do Sul (*) (8)	281.730,223 (9)
FONTE: IBGE (10)	

Nota: A data da medição não foi informada. (11)

(\*) Inclusive 10.152,251km<sup>2</sup> e 2.811,552km<sup>2</sup> referentes às Lagoas dos Patos e Mirim, respectivamente, incorporadas à área do Estado segundo Constituição Estadual de 1988, não constituindo área municipal. (12)

- R.:(1): Título  
(2), (4), (6) e (8): Coluna Indicadora  
(2) e (3): cabeçalho  
(4): célula  
(4) e (5): linha  
(4), (6) e (8): coluna  
De (4) a (9): corpo  
(10): fonte  
(11): nota  
(12): chamada

## UNIDADE 2

### TÓPICO 1

- 1 Um banco instalou um caixa eletrônico em um posto de combustível e está observando o número de usuários que vêm utilizando o serviço. Diariamente, o número de clientes que utilizou o serviço nos últimos 32 dias foi:**

15 17 16 15 17 14 17 16 16 17 15  
18 14 17 15 14 15 14 15 16 17 18  
18 17 15 14 18 18 16 15 14

- Primeiro passo: ordenar os dados (rol) – preferível em ordem crescente: achar o menos número = 14 e seguir, verificando quantas vezes um mesmo valor se repete.

## MÉTODOS QUANTITATIVOS

14	14	14	14	14	14
15	15	15	15	15	15
15	15	16	16	16	16
16	16	17	17	17	17
17	17	17	18	18	18
18	18				

Tamanho da amostra = 32, então  
total  $n_i = 32$  (foram 32 dias de  
observações)

**a) Organize uma tabela de frequências (utilize 4 casas decimais).**

R.: (cuidar com regras de arredondamento)

**Montar uma tabela com a frequência absoluta ( $n_i$ ), frequência relativa ( $f_i$ ) e frequência relativa acumulada ( $f_{ai}$ )** - precisa-se da  $f_{ai}$  para responder a letra “b” da questão e para saber a  $f_{ai}$ , precisa-se do  $f_i$ .

N° usuários	$n_i$	$f_i = n_i/\sum n_i$	$f_{ai}$
14	6	0,1875	0,1875
15	8	0,2500	0,4375
16	6	0,1875	0,6250
17	7	0,2188	0,8438
18	5	0,1562	1,0000
TOTAL ( $\Sigma$ )	32	1.0000	

$n_i$  = frequência de dias em que apareceram n° “X” de usuários no caixa eletrônico.

$f_{ai}$  = acumulado de  $f_i$

**b) Qual é a porcentagem das observações está abaixo de 16 dias?**

R.: 43,75% das observações (durante 6 dias obteve-se 14 usuários/dia e durante 8 dias obteve-se 15 usuários/dia).

$6 + 8 = 14$  dias (abaixo de 16 dias = 43,75% = valor da  $f_{ai}$ )

$6 + 8 + 6 = 20$  dias (entre 6 e 20 dias, ou abaixo de 21 dias = **62,50%**)

**2 Um posto de saúde de certo bairro mantém um arquivo com o número de pacientes que procuram o consultório odontológico diariamente. Os dados são os seguintes:**

3, 4, 3, 4, 5, 1, 6, 3, 4, 5, 3, 4, 3, 3, 4, 3, 5, 5, 5, 5, 6, 11, 10, 2, 1, 2, 3, 1, 5, 2.

# MÉTODOS QUANTITATIVOS

## Organize uma tabela de frequência.

R.: Primeiro passo: ordenar os dados (rol) – preferível em ordem crescente: achar o menor número = 1 e seguir, verificando quantas vezes um mesmo valor se repete.

1	1	1	2	2	2
3	3	3	3	3	3
3	3	4	4	4	4
4	5	5	5	5	5
5	5	6	6	10	11

Utilizou-se 2 casas após a vírgula

frequência de dias, ou seja, em 3 dias (dos 30 dias analisados) foram atendidos 1 usuário por dia, em outros 3 dias (dos 30 dias) foram atendidos mais 2 usuários por dia...

usuários	ni	fi	fai
1	3	0,10	0,10
2	3	0,10	0,20
3	8	0,27	0,47
4	5	0,17	0,63
5	7	0,23	0,87
6	2	0,07	0,93
10	1	0,03	0,97
11	1	0,03	1,00
	30	1,00	

quantidade de usuários

em 3% dos dias trabalhados se atendeu 10 pacientes por dia

ni = frequência de dias em que apareceram n° "x" de pacientes /usuários no posto

fi = ni/Σni

fai = acumulado de fi

## MÉTODOS QUANTITATIVOS

**3 (MAGALHAES, 2010) O tempo de utilização de caixas eletrônicos depende de cada usuário e das operações efetuadas. Foram coletadas 26 medidas desse tempo (minutos):**

1,1 1,2 1,7 1,5 0,9 1,3 1,4 1,6 1,7 1,6 1,0 0,8 1,5  
1,3 1,7 1,6 1,4 1,2 1,2 1,0 0,9 1,8 1,7 1,5 1,3 1,5

**a) Organize uma tabela de frequência sem agrupar dados.**

R.: Tempo de uso dos caixas eletrônicos:

Tempo de uso (minutos)	ni	fi = ni/Σni	fai
0,8	1	0,038	0,038
0,9	2	0,077	0,115
1	2	0,077	0,192
1,1	1	0,038	0,231
1,2	3	0,115	0,346
1,3	3	0,115	0,462
1,4	2	0,077	0,538
1,5	4	0,154	0,692
1,6	3	0,115	0,808
1,7	4	0,154	0,962
1,8	1	0,038	1
TOTAL (Σ)	26	1	

FONTES: A autora

Rol:

0,8	0,9	0,9	1,0	1,0	1,1
1,2	1,2	1,2	1,3	1,3	1,3
1,4	1,4	1,5	1,5	1,5	1,5
1,6	1,6	1,6	1,7	1,7	1,7
1,7	1,8				

**b) Agrupe os dados em faixas de 0,2 minutos a partir de 0,8 e obtenha uma nova tabela de frequência.**

**OBS.: Utilize 3 casas decimais depois da vírgula.**

R.: a partir de 0,8 (limite inferior) e 1,0 (limite superior) – na primeira classe.

## MÉTODOS QUANTITATIVOS

Tempo de uso (minutos)	$n_i$	$f_i = n_i/\Sigma n_i$	fai
0,8 † 1,0	3	0,115	0,115
1,0 † 1,2	3	0,115	0,230
1,2 † 1,4	6	0,231	0,461
1,4 † 1,6	6	0,231	0,692
1,6 † 1,8	7	0,269	0,961
1,8 † 2,0	1	0,038	1,000
TOTAL ( $\Sigma$ )	26	1,00	

Rol:

0,8	0,9	0,9	1,0	1,0	1,1
1,2	1,2	1,2	1,3	1,3	1,3
1,4	1,4	1,5	1,5	1,5	1,5
1,6	1,6	1,6	1,7	1,7	1,7
1,7	1,8				

**4 (MAGALHAES, 2010) O valor médio de comercialização da saca de milho de 60 quilos na BM&F é apresentado a seguir, em reais, para os últimos 40 meses.**

6,1 6,2 6,7 6,5 6,9 6,3 7,4 7,6 7,7 7,6  
 7,3 7,7 7,6 7,4 7,2 7,2 7,3 7,6 7,5 7,4  
 7,5 7,7 8,2 8,3 8,1 8,1 8,1 7,9 7,8 7,4  
 7,5 7,6 7,5 7,6 7,4 7,3 7,4 7,5 7,5 7,4

**Organize os dados em faixas de tamanho 0,4 a partir de 6, utilizando quatro casas decimais após a vírgula.**

R.: Rol:

6,1	6,2	6,3	6,5	6,7	6,9	7,2	7,2	7,3	7,3
7,3	7,4	7,4	7,4	7,4	7,4	7,4	7,4	7,5	7,5
7,5	7,5	7,5	7,5	7,6	7,6	7,6	7,6	7,6	7,6
7,7	7,7	7,7	7,8	7,9	8,1	8,1	8,1	8,2	8,3

# MÉTODOS QUANTITATIVOS

R.: Valor médio de comercialização da saca de milho na BM&P, em reais

2014			
	ni	fi	fai
6,0 f 6,4	3	0,0750	0,0750
6,4 f 6,8	2	0,0500	0,1250
6,8 f 7,2	1	0,0250	0,1500
7,2 f 7,6	18	0,4500	0,6000
7,6 f 8,0	11	0,2750	0,8750
8,0 f 8,4	5	0,1250	1,0000
TOTAL	40	1,0000	

## TÓPICO 2

**1 Considere as seguintes tabelas. Qual é o gráfico mais indicado para uma das situações?**

a) GRAU DE INSTRUÇÃO DOS USUÁRIOS DO DESODORANTE X  
- FEVEREIRO 2012

Grau de instrução	Número de usuários
Fundamental	21
Médio	92
Superior	84
Não responderam	3

FONTE: Dados fictícios

R.: Gráfico de setores, colunas ou barras.

b) PERIODICIDADE DE ATIVIDADE FÍSICA DOS USUÁRIOS DO  
DESODORANTE X  
- FEVEREIRO 2012

## MÉTODOS QUANTITATIVOS

Periodicidade	Número de usuários
0	19
1	10
2	56
3	44
4	30
5	23
6	14
7	4

FONTE: Dados fictícios

R.: Gráfico delinha (para setores, têm muitas classificações).

### c) USO DE UM PRODUTO EM DETERMINADA SEMANA NA INGLATERRA

Região	Taxa de mortalidade infantil (%)		
	Não usou	Usou uma vez	Usou mais de uma vez
Norte	20	40	60
Central	60	35	35
Sul	100	60	20
Escócia	10	15	10
País de Gales	25	30	30

FONTE: Silver (2000)

R.: Gráfico de colunas ou barras.

## 2 Trace o gráfico de linha baseado na tabela a seguir:

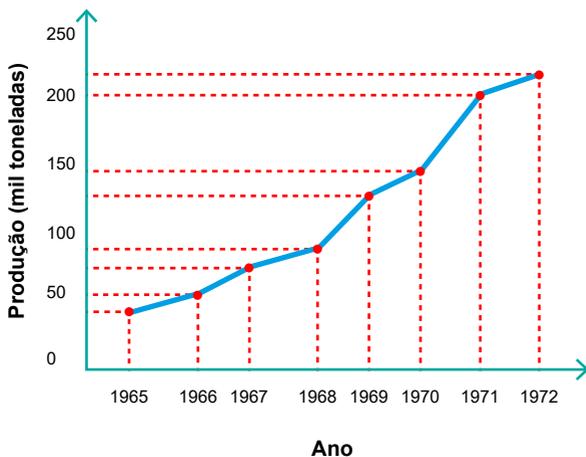
### a) MÉDIA MENSAL DA PRODUÇÃO BRASILEIRA DE CARVÃO BETUMINOSO – 1965 - 1972

Ano	Produção (mil toneladas)
1965	45
1966	50
1967	70
1968	80
1969	130
1970	150
1971	200
1972	210

FONTE: Oliveira (2010)

R.: 1°) traçar os eixos X e Y do plano cartesiano em que os valores de X serão o ano e de Y a quantidade produzida:

## MÉTODOS QUANTITATIVOS



- 2°) juntar ano com quantidade produzida  
 3°) a partir da união dos pontos, traçar a linha

### 3 Trace o gráfico de setores baseado na tabela a seguir:

#### ESTIMATIVAS POPULACIONAIS DO BRASIL - GRANDES REGIÕES - 2000

Região	População
Norte	12.900.704
Nordeste	47.741.711
Sudeste	72.412.411
Sul	25.107.616
Centro-Oeste	11.636.728

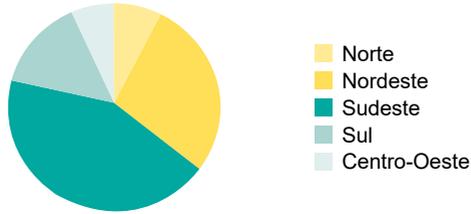
FONTE: CASTANHEIRA, 2008

R.:

Região	ni	fi	(%)	Graus
Norte	12.900.704	0,0760	7,60	27,3
Nordeste	47.741.711	0,2812	28,12	101,2
Sudeste	72.412.411	0,4264	42,64	153,5
Sul	25.107.616	0,1479	14,79	53,2
Centro-Oeste	11.636.728	0,0685	6,85	24,8
TOTAL	169.799.170	1,0000	100	360

# MÉTODOS QUANTITATIVOS

ESTIMATIVAS POPULACIONAIS DO BRASIL – GRANDES REGIÕES – 2000



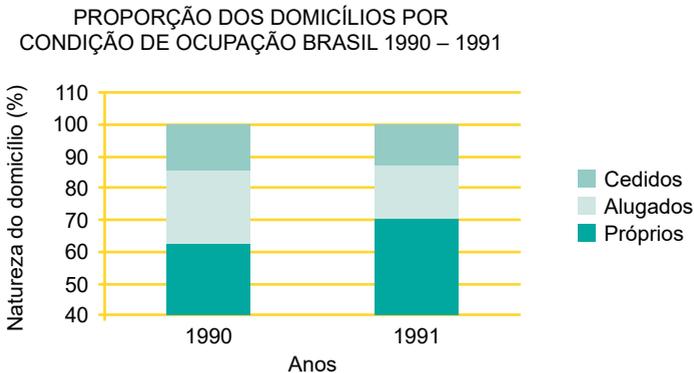
4 (CRESPO, 2005) Represente a tabela por meio de um gráfico de colunas múltiplas.

PROPORÇÃO DOS DOMICÍLIOS POR CONDIÇÃO DE OCUPAÇÃO BRASIL 1990 – 1991

Anos	Natureza		
	Próprios (%)	Alugados (%)	Cedidos (%)
1990	62,7	22,9	14,4
1991	70,3	16,5	13,2

FONTE: IBGE

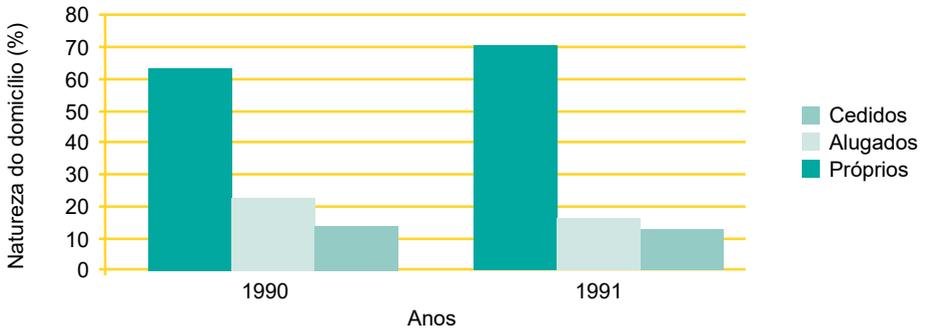
R.:



Ou

# MÉTODOS QUANTITATIVOS

PROPORÇÃO DOS DOMICÍLIOS POR  
CONDIÇÃO DE OCUPAÇÃO BRASIL 1990 – 1991



## TÓPICO 3

1 Estudando uma nova técnica de sutura, foram contados os dias necessários para a completa cicatrização de determinada cirurgia. Os resultados de 25 pacientes foram os seguintes:

6 8 9 7 6 6 7 8 9 10 7 8 10  
9 9 9 7 6 5 7 7 7 8 10 11

Organize os dados numa distribuição de frequências e calcule o tempo médio necessário para a completa cicatrização.

R.:  $n_i$  total = 25

Rol: iniciar pelo menor tempo de cicatrização (neste caso 5 dias), verificando quantas vezes um mesmo valor se repete.

5	6	6	6	6
7	7	7	7	7
7	7	8	8	8
8	9	9	9	9
9	10	10	10	11

## MÉTODOS QUANTITATIVOS

a) Organizar em uma tabela com distribuição de frequência:

$X_i$	$n_i$
5	1
6	4
7	7
8	4
9	5
10	3
11	1
TOTAL	25

$$\bar{x} = \frac{(x_1 \cdot n_1) + (x_2 \cdot n_2) + \dots}{\sum n_i}$$

$$\bar{x} = \frac{(5 \cdot 1) + (6 \cdot 4) + (7 \cdot 7) + (8 \cdot 4) + (9 \cdot 5) + (10 \cdot 3) + (11 \cdot 1)}{25}$$

$\bar{x} = \frac{196}{25} = 7,84$  dias é a média de tempo de cicatrização.

- 1 paciente levou 5 dias para cicatrizar
- 4 pacientes levaram 6 dias
- 7 pacientes levaram 7 dias...

$X_i$  = quantidade de dias para cicatrizar

$n_i$  = frequência (quantidade de pacientes analisados)

b) Achar o Tempo médio = cálculo da média ( $\bar{x}$ ) de tempo para cicatrizar: ou é possível criar uma nova coluna simplificando a fórmula

$x_i$	$n_i$	$x_i \cdot n_i$
5	1	5
6	4	24
7	7	49
8	4	32
9	5	45
10	3	30
11	1	11
TOTAL	25	196

## MÉTODOS QUANTITATIVOS

$$\bar{x} = \frac{\sum(x_i \cdot n_i)}{\sum n_i}$$

$$\bar{x} = \frac{196}{25}$$

**2 O entroncamento entre duas ruas em uma determinada cidade tem alto índice de acidentes de trânsito, conforme pode ser constatado nos últimos 12 meses:**

5 4 7 8 5 6 4 7 9 7 6 8

**Determine a média do número de acidentes mensais neste local.**

R.:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n}$$

x = n° de acidentes por mês

n = total de meses analisados (pois quer a média mensal de acidentes)

$$\bar{x} = \frac{5+4+7+8+5+6+4+7+9+7+6+8}{12} = \frac{76}{12} = 6,33 \text{ acidentes/mês}$$

**3 Em uma clínica cardíaca foram anotados os níveis de colesterol (em mg/100ml) para trinta pacientes, homens com idade entre 40 e 60 anos que foram à clínica fazer um check-up.**

Paciente	1	2	3	4	5	6	7	8	9	10
Colesterol	160	160	161	163	167	170	172	172	173	177

Paciente	11	12	13	14	15	16	17	18	19	20
Colesterol	178	181	181	182	185	186	194	197	199	203

Paciente	21	22	23	24	25	26	27	28	29	30
Colesterol	203	205	206	206	208	209	211	214	218	225

## MÉTODOS QUANTITATIVOS

- Calcule a média, moda e mediana a partir da tabela de frequência.
- Organize os dados em uma tabela de frequência com faixas de tamanho 10 a partir de 160.
- Refaça o item a. utilizando a tabela com intervalos de classe.

R.: Não tem necessidade de fazer o rol (organizar os dados), pois os níveis de colesterol já estão apresentados em ordem crescente. Temos um **ni total de 30 pacientes**.

### a) Achar a média, a moda e a mediana a partir da tabela de frequência:

1°) montar uma tabela de frequência com os níveis de colesterol e sua incidência (frequência).

níveis de colesterol (x)	ni	fi = ni/Σni	fai
160	2	0,067	0,067
161	1	0,033	0,100
163	1	0,033	0,133
167	1	0,033	0,166
170	1	0,033	0,199
172	2	0,067	0,260
173	1	0,033	0,299
177	1	0,033	0,332
178	1	0,033	0,365
181	2	0,067	0,432
182	1	0,033	0,465
185	1	0,033	0,498
186	1	0,033	0,531
194	1	0,033	0,564
197	1	0,033	0,597
199	1	0,033	0,630
203	2	0,067	0,697
205	1	0,033	0,730
206	2	0,067	0,797
208	1	0,033	0,830
209	1	0,033	0,863
211	1	0,033	0,896
214	1	0,033	0,929
218	1	0,033	0,962
225	1	0,033	1,000
TOTAL = Σ	30	1	

## MÉTODOS QUANTITATIVOS

**x = dado analisado (neste caso os níveis de colesterol)**

$$\text{MÉDIA } (\bar{x}) = \frac{(x1 \cdot n1) + (x2 \cdot n2) + \dots}{\sum ni}$$

$$\bar{x} = \frac{(160 \cdot 2) + (161 \cdot 1) + (163 \cdot 1) + \dots}{30}$$

$$\bar{x} = \frac{5666}{30} = 188,9$$

MODA (MO) = valor ou atributo analisado que ocorre com maior frequência.

**MO = 160, 172, 181, 203, 206 (MULTIMODAL = mais de duas modas).**

MEDIANA (Md): é o valor da variável que ocupa a posição central de um conjunto de *n* dados ordenados (50% dos dados estão acima dela e 50% dos dados estão abaixo dela).

OLHAR O ROL:

- IDENTIFICAR O VALOR DO MEIO (QUANDO O N° DA AMOSTRA É ÍMPAR)
- CASO O NÚMERO DA AMOSTRA SEJA PARA PEGAR OS DOIS VALORES CENTRAIS, SOMAR E DIVIDIR POR 2.

**Neste caso temos uma amostra de 30 (par), no rol vamos achar os dois valores centrais, somar e dividir por 2:**

$$\text{Md} = \frac{185 + 186}{2} = \frac{371}{2} = 185,5$$

Paciente	1	2	3	4	5	6	7	8	9	10
Colesterol	160	160	161	163	167	170	172	172	173	177

Paciente	11	12	13	14	15	16	17	18	19	20
Colesterol	178	181	181	182	185	186	194	197	199	203

Paciente	21	22	23	24	25	26	27	28	29	30
Colesterol	203	205	206	206	208	209	211	214	218	225

## MÉTODOS QUANTITATIVOS

### b) Organize os dados em uma tabela de frequência com faixas de tamanho 10 a partir de 160:

1. Fazer uma nova tabela com intervalos de classe de 10 para os valores do colesterol,

E CALCULAR: (vamos precisar desses valores para média, moda e mediana – item c)

fac = acúmulo de ni

fi =  $ni/\sum ni$

fai = acúmulo de fi

xi = ponto médio da classe =  $\frac{\text{lim sup} + \text{lim inf}}{2}$

xi . ni

níveis colesterol (x)	ni	fac	fi = ni/Σni	fai	xi = $\frac{\text{lim sup} + \text{lim inf}}{2}$	ni . xi
160 + 170	5	5	0,17	0,17	165,00	825
170 + 180	6	11	0,20	0,37	175,00	1050
180 + 190	5	16	0,17	0,54	185,00	925
190 + 200	3	19	0,10	0,64	195,00	585
200 + 210	7	26	0,23	0,87	205,00	1435
210 + 220	3	29	0,10	0,97	215,00	645
220 + 230	1	30	0,03	1	225,00	225
total = Σ	30		1			5690

### c) Calcular média, moda e mediana – agora com intervalo de classe (variável contínua):

MÉDIA ARITMÉTICA PARA DADOS AGRUPADOS (variável contínua)

$$\dot{x} = \frac{\sum (xi \cdot ni)}{\sum ni}$$

$$\dot{x} = \frac{5690}{30} = 189,65$$

## MÉTODOS QUANTITATIVOS

MODA PARA DADOS AGRUPADOS (variável contínua)

1°) achar a classe modal (com maior ni)

níveis colesterol (x)	ni	fac	fi = ni/Σni	fai	xi = $\frac{\text{lim sup} + \text{lim inf}}{2}$	ni . xi
160 + 170	5	5	0,17	0,17	165,00	825
170 + 180	6	11	0,20	0,37	175,00	1050
180 + 190	5	16	0,17	0,54	185,00	925
190 + 200	3	19	0,10	0,64	195,00	585
200 + 210	7	26	0,23	0,87	205,00	1435
210 + 220	3	29	0,10	0,97	215,00	645
220 + 230	1	30	0,03	1	225,00	225
total = Σ	30		1			5690

2°) utilizar a fórmula:

$$Mo = \text{linf} + \left( \frac{\Delta a}{\Delta a + \Delta p} \right) \cdot h$$

*l inf* = limite inferior da classe modal = **200**

$\Delta a$  = diferença entre ni da classe modal e a ni da classe anterior = **7 - 3 = 4**

$\Delta p$  = diferença entre a ni da classe modal e a ni da classe posterior = **7 - 3 = 4**

$h$  = amplitude da classe modal (limite superior - limite inferior da classe modal) = **210 - 200 = 10**

$$Mo = 200 + \left( \frac{4}{4 + 4} \right) \cdot 10$$

$$Mo = 200 + \frac{4}{8} \cdot 10$$

$$Mo = 200 + 0,5 \cdot 10$$

$$Mo = 200 + 5$$

$$\mathbf{Mo = 205}$$

## MÉTODOS QUANTITATIVOS

### MEDIANA PARA DADOS COM INTERVALO DE CLASSE (VARIÁVEL CONTÍNUA)

1°) criar a coluna Fac (frequência acumulada simples = acúmulo do ni) ok

2°) achar a classe mediana (valor que divide a pesquisa em 2) =  $\Sigma ni/2$

**$30/2 = 15$  (temos na classe mediana o 15° elemento da amostra)**

3°) encontrar o 15° elemento na coluna Fac

níveis colesterol (x)	ni	fac	fi = ni/ $\Sigma ni$	fai	$xi = \frac{\text{lim sup} + \text{lim inf}}{2}$	ni . xi
160 f 170	5	5	0,17	0,17	165,00	825
170 f 180	6	11	0,20	0,37	175,00	1050
180 f 190	5	16	0,17	0,54	185,00	925
190 f 200	3	19	0,10	0,64	195,00	585
200 f 210	7	26	0,23	0,87	205,00	1435
210 f 220	3	29	0,10	0,97	215,00	645
220 f 230	1	30	0,03	1	225,00	225
total = $\Sigma$	30		1			5690

- **O 15° elemento está na linha 3 com o 16°, pois na linha 2 temos o 11° e na linha 4 o 19° elemento.**

4°) Após identificar a classe mediana (onde está o 15° elemento) vamos substituir na fórmula:

$$md = Li_{md} + \frac{\left[ \left( \frac{\Sigma ni}{2} \right) - \text{fac ant} \right]}{ni_{md}} * h$$

EM QUE:

Li md = limite inferior da classe mediana = **180**

ni md = frequência absoluta da classe mediana = **5**

Fac ant = frequência acumulada simples anterior a classe mediana = **11**

h = amplitude da classe mediana (lim sup – lim inf) = **10**

$$Md = 180 + \frac{[15 - 11]}{5} \cdot 10$$

$$Md = 180 + \frac{4}{5} \cdot 10$$

## MÉTODOS QUANTITATIVOS

$$Md = 180 + 0,8 \cdot 10$$

$$Md = 180 + 8 = 188$$

**4 O índice de germinação é um dos principais fatores para definir a qualidade das sementes. Ele é determinado em experimento científico conduzido pelo fabricante e regulamentado pelos órgãos fiscalizadores. Um fabricante afirma que o índice de germinação de suas sementes de milho é de 85%. Para verificar tal afirmação, uma cooperativa de agricultores sorteou 100 amostras com 100 sementes em cada uma e anotou a porcentagem de germinação em cada amostra.**

Germinação (%)		Frequência
60	- 75	8
75	- 80	20
80	- 85	42
85	- 90	18
90	- 95	10
95	- 100	2

- Calcule a média, a moda, a mediana e os quartis para estes dados.
- Construa o *box plot*.
- Comente a afirmação do fabricante.

R.:

1º passo: calcular:

fac = acúmulo de  $n_i$

$f_i = n_i / \sum n_i$

fai = acúmulo de  $f_i$

$x_i = \text{ponto médio da classe} = \frac{\text{lim sup} + \text{lim inf}}{2}$

## MÉTODOS QUANTITATIVOS

germinação (%)	frequência (ni)	fac	fi	fai	xi	ni . Xi
70 † 75	8	8	0,080	0,08	67,5	540
75 † 80	20	28	0,200	0,280	<b>77,5</b>	<b>1550</b>
80 † 85	<b>42</b>	<b>70</b>	<b>0,420</b>	<b>0,700</b>	<b>82,5</b>	<b>3465</b>
85 † 90	18	88	0,180	0,880	87,5	1575
90 † 95	10	98	0,100	0,980	92,5	925
95 † 100	2	100	0,020	1,000	97,5	195
total	100					8250

a) Calcular a média, a moda, a mediana e os quartis:

**Média para dados com intervalo de classe (variável contínua):**

$$\bar{x} = \frac{\sum (xi \cdot ni)}{\sum ni}$$

$$X = 8250/100 = 82,5$$

A média é 82,5

**Moda para dados agrupados:**

1º) achar a classe modal (com maior ni)

**3ª classe, ni = 42**

*l inf* = limite inferior da classe modal = **80**

$\Delta a$  = diferença entre ni da classe modal e a ni da classe anterior = **42 - 20 = 22**

$\Delta p$  = diferença entre a ni da classe modal e a ni da classe posterior = **42 - 18 = 24**

*h* = amplitude da classe modal (limite superior – limite inferior da classe modal) = **85 - 80 = 5**

2º) utilizar a fórmula:

$$Mo = \text{linf} + \left( \frac{\Delta a}{\Delta a + \Delta p} \right) \cdot h$$

$$Mo = 80 + \left( \frac{22}{22 + 24} \right) \cdot 5 = 80 + 0,48 \cdot 5$$

$$Mo = 80 + 2,4 = 82,4$$

## MÉTODOS QUANTITATIVOS

**Mediana (Md) para dados com intervalo de classe (variável contínua):**

1°) criar a coluna Fac (frequência acumulada simples = acúmulo do ni) = **ok**  
(ver tabela)

2°) achar a classe mediana =  $\Sigma ni/2$

**$\Sigma ni/2 = 100/2 = 50$  = classe mediana é onde se encontra o 50° elemento (valor que divide a pesquisa em 2 partes). Encontrar o 50° elemento olhando a coluna da Fac:**

- Na primeira linha temos até 8 elementos;
- Na segunda linhas temos até 28 elementos;
- Na terceira linha temos até 70 elementos (encontramos a nossa classe, pois se nesta temos até 70 elementos, o 50° elemento se encontra aí, já que nas anteriores estão elementos menores que 50).

3°) usar a fórmula:

$$md = Li_{md} + \frac{\left[ \left( \frac{\Sigma ni}{2} \right) - fac_{ant} \right]}{ni_{md}} \cdot h$$

Li md = limite inferior da classe mediana = **80**

ni md = frequência absoluta da classe mediana = **42**

Fac ant = frequência acumulada simples anterior a classe mediana = **28**

h = amplitude da classe mediana = **85 - 80 = 5**

$$Md = 80 + \frac{[50 - 28]}{42} \cdot 5$$

$$Md = 80 + \frac{22}{42} \cdot 5$$

$$Md = 80 + 0,5238 \cdot 5$$

$$Md = 80 + 2,619$$

$$Md = 82,62 = 2^\circ \text{ quartis}$$

1°, 2° e 3° Quartis: quero saber em qual grau de germinação se encontram 25%, 50% e 75% da amostra, respectivamente.

## MÉTODOS QUANTITATIVOS

Analisar a frequência relativa acumulada (fai):

- 8% está entre 60 e 75
- 28% está entre 60 e 80 (Primeiro Quartil)
- 70% está entre 60 e 85 (Segundo Quartil = mediana)
- 88% está entre 60 e 90 (Terceiro quartil)

1° quartil (Q1):

- Queremos saber o 1° quartil, ou seja, queremos saber qual o grau de germinação de até 25% da amostra (1/4):
  - Deve-se analisar a fai (frequência relativa acumulada) para identificar o valor onde podem estar estes 25%:
- Na segunda linha encontra-se o 1° quartil, pois até 28% da amostra está entre 60 e 80, precisamos achar qual é o valor deste quartil:

$80 - 75 = 5$  (limite sup – lim inf da classe)

5 ---- 20% (fi)

X ----- 25% - 8% (o que eu quero de fai menos a anterior)

Então: 5 ---- 20%

          x ---- 17%

$$20x = 17 \cdot 5$$

$$20x = 85$$

$$X = 85/20 = 4,25$$

**1° quartil = 4,25 + 75**

**Q1 = 79,25% de germinação (1/4 da amostra = 25% tem até 79,25 % de germinação)**

**Soma o resultado da regra de três ao limite inferior da classe onde estão os 25%**

2° quartil (Q2):

Queremos achar qual o grau de germinação de até 50% da amostra (Q2):

- Deve-se analisar a fai (frequência relativa acumulada) para identificar o valor onde podem estar estes 50%:
- Na terceira linha se encontra o 2° quartil, pois até 70% da amostra está entre 60 e 85, precisamos achar qual é o valor deste quartil:

**$85 - 80 = 5$  (lim sup – lim inf da classe)**

Então:

5 ---- 42% (fi)

X ---- (50% -28%) o que eu quero de fai menos anterior

Então: 5 ----- 42%

x ----- 22%

$X = 2,62$

**Q2 = 2,62 + 80 (lim inf da classe) = 82,62% = mediana** (50% da amostra se encontra com grau de germinação de até 82,62%)

3° quartil (Q3):

Queremos achar qual o grau de germinação de até 75% da amostra (Q3):

- Deve-se analisar a fai (frequência relativa acumulada) para identificar o valor onde podem estar estes 75%:
- Na quarta linha se encontra o 3° quartil, pois até 88% da amostra está entre 60 e 90, precisamos achar qual é o valor deste quartil:

**90 - 85 = 5 (lim sup – lim inf da classe)**

Então:

5 ---- 18% (fi)

X ---- (75% - 70%) o que eu quero de fai menos anterior

Então: 5 ----- 18%

x ----- 5%

$X = 1,39$

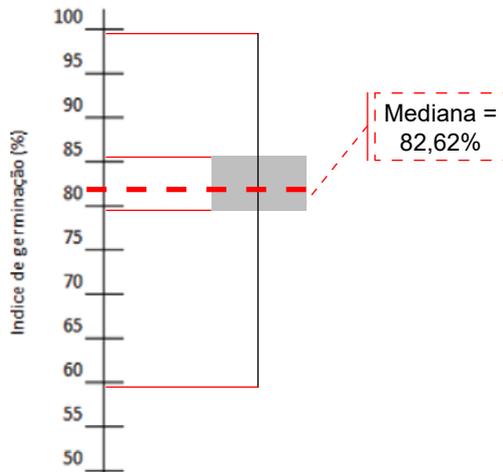
**Q3 = 1,39 + 85 (lim inf da classe) = 86,39%** (75% da amostra se encontra com grau de germinação de até 86,39%)

Construindo o gráfico box plot:

Para construir este diagrama é necessário:

- Valor mínimo observado = 60%
- 1° quartil = 79,25%
- Mediana = 2° quartil = 82,62%
- 3° quartil = 86,39%
- Valor máximo observado = 100%

## Germinação das Sementes



O quadrado corresponde às observações centrais 1º Q ao 3º Q (onde as germinações das sementes ficam entre 79,25% e 86,39%). Percebe-se que a germinação dentro do quadrado não está bem distribuída, pois a mediana não divide o quadrado ao meio.

c)

R.: De acordo com o que é observado na pesquisa, 50% das sementes têm índice de germinação inferior ou igual a 82,62%, enquanto 75% das sementes têm germinação inferior a 86,4%.

Na verdade, de acordo com a tabela, 70% das sementes têm índice de germinação menor do que 85%. Assim, a informação do fabricante não expressa exatamente o que é observado.

---

## TÓPICO 4

**1 No tópico anterior, você organizou os dados relativos ao tempo em dias necessário para a completa cicatrização de determinada cirurgia (exercício 1). Calcule para o mesmo exercício:**

- a) A amplitude dos dados.
- b) O desvio padrão amostral.

## MÉTODOS QUANTITATIVOS

- c) O coeficiente de assimetria.
- d) O coeficiente de curtose.

R.: Usar dados e informações da Questão 1 da página 102 para resolver essa questão.

1°) pegar o valor da média do tempo de cicatrização já calculado:

$$\text{MÉDIA } (\bar{x}) = \frac{(x_1 \cdot n_1) + (x_2 \cdot n_2) + \dots}{\sum n_i}$$

$x_1$  = quantidade de dias para cicatrizar

$n_1$  = frequência (quantidade de pacientes analisados)

$X_i$	$n_i$
5	1
6	4
7	7
8	4
9	5
10	3
11	1
TOTAL	25

$$\bar{x} = \frac{(5 \cdot 1) + (6 \cdot 4) + (7 \cdot 7) + (8 \cdot 4) + (9 \cdot 5) + (10 \cdot 3) + (11 \cdot 1)}{25}$$

$$\bar{x} = \frac{196}{25} = 7,84 \text{ dias é a média de tempo de cicatrização}$$

2°) para achar o desvio padrão amostral precisamos da variância amostral:

$$\text{Var}(x) =$$

$$\frac{(x_1 - \bar{X})^2 + (x_2 - \bar{X})^2 + (x_n - \bar{X})^2}{n - 1}$$

Em que:

$x_i$  = quantidade de dias para cicatrizar

$n$  = total frequência absoluta ( $n_i$ )

$\bar{x}$  = média de dias para cicatrização

## MÉTODOS QUANTITATIVOS

Então monta-se uma nova tabela contendo os valores de  $x_i - \bar{x}$  (quantidade de dias menos a média de dias para cicatrizar) e depois  $(x_i - \bar{x})^2$  para facilitar o uso na fórmula da variância.

Necessário para cálculo da média ( $n_i - x_i$ )

$X_i$	$n_i$	$n_i \cdot x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
5	1	5	-2,84	8,07
6	4	24	-1,84	3,39
7	7	49	-0,84	0,71
8	4	32	0,16	0,03
9	5	45	1,16	1,35
10	3	30	2,16	4,67
11	1	11	3,16	9,97
TOTAL	25	196		28,19

Achando a variância populacional:

$$\text{Var}(x) =$$

$$\frac{(x_1 - \bar{X})^2 + (x_2 - \bar{X})^2 + \dots + (x_n - \bar{X})^2}{n - 1}$$

$$\text{Var}(x) = \frac{8,07 + 3,39 + 0,71 + 0,03 + 1,35 + 4,67 + 9,97}{25 - 1} = 28,19/24 = \mathbf{1,17}$$

a) Amplitude dos dados ( $\Delta$ ): diferença entre o maior e o menor valor observado:

Maior valor: 11 dias

Menor valor: 1 dia

$$11 - 1 = 10 \quad (\Delta = \mathbf{10})$$

b) Desvio padrão amostral ( $S_x$ ) = raiz quadrada da variância amostral

$$S(x) = \sqrt{\text{var}(x)}$$

$$S(x) = \sqrt{1,17}$$

$$\mathbf{S(x) = 1,08}$$

c) Achar o coeficiente de assimetria (precisamos da mediana):

## MÉTODOS QUANTITATIVOS

MEDIANA (Md): é o valor da variável que ocupa a **posição** central de um conjunto de  $n$  dados ordenados (50% dos dados estão acima dela e 50% dos dados estão abaixo dela)

COMO NÃO SÃO DADOS AGRUPADOS (COM INTERVALO DE CLASSE), OLHAR O ROL:

- IDENTIFICAR O VALOR DO MEIO (QUANDO O N° DA AMOSTRA É ÍMPAR)
- CASO O NÚMERO DA AMOSTRA SEJA PARA PEGAR OS DOIS VALORES CENTRAIS, SOMAR E DIVIDIR POR 2.

5	6	6	6	6
7	7	7	7	7
7	7	8	8	8
8	9	9	9	9
9	10	10	10	11

Neste caso temos uma amostra de 25 (ímpar), no rol vamos achar o valor central, ou seja, o valor que divide a amostra em duas partes iguais (menos de 50% e mais de 50% do total da amostra).

**Md = 8**

Coefficiente de Assimetria (As):

$$As = \frac{3(\bar{X} - Md)}{s}$$

Desvio padrão amostral:

$$S(x) = \sqrt{\frac{(x_1 - \bar{X})^2 + (x_2 - \bar{X})^2 + \dots + (x_n - \bar{X})^2}{n - 1}}$$

$$S(x) = \sqrt{\frac{28,19}{25 - 1}}$$

$$S(x) = \sqrt{1,17}$$

$$S(x) = 1,08$$

## MÉTODOS QUANTITATIVOS

Em que:  $\bar{x}$  é a média = **7,84**

Md= mediana = **8**

S = desvio padrão amostral = **1,08**

**Quando:**

**“AS” menor/igual 0 e menor que 0,15 assimetria leve**

**“AS” maior/igual a 0,15 até 1 assimetria moderada**

**“AS” maior que 1 assimetria forte**

Então: Coeficiente de Assimetria (As):

$$As = \frac{3(\bar{X} - Md)}{s}$$

$$As = \frac{3 \cdot (7,84 - 8)}{1,08}$$

$$As = \frac{3 \cdot (-0,16)}{1,08}$$

$$As = -0,48/1,08 = -0,44$$

Assimetria leve, pois “As” é menor/igual zero e menor que 0,15.

d) Coeficiente de Curtose (C):

$$C = \frac{Q3 - Q1}{2(P90 - P10)}$$

Em que Q3 = terceiro quartis

Q1 = primeiro quartis

P90 = 90° percentis

P10 = 10° percentis

Para achar Q1, Q3, P10 e P90 precisamos do Fai (frequência relativa acumulada):

## MÉTODOS QUANTITATIVOS

xi	ni	fi	fai
5	1	0,040	0,040
6	4	0,160	0,200
7	7	0,280	0,480
8	4	0,160	0,640
9	5	0,200	0,840
10	3	0,120	0,960
11	1	0,040	1,000
TOTAL	25	1,000	

Q1 = separa os 25% primeiros dados do restante

Q3 = separa os 75% primeiros dados do restante

90° percentil = separa os 90% primeiros dados do restante

10° Percentil = separa os 10% primeiros dados do restante.

Para encontrar estes valores, identifica-se a classe onde se **encontram as devidas porcentagens (pelo valor da Fai), pegando o valor de xi da classe:**

Q1 = 7 (pois até 48% dos primeiros valores se encontram nesta classe, incluindo os 25% primeiros)

Q3 = 9 (pois até 84% dos primeiros valores se encontram nesta classe, incluindo os 75% primeiros)

P10 = 6 (pois até 20% dos primeiros valores se encontram nesta classe, incluindo os 10% primeiros)

P90 = 10 (pois até 96% dos primeiros valores se encontram nesta classe, incluindo os 90% primeiros)

$$C = \frac{Q3 - Q1}{2(P90 - P10)}$$

$$C = \frac{9 - 7}{2(10 - 6)}$$

$$C = \frac{2}{2 \cdot 4}$$

$$C = 2/8 = 0,25$$

**Neste caso temos uma curva leptocúrtica (C menor que 0,263)**

Quando C = 0,263 A CURVA É MESOCÚRTICA

Quando C < 0,263 A CURVA É LEPTOCÚRTICA

## MÉTODOS QUANTITATIVOS

Quando  $C > 0,263$  A CURVA É PLATICÚRTICA

**2 No tópico anterior, você calculou a média de acidentes nos últimos 12 meses em um entroncamento numa certa cidade.**

5 4 7 8 5 6 4 7 9 7 6 8

Encontre:

- O desvio padrão amostral.
- A distribuição é assimétrica? Justifique sua resposta.
- Indique que tipo de curtose ocorre nesta distribuição.

R.: Usar dados da Questão 2 (p. 102).

MÉDIA DE ACIDENTES/MÊS

$$\bar{x} = \frac{5+4+7+8+5+6+4+7+9+7+6+8}{12} = \frac{76}{12} = 6,333 \text{ acidentes/mês}$$

- Desvio padrão amostral: primeiro deve-se achar a variância amostral. Para isso precisamos montar a tabela:

N° acid. (xi)	Freq ni	fac	fi (ni/Σni)	fai	xi . ni	xi - $\bar{x}$	(xi - $\bar{x}$ ) <sup>2</sup>
4	2	2	0,167	0,167	8	-2,333	5,443
5	2	4	0,167	0,334	10	-1,333	1,777
6	2	6	0,167	0,501	12	-0,333	0,111
7	3	9	0,250	0,751	21	0,667	0,445
8	2	11	0,167	0,918	16	1,667	2,779
9	1	12	0,083	1,000	9	2,667	7,113
total	12				76		17,668

Rol:

4	4	5	5	6	6
7	7	7	8	8	9

Total  $n_i = 12$

Média  $(\bar{x}) = 6,333$

Variância amostral:

$\text{Var}(x) =$

$$\frac{(x_1 - \bar{X})^2 + (x_2 - \bar{X})^2 + \dots + (x_n - \bar{X})^2}{n - 1}$$

$$\text{Var}(x) = \frac{17,668}{12 - 1} = 1,606$$

**a) Desvio padrão amostral ( $S_x$ ) = raiz quadrada da variância amostral:**

$$S(x) = \sqrt{\text{var}(x)}$$

$$S(x) = \sqrt{1,606}$$

$$S(x) = 1,27$$

**b) A distribuição é assimétrica? Justifique sua resposta. (não está pedindo o coeficiente de assimetria):**

Para saber, precisamos achar a moda ( $M_o$ ) = valor observado com maior frequência.

Então  $M_o = 7$  (a quantidade de 7 acidentes se repete 3 vezes na observação)

Se:

**Média –  $M_o = 0$  é simétrica**

**Média –  $M_o < 0$  assimétrica à esquerda (negativa)**

**Média –  $M_o > 0$  assimétrica à direita (positiva)**

## MÉTODOS QUANTITATIVOS

Neste caso:

média = 6,333

Moda = 7

**Média – Mo = - 0,667 (menor que zero, portanto, a distribuição é assimétrica à esquerda ou negativa).**

**c) Indique que tipo de curtose ocorre nesta distribuição:**

$$C = \frac{Q3 - Q1}{2(P90 - P10)}$$

Em que Q3 = terceiro quartis

Q1 = primeiro quartis

P90 = 90° percentis

P10 = 10° percentis

Para achar Q1, Q3, P10 e P90 precisamos do Fai (frequência relativa acumulada):

N° acid. (xi)	Freq ni	fac	fi (ni/Σni)	fai	xi . ni	xi - $\bar{x}$	(xi - $\bar{x}$ ) <sup>2</sup>
4	2	2	0,167	0,167	8	-2,333	5,443
5	2	4	0,167	0,334	10	-1,333	1,777
6	2	6	0,167	0,501	12	-0,333	0,111
7	3	9	0,250	0,751	21	0,667	0,445
8	2	11	0,167	0,918	16	1,667	2,779
9	1	12	0,083	1,000	9	2,667	7,113
total	12				76		17,668

- Q1 = separa os 25% primeiros dados do restante
- Q3 = separa os 75% primeiros dados do restante
- 90° percentil = separa os 90% primeiros dados do restante
- 10° Percentil = separa os 10% primeiros dados do restante.

Neste caso:

Q1 = 5

Q3 = 7

P10 = 4

P90 = 8

## MÉTODOS QUANTITATIVOS

Para encontrar estes valores (Q1, Q3, P10 e P90), identifica-se a classe onde se **encontram as devidas porcentagens (pelo valor da Fai), pegando o valor de xi da classe:**

Q1 = 5 (pois até 33,4% dos primeiros valores se encontram nesta classe, incluindo os 25% primeiros)

Q3 = 7 (pois até 75,1% dos primeiros valores se encontram nesta classe, incluindo os 75% primeiros)

P10 = 4 (pois até 16,7% dos primeiros valores se encontram nesta classe, incluindo os 10% primeiros)

P90 = 8 (pois até 91,8% dos primeiros valores se encontram nesta classe, incluindo os 90% primeiros)

$$C = \frac{Q3 - Q1}{2(P90 - P10)}$$

$$C = \frac{7 - 5}{2(8 - 4)}$$

$$C = \frac{2}{8} = 0,25$$

**Neste caso temos uma curva leptocúrtica (C menor que 0,263).**

Quando  $C = 0,263$  A CURVA É MESOCÚRTICA

Quando  $C < 0,263$  A CURVA É LEPTOCÚRTICA

Quando  $C > 0,263$  A CURVA É PLATICÚRTICA

**3 Em uma clínica cardíaca, foram anotados os níveis de colesterol (em mg/100ml) para trinta pacientes, homens com idade entre 40 e 60 anos que foram à clínica fazer um check-up.**

Paciente	1	2	3	4	5	6	7	8	9	10
Colesterol	160	160	161	163	167	170	172	172	173	177

Paciente	11	12	13	14	15	16	17	18	19	20
Colesterol	178	181	181	182	185	186	194	197	199	203

Paciente	21	22	23	24	25	26	27	28	29	30
Colesterol	203	205	206	206	208	209	211	214	218	225

## MÉTODOS QUANTITATIVOS

No tópico anterior, você organizou estes dados em uma distribuição de frequências simples. Utilize esta tabela para calcular:

- O desvio padrão amostral.
- O coeficiente de assimetria (se houver assimetria).
- O coeficiente de curtose.

R.: Dados da Questão 3 (p. 102).

níveis de colesterol (x)	ni	fi = ni/Σni	fai
160	2	0,067	0,067
161	1	0,033	0,100
163	1	0,033	0,133
167	1	0,033	0,166
170	1	0,033	0,199
172	2	0,067	0,260
173	1	0,033	0,299
177	1	0,033	0,332
178	1	0,033	0,365
181	2	0,067	0,432
182	1	0,033	0,465
185	1	0,033	0,498
186	1	0,033	0,531
194	1	0,033	0,564
197	1	0,033	0,597
199	1	0,033	0,630
203	2	0,067	0,697
205	1	0,033	0,730
206	2	0,067	0,797
208	1	0,033	0,830
209	1	0,033	0,863
211	1	0,033	0,896
214	1	0,033	0,929
218	1	0,033	0,962
225	1	0,033	1,000
TOTAL =Σ	30	1	

## MÉTODOS QUANTITATIVOS

$$\text{MÉDIA } (\bar{x}) = \frac{(x_1 \cdot n_1) + (x_2 \cdot n_2) + \dots}{\sum n_i}$$

$$\bar{x} = \frac{(160 \cdot 2) + (161 \cdot 1) + (163 \cdot 1) + \dots}{30}$$

$$\bar{x} = \frac{5666}{30} = 188,9$$

MODA (MO) = valor ou atributo analisado que ocorre com maior frequência.  
MO = 160, 172, 181, 203, 206 (MULTIMODAL = mais de duas modas)

MEDIANA (Md): é o valor da variável que ocupa a **posição** central de um conjunto de *n* dados ordenados (50% dos dados estão acima dela e 50% dos dados estão abaixo dela)

OLHAR O ROL:

- IDENTIFICAR O VALOR DO MEIO (QUANDO O N° DA AMOSTRA É ÍMPAR)
- CASO O NÚMERO DA AMOSTRA SEJA PARA PEGAR OS DOIS VALORES CENTRAIS, SOMAR E DIVIDIR POR 2.

Neste caso temos uma amostra de 30 (par), no rol vamos achar os dois valores centrais, somar e dividir por 2:

$$\text{Md} = \frac{185 + 186}{2} = \frac{371}{2} = 185,5$$

Completar a tabela com os valores de  $(x - \bar{x})$  e  $(x - \bar{x})^2$

## MÉTODOS QUANTITATIVOS

níveis de colesterol (x)	ni	fi = ni/Σni	fai	x - $\bar{x}$	(x - $\bar{x}$ ) <sup>2</sup>
160	2	0,067	0,067	-28,900	835,210
161	1	0,033	0,100	-27,900	778,410
163	1	0,033	0,133	-25,900	670,810
167	1	0,033	0,166	-21,900	479,610
170	1	0,033	0,199	-18,900	357,210
172	2	0,067	0,260	-16,900	285,610
173	1	0,033	0,299	-15,900	252,810
177	1	0,033	0,332	-11,900	141,610
178	1	0,033	0,365	-10,900	118,810
181	2	0,067	0,432	-7,900	62,410
182	1	0,033	0,465	-6,900	47,610
185	1	0,033	0,498	-3,900	15,210
186	1	0,033	0,531	-2,900	8,410
194	1	0,033	0,564	5,100	26,010
197	1	0,033	0,597	8,100	65,610
199	1	0,033	0,630	10,100	102,010
203	2	0,067	0,697	14,100	198,810
205	1	0,033	0,730	16,100	259,210
206	2	0,067	0,797	17,100	292,410
208	1	0,033	0,830	19,100	364,810
209	1	0,033	0,863	20,100	404,010
211	1	0,033	0,896	22,100	488,410
214	1	0,033	0,929	25,100	630,010
218	1	0,033	0,962	29,100	846,810
225	1	0,033	1,000	36,100	1303,210
TOTAL =Σ	30	1		21,500	9035,050

## a) O desvio padrão amostral:

Para achar o desvio padrão amostral temos que achar a variância amostral:

$$\text{Var}(x) =$$

$$\frac{(x_1 - \bar{X})^2 + (x_2 - \bar{X})^2 + \dots + (x_n - \bar{X})^2}{n - 1}$$

$$\text{Var}(x) = \frac{9035,050}{30 - 1} =$$

$$\text{Var}(x) = \frac{9035,050}{29} = 821,368$$

a) Desvio padrão amostral ( $S_x$ ) = raiz quadrada da variância amostral

$$S(x) = \sqrt{\text{var}(x)}$$

$$S(x) = \sqrt{821,368}$$

$$S(x) = 28,66$$

## b) Coeficiente de assimetria (se houver):

$$A_s = \frac{3(\bar{X} - Md)}{s}$$

Em que:  $\bar{x}$  é a média,

Md= mediana

S = desvio padrão amostral

$$A_s = \frac{3(188,9 - 185,5)}{28,66}$$

$$A_s = 10,2/28,66 = 0,36$$

## c) Coeficiente de curtose:

$$C = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})}$$

## MÉTODOS QUANTITATIVOS

Em que: Q3 = terceiro quartis

Q1 = primeiro quartis

P90 = 90° percentis

P10 = 10° percentis

Para achar Q1, Q3, P10 e P90 precisamos do Fai (frequência relativa acumulada):

Neste caso:

Q1 = 172

Q3 = 206

P10 = 161

P90 = 214

$$C = \frac{206 - 172}{2(214 - 161)}$$

$$C = 34/106 = 0,321$$

- Q1 = separa os 25% primeiros dados do restante
- Q3 = separa os 75% primeiros dados do restante
- 90° percentil = separa os 90% primeiros dados do restante
- 10° Percentil = separa os 10% primeiros dados do restante.

Quando  $C = 0,263$  A CURVA É MESOCÚRTICA

Quando  $C < 0,263$  A CURVA É LEPTOCÚRTICA

Quando  $C > 0,263$  A CURVA É PLATICÚRTICA

**4 Um banco instalou um caixa eletrônico em um posto de combustível e está observando o número de usuários que vêm utilizando o serviço. Diariamente, número de clientes que utilizou o serviço nos últimos 32 dias foi:**

15 17 16 15 17 14 17 16 16 17 15  
18 14 17 15 14 15 14 15 16 17 18  
18 17 15 16 14 18 18 16 15 14.

**Calcule:**

**a) A variância amostral para esta distribuição.**

**b) Qual o tipo de curtose que ocorre?**

R.:

Rol:

## MÉTODOS QUANTITATIVOS

14	14	14	14	14	14
15	15	15	15	15	15
15	15	16	16	16	16
16	16	17	17	17	17
17	17	17	18	18	18
18	18				

N° usuários	ni	fi = ni/Σni	fai
14	6	0,1875	0,1875
15	8	0,2500	0,4375
16	6	0,1875	0,6250
17	7	0,2188	0,8438
18	5	0,1562	1,0000
TOTAL (Σ)	32	1,0000	

N° usuários (X)	ni	fac	fi = ni/Σni	fai	ni . X	x - $\bar{x}$	(x - $\bar{x}$ ) <sup>2</sup>
14	6	6	0,1875	0,1875	84	-1,9100	3,6481
15	8	14	0,25	0,4375	120	-0,9100	0,8281
16	6	20	0,1875	0,625	96	0,0900	0,0081
17	7	27	0,2188	0,8438	119	1,0900	1,1881
18	5	32	0,1562	1	90	2,0900	4,3681
TOTAL (Σ)	32		1		509		10,0405

MÉDIA:

$$\bar{x} = \frac{(x1 \cdot n1) + (x2 \cdot n2) + \dots}{\Sigma ni}$$

$$\bar{x} = 509/32 = 15,91$$

a) Variância amostral = Var (x)

Var(x) =

$$\frac{(x_1 - \bar{X})^2 + (x_2 - \bar{X})^2 + \dots + (x_n - \bar{X})^2}{n - 1}$$

$$\text{Var}(x) = \frac{10,040}{32 - 1} =$$

$$\text{Var}(x) = 10,040/31 = \mathbf{0,32}$$

**b) Qual o tipo de curtose?**

$$C = \frac{Q3 - Q1}{2(P90 - P10)}$$

Em que Q3 = terceiro quartis

Q1 = primeiro quartis

P90 = 90° percentis

P10 = 10° percentis

Para achar Q1, Q3, P10 e P90 precisamos do Fai (frequência relativa acumulada):

Neste caso:

Q1 = 15

Q3 = 17

P10 = 14

P90 = 18

$$C = \frac{17 - 15}{2(18 - 14)}$$

$$C = \mathbf{2/8 = 0,25}$$

0,25 < 0,263 a curva é leptocúrtica

- Q1 = separa os 25% primeiros dados do restante
- Q3 = separa os 75% primeiros dados do restante
- 90° percentil = separa os 90% primeiros dados do restante
- 10° Percentil = separa os 10% primeiros dados do restante.

## MÉTODOS QUANTITATIVOS

Quando  $C = 0,263$  A CURVA É MESOCÚRTICA

Quando  $C < 0,263$  A CURVA É LEPTOCÚRTICA

Quando  $C > 0,263$  A CURVA É PLATICÚRTICA

**5 Um posto de saúde de certo bairro mantém um arquivo com o número de pacientes que procuram o consultório odontológico diariamente. Os dados são obtidos no último mês foram os seguintes:**

3 4 3 4 5 1 6 3 4 5 3 4  
 3 3 4 3 5 5 5 5 6 11 10 2  
 1 2 3 1 5 2.

**Calcule a variância amostral para esta distribuição. Há simetria? Por quê?**

R.:

Rol:

1	1	1	2	2	2
3	3	3	3	3	3
3	3	4	4	4	4
4	5	5	5	5	5
5	5	6	6	10	11

usuários (x)	ni	fac	fi = ni/Σni	fai	x. n	x - $\bar{x}$	(x - $\bar{x}$ ) <sup>2</sup>
1	3	3	0,100	0,100	3	-3,03	9,1809
2	3	6	0,100	0,200	6	-2,03	4,1209
3	8	14	0,267	0,467	24	-1,03	1,0609
4	5	19	0,167	0,633	20	-0,03	0,0009
5	7	26	0,233	0,867	35	0,97	0,9409
6	2	28	0,067	0,933	12	1,97	3,8809
10	1	29	0,033	0,967	10	5,97	35,6409
11	1	30	0,033	1,000	11	6,97	48,5809
total	30				121		103,4072

## MÉTODOS QUANTITATIVOS

$$\text{Média} = 121/30 = 4,03$$

Variância amostral:

$$\text{Var}(x) =$$

$$\frac{(x_1 - \bar{X})^2 + (x_2 - \bar{X})^2 + \dots + (x_n - \bar{X})^2}{n - 1}$$

$$\text{Var}(x) = \frac{103.4072}{30 - 1} =$$

$$\text{Var}(x) = \frac{103.4072}{29} = 3,56$$

$$\text{Desvio padrão Amostral} = \sqrt{3,56} = 1,89$$

Coeficiente de assimetria:

$$As = \frac{3(\bar{X} - Md)}{s}$$

Em que:  $\bar{x}$  é a média,

Md= mediana

S = desvio padrão amostral

MEDIANA:

OLHAR O ROL:

- IDENTIFICAR O VALOR DO MEIO (QUANDO O N° DA AMOSTRA É ÍMPAR)
- CASO O NÚMERO DA AMOSTRA SEJA PARA PEGAR OS DOIS VALORES CENTRAIS, SOMAR E DIVIDIR POR 2.

Neste caso temos uma amostra de 30 (par), no rol vamos achar os dois valores centrais, somar e dividir por 2:

$$\text{Md} = \frac{4 + 4}{2} = 4$$

$$As = \frac{3(4,03 - 4)}{1,89}$$

$$As = 0,048$$

Assimetria leve

- Quando:
- “AS” menor/igual 0 e menor que 0,15 assimetria leve
- “AS” maior/igual a 0,15 até 1 assimetria moderada
- “AS” maior que 1 assimetria forte

**6 (MAGALHÃES, 2010) O tempo de utilização de caixas eletrônicos depende de cada usuário e das operações efetuadas. Foram coletadas 26 medidas desse tempo (minutos):**

1,1	1,2	1,7	1,5	0,9	1,3	1,4	1,6	1,7	1,6	1,0	0,8	1,5
1,3	1,7	1,6	1,4	1,2	1,2	1,0	0,9	1,8	1,7	1,5	1,3	1,5

Utilize as informações da amostra e calcule:

- O desvio padrão amostral.
- O coeficiente de assimetria (se houver assimetria).
- O coeficiente de curtose.

R.:

Tempo de uso (minutos) = x	ni	fi = ni/Σni	fai	x. n	x - $\bar{x}$	(x - $\bar{x}$ ) <sup>2</sup>
0,8	1	0,038	0,038	0,80	-0,560	0,3136
0,9	2	0,077	0,115	1,80	-0,460	0,2116
1	2	0,077	0,192	2,00	-0,360	0,1296
1,1	1	0,038	0,231	1,10	-0,260	0,0676
1,2	3	0,115	0,346	3,60	-0,160	0,0256
1,3	3	0,115	0,462	3,90	-0,060	0,0036
1,4	2	0,077	0,538	2,80	0,040	0,0016
1,5	4	0,154	0,692	6,00	0,140	0,0196
1,6	3	0,115	0,808	4,80	0,240	0,0576
1,7	4	0,154	0,962	6,80	0,340	0,1156
1,8	1	0,038	1	1,80	0,440	0,1936
TOTAL (Σ)	26	1		35,40		1,1396

## MÉTODOS QUANTITATIVOS

$$\bar{x} = \frac{(x_1 \cdot n_1) + (x_2 \cdot n_2) \dots}{\sum n_i}$$

$$\text{Média } (\bar{x}) = 35,40/25 = 1,36$$

**MEDIANA (Md):** é o valor da variável que ocupa a posição central de um conjunto de  $n$  dados ordenados (50% dos dados estão acima dela e 50% dos dados estão abaixo dela).

OLHAR O ROL:

- IDENTIFICAR O VALOR DO MEIO (QUANDO O N° DA AMOSTRA É ÍMPAR)
- CASO O NÚMERO DA AMOSTRA SEJA PARA PEGAR OS DOIS VALORES CENTRAIS, SOMAR E DIVIDIR POR 2.

Neste caso temos uma amostra de 26 (par), no rol vamos achar os dois valores centrais, somar e dividir por 2:

$$\text{Md} = \frac{1,4 + 1,4}{2} = \frac{2,8}{2} = 1,4$$

Rol:

0,8	0,9	0,9	1,0	1,0	1,1
1,2	1,2	1,2	1,3	1,3	1,3
1,4	1,4	1,5	1,5	1,5	1,5
1,6	1,6	1,6	1,7	1,7	1,7
1,7	1,8				

**a) Desvio padrão amostral:** precisamos calcular a variância amostral:

Variância amostral:

$$\text{Var}(x) =$$

$$\frac{(x_1 - \bar{X})^2 + (x_2 - \bar{X})^2 + \dots + (x_n - \bar{X})^2}{n - 1}$$

$$\text{Var}(x) = \frac{1.1396}{26-1} =$$

$$\text{Var}(x) = \frac{1.1396}{25} = 0,046$$

$$\text{Desvio padrão Amostral} = \sqrt{0,046} = \mathbf{0,214}$$

### b) Coeficiente de assimetria:

$$As = \frac{3(\bar{X} - Md)}{s}$$

Em que:  $\bar{x}$  é a média,

Md= mediana

S = desvio padrão amostral

$$As = \frac{3(1,36 - 1,4)}{0,214}$$

$$As = - 2,5$$

**Assimetria leve**

- Quando:
- “AS” menor/igual 0 e menor que 0,15 assimetria leve
- “AS” maior/igual a 0,15 até 1 assimetria moderada
- “AS” maior que 1 assimetria forte

### c) Curtose:

$$C = \frac{Q3 - Q1}{2(P90 - P10)}$$

Em que Q3 = terceiro quartis

Q1 = primeiro quartis

P90 = 90° percentis

P10 = 10° percentis

Para achar Q1, Q3, P10 e P90 precisamos do Fai (frequência relativa acumulada):

Neste caso:

$$Q1 = 1,2$$

$$Q3 = 1,6$$

$$P10 = 0,9$$

$$P90 = 1,7$$

$$C = \frac{1,6 - 1,2}{2(1,7 - 0,9)}$$

$$C = 0,04/1,6 = \mathbf{0,025}$$

**0,025 < 0,263 a curva é leptocúrtica**

- Q1 = separa os 25% primeiros dados do restante
- Q3 = separa os 75% primeiros dados do restante
- 90° percentil = separa os 90% primeiros dados do restante
- 10° Percentil = separa os 10% primeiros dados do restante.

Quando  $C = 0,263$  A CURVA É MESOCÚRTICA

Quando  $C < 0,263$  A CURVA É LEPTOCÚRTICA

Quando  $C > 0,263$  A CURVA É PLATICÚRTICA

## UNIDADE 3

### TÓPICO 1

- 1 Uma empresa deseja investigar quantos dos seus funcionários são fumantes e qual o seu perfil. Neste sentido, aplicou um questionário para seus 24 funcionários. A seguir, você pode ver o resultado da pesquisa:

## MÉTODOS QUANTITATIVOS

Indivíduo	Sexo	Hábito de fumar	Idade	Escolaridade	Cargo	Salário (em R\$)
1	M	Sim	18	Fundamental	Ajudantes	900,00
2	M	Não	17	Médio	Ajudantes	1056,00
3	M	Não	18	Médio	Ajudantes	1356,00
4	M	Não	19	Fundamental	Ajudantes	1500,00
5	F	Não	32	Superior	Chefia	10500,00
6	F	Não	35	Superior	Chefia	11400,00
7	M	Sim	15	Fundamental	Ajudantes	1056,00
8	F	Sim	17	Fundamental	Ajudantes	1056,00
9	M	Não	19	Médio	Escriturários	2430,00
10	M	Não	20	Médio	Escriturários	2550,00
11	F	Não	22	Médio	Escriturários	2940,00
12	M	Sim	45	Superior	Chefia	13500,00
13	M	Sim	18	Fundamental	Ajudantes	1110,00
14	M	Não	17	Médio	Ajudantes	1056,00
15	M	Não	18	Médio	Ajudantes	1350,00
16	M	Não	19	Fundamental	Ajudantes	1500,00
17	F	Não	36	Superior	Chefia	10500,00
18	F	Não	35	Superior	Chefia	11400,00
19	M	Sim	15	Fundamental	Ajudantes	1056,00
20	F	Sim	17	Fundamental	Ajudantes	1056,00
21	M	Não	19	Médio	Escriturários	2430,00
22	M	Não	20	Médio	Escriturários	2430,00
23	F	Não	22	Médio	Escriturários	2940,00
24	M	Sim	40	Superior	Chefia	13500,00

Com base nos dados da tabela, monte as seguintes distribuições conjuntas e conclua se há dependência entre as variáveis:

**a) Hábito de fumar em relação ao gênero.**

R.:

## MÉTODOS QUANTITATIVOS

(Fuma, Gênero)	ni
(Sim, Feminino)	2
(Sim, Masculino)	6
(Não, Feminino)	6
(Não, Masculino)	10
TOTAL	24

### PERFIL DOS FUNCIONÁRIOS FUMANTES DA EMPRESA

Hábito de fumar	Gênero		TOTAL
	Feminino	Masculino	
Sim	2	6	8
Não	6	10	16
TOTAL	8	16	24

### b) Idade em relação ao hábito de fumar.

R.:

(Idade, Fuma)	ni
(15,sim)	2
(15,não)	0
(17,sim)	2
(17,não)	2
(18,sim)	2
(18,não)	2
(19,sim)	0
(19,não)	4
(20,sim)	0
(20,não)	2
(22,sim)	0
(22,não)	2
(32,sim)	0
(32,não)	1
(35,sim)	0
(35,não)	2
(36,sim)	0
(36,não)	1
(40,sim)	1
(40,não)	0
(45,sim)	1
(45,não)	0
TOTAL	24

## MÉTODOS QUANTITATIVOS

PERFIL DOS FUNCIONÁRIOS FUMANTES DA EMPRESA

Idade	Hábito de fumar		TOTAL
	Sim	Não	
15	2	0	2
17	2	2	4
18	2	2	4
19	0	4	4
20	0	2	2
22	0	2	2
32	0	1	1
35	0	2	2
36	0	1	1
40	1	0	1
45	1	0	1
TOTAL	8	16	24

### c) Idade em relação ao gênero.

R.:

(Idade, Gênero)	ni
(15, F)	0
(15, M)	2
(17, F)	2
(17, M)	2
(18, F)	0
(18, M)	4
(19, F)	0
(19, M)	4
(20, F)	0
(20, M)	2
(22, F)	2
(22, M)	0
(32, F)	1
(32, M)	0
(35, F)	2
(35, M)	0
(36, F)	1
(36, M)	0
(40, F)	0
(40, M)	1
(45, F)	0
(45, M)	1
TOTAL	24

## MÉTODOS QUANTITATIVOS

### PERFIL DOS FUNCIONÁRIOS FUMANTES DA EMPRESA

Idade	Gênero		TOTAL
	Feminino	Masculino	
15	0	2	2
17	2	2	4
18	0	4	4
19	0	4	4
20	2	0	2
22	0	2	2
32	1	0	1
35	2	0	2
36	1	0	1
40	0	1	1
45	0	1	1
TOTAL	8	16	24

#### d) Cargo em relação ao hábito de fumar.

R.:	(Cargo, Fuma)	ni
	(Ajudante, Sim)	6
	(Ajudante, não)	6
	(Escriturário, sim)	0
	(Escriturário, não)	6
	(Chefia, sim)	2
	(Chefia, não)	4
	TOTAL	24

### PERFIL DOS FUNCIONÁRIOS FUMANTES DA EMPRESA

Cargo	Hábito de fumar		TOTAL
	Sim	Não	
Ajudante	6	6	12
Escriturário	0	6	6
Chefia	2	4	6
TOTAL	8	16	24

#### e) Sexo em relação ao cargo.

## MÉTODOS QUANTITATIVOS

R.:

PERFIL DOS FUNCIONÁRIOS FUMANTES DA EMPRESA			
Cargo	Gênero		TOTAL
	Feminino	Masculino	
Ajudante	2	10	12
Escriturário	2	4	6
Chefia	4	2	6
TOTAL	8	16	24

### f) Faixa salarial em relação ao hábito de fumar.

R.:

Idade	Hábito de Fumar		TOTAL
	Sim	Não	
900	1	0	1
1056	4	2	6
1110	1	0	1
1350	0	1	1
1356	0	1	1
1500	0	2	2
2430	0	3	3
2550	0	1	1
2940	0	2	2
10500	0	2	2
11400	0	2	2
13500	2	0	2
TOTAL	8	16	24

FONTE: Adaptado de: Farhat e Elian (2006).

- 2** Sabe-se que alguns supermercados colocam uma mercadoria em oferta e aumentam levemente o preço das outras que podem ser compradas como complemento. Assim, quem compra macarrão pode necessitar também do molho do tomate. Uma pesquisa teve o objetivo de observar se havia correlação entre o preço do macarrão em oferta e do molho de tomate em dez supermercados. Os preços praticados estão na tabela a seguir:

## MÉTODOS QUANTITATIVOS

Preço do Macarrão	Preço do molho de tomate
2,00	1,20
2,30	1,40
2,50	1,75
2,60	2,00
2,80	2,20
3,00	2,40
3,20	2,70
3,25	3,00
3,30	3,20
3,50	4,00

As duas séries de preços estão correlacionadas? Justifique.

R.:

X: preço do macarrão em oferta

Y: preço do molho de tomate

I	$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i \cdot y_i$
1	2,00	1,20	4,00	1,44	2,40
2	2,30	1,40	5,29	1,96	3,22
3	2,50	1,75	6,25	3,06	4,38
4	2,60	2,00	6,76	4,00	5,20
5	2,80	2,20	7,84	4,84	6,16
6	3,00	2,40	9,00	5,76	7,20
7	3,20	2,70	10,24	7,29	8,64
8	3,25	3,00	10,56	9,00	9,75
9	3,30	3,20	10,89	10,24	10,56
10	3,50	4,00	12,25	16,00	14,00
Total	28,45	23,85	83,08	63,59	71,51

$$r = \frac{n \cdot \left( \sum_{i=1}^n x_i \cdot y_i \right) - \left( \sum_{i=1}^n x_i \right) \cdot \left( \sum_{i=1}^n y_i \right)}{\sqrt{\left[ n \cdot \left( \sum_{i=1}^n x_i^2 \right) - \left( \sum_{i=1}^n x_i \right)^2 \right] \cdot \left[ n \cdot \left( \sum_{i=1}^n y_i^2 \right) - \left( \sum_{i=1}^n y_i \right)^2 \right]}}$$

$$= \frac{10 \cdot 71,51 - 28,45 \cdot 23,85}{\sqrt{\left[ 10 \cdot 83,08 - (28,45)^2 \right] \cdot \left[ 10 \cdot 63,59 - (23,85)^2 \right]}} = 0,963$$

## MÉTODOS QUANTITATIVOS

As duas séries estão fortemente correlacionadas positivamente, pois o coeficiente linear de ambas é muito próximo de 1.

FONTE: Adaptado de Novaes e Coutinho (2009).

**3 A tabela a seguir apresenta os dados referentes a áreas de terrenos em metros quadrados com seus respectivos preços de venda em mil reais de uma determinada região de São Paulo.**

Área dos terrenos (m <sup>2</sup> )	Preço de venda (mil reais)
40	42
60	58
60	58
60	70
80	80
100	90
115	100
130	102
138	130
150	130
160	140

**Há correlação entre os dados? Justifique.**

R.:

X: áreas de terrenos em metros quadrados

Y: preços de venda em mil reais

i	$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i \cdot y_i$
1	42	40	1764	1600	1680
2	58	60	3364	3600	3480
3	58	60	3364	3600	3480
4	70	60	4900	3600	4200
5	80	80	6400	6400	6400
6	90	100	8100	10000	9000
7	100	115	10000	13225	11500
8	102	130	10404	16900	13260
9	130	138	16900	19044	17940
10	130	150	16900	22500	19500
11	140	160	19600	25600	22400
Total	1000	1093	101696	126069	112840

$$r = \frac{n \cdot \left( \sum_{i=1}^n x_i \cdot y_i \right) - \left( \sum_{i=1}^n x_i \right) \cdot \left( \sum_{i=1}^n y_i \right)}{\sqrt{\left[ n \cdot \left( \sum_{i=1}^n x_i^2 \right) - \left( \sum_{i=1}^n x_i \right)^2 \right] \cdot \left[ n \cdot \left( \sum_{i=1}^n y_i^2 \right) - \left( \sum_{i=1}^n y_i \right)^2 \right]}}$$

$$= \frac{11 \cdot 112840 - 1000 \cdot 1093}{\sqrt{\left[ 11 \cdot 101696 - (1000)^2 \right] \cdot \left[ 11 \cdot 126069 - (1093)^2 \right]}} = 0,981$$

Há forte correlação positiva entre as variáveis, pois o coeficiente de correlação linear é muito próximo de 1.

FONTE: Adaptado de Farhat e Elian (2006).

## TÓPICO 2

**1 Uma agência de turismo, especializada em oferecer passeios opcionais para turistas que visitam determinada região, está estudando a variação na adesão a determinado pacote quando são acrescentados ou tirados percursos do preço cobrado, obtendo as seguintes informações:**

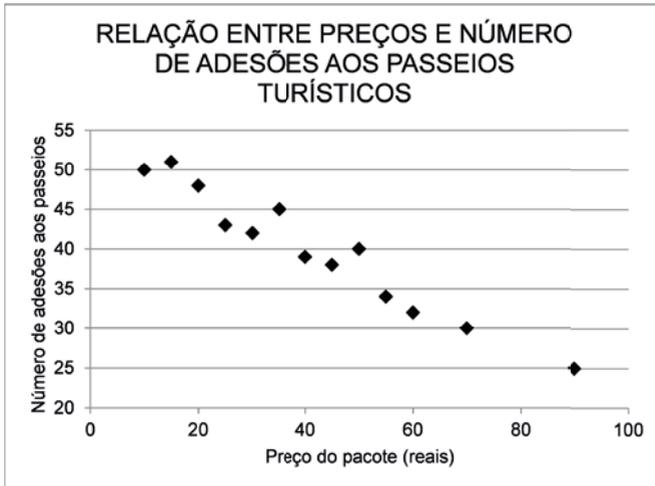
Preço (\$)	Número de adesões
10	50
15	51
20	48
25	43
30	42
35	45
40	39
45	38
50	40
55	34
60	32
70	30
90	25

X: Preço (variável independente)

Y: Número de adesões (variável dependente)

a) Faça o gráfico de dispersão.

R.:



b) Há correlação linear entre os dados? Justifique.

R.: Aparentemente, sim. Pelo gráfico, quanto maior o preço do pacote, menor o número de adesões. Calculando a correlação entre as variáveis, temos:

i	$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i \cdot y_i$
1	10	50	100	2500	500
2	15	51	225	2601	765
3	20	48	400	2304	960
4	25	43	625	1849	1075
5	30	42	900	1764	1260
6	35	45	1225	2025	1575
7	40	39	1600	1521	1560
8	45	38	2025	1444	1710
9	50	40	2500	1600	2000
10	55	34	3025	1156	1870
11	60	32	3600	1024	1920
12	70	30	4900	900	2100
13	90	25	8100	625	2250
TOTAL	545	517	29225	21313	19545

$$r = -0,972$$

De fato, há forte correlação negativa.

- c) Encontre a função matemática que explique a dependência entre o número de adesões e o preço do passeio opcional.

R.:

$$\beta = \frac{n \cdot \sum X \cdot Y - \sum X \cdot \sum Y}{n \cdot \sum X^2 - (\sum X)^2} = \frac{13 \cdot 19545 - 545 \cdot 517}{13 \cdot 29225 - (545)^2} = -0,3339$$

$$\alpha = \bar{Y} - \hat{\beta} \cdot \bar{X} = \frac{545}{13} - (-0,339) \cdot \frac{517}{13} = 53,767$$

A equação da reta que melhor aproxima os dados minimizando os quadrados dos resíduos é dada por:

$$\hat{Y} = 53,767 - 0,339 \cdot X$$

- d) Estime o número de pessoas que farão o passeio opcional se o valor cobrado for de 80 reais.

$$R.: \hat{Y} = 53,767 - 0,339 \cdot 80 = 27,06 \rightarrow 27 \text{ pessoas.}$$

- e) Encontre a proporção com que a variação no número de ações é explicada linearmente pelo preço da mesma.

$$R.: R^2 = r^2 = (-0,972)^2 = 0,945$$

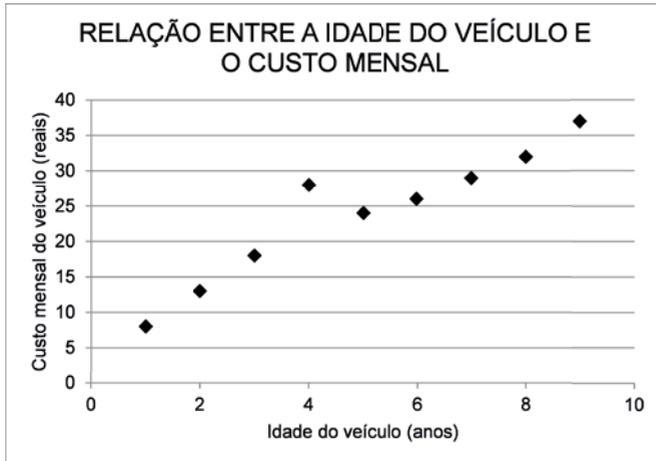
FONTE: Adaptado de Novaes e Coutinho (2009)

- 2 O custo mensal de manutenção de determinado tipo de automóvel (excluindo-se combustível e troca de óleo) está sendo analisado em função da idade do veículo. Nove automóveis fabricados em diferentes anos tiveram o custo averiguado e os dados obtidos foram os seguintes:

Idade do veículo (anos)	Custo mensal (reais)
1	8
2	13
3	18
4	28
5	24
6	26
7	29
8	32
9	37

a) Trace o gráfico de dispersão.

R.:



b) Calcule o coeficiente de determinação.

R.:

i	$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i \cdot y_i$
1	1	8	1	64	8
2	2	13	4	169	26
3	3	18	9	324	54
4	4	28	16	784	112
5	5	24	25	576	120
6	6	26	36	676	156
7	7	29	49	841	203
8	8	32	64	1024	256
9	9	37	81	1369	333
TOTAL	45	215	285	5827	1268

$$r = 0,948 \longrightarrow R^2 = r^2 = (0,948)^2 = 0,899$$

c) Faça a regressão linear e encontre a equação da reta melhor ajustada.

R.:

$$\hat{\beta} = \frac{n \cdot \sum X \cdot Y - \sum X \cdot \sum Y}{n \cdot \sum X^2 - (\sum X)^2} = \frac{9 \cdot 1268 - 45 \cdot 215}{9 \cdot 285 - (45)^2} = 3,217$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \cdot \bar{X} = \frac{215}{9} - 3,217 \cdot \frac{45}{9} = 7,806$$

A equação da reta que melhor aproxima os dados minimizando os quadrados dos resíduos é dada por:

$$\hat{Y} = 7,806 + 3,217 \cdot X$$

**d) Com base no modelo de regressão linear, qual é o custo mensal de um carro com 12 anos de uso?**

R.:  $\hat{Y} = 7,806 + 3,217 \cdot 12 = 46,41$

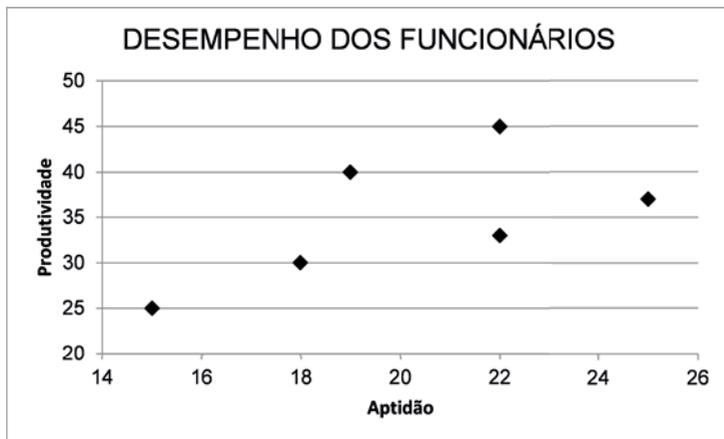
FONTE: Adaptado de Magalhães e Lima (2010).

**3 Uma indústria submete seus novos operários a um teste de aptidão (X) e três meses depois mede a produtividade destes operários (Y). Os resultados estão na tabela a seguir:**

Operário	Aptidão (X)	Produtividade (Y)
A	22	45
B	25	37
C	15	25
D	19	40
E	22	33
F	18	30

**a) Faça o diagrama de dispersão e calcule o coeficiente de correlação.**

R.:



b) Encontre a equação da reta de regressão.

R.:  $\hat{Y} = 9,003 + 1,289 \cdot X$

c) Para um indivíduo cujo resultado no teste de aptidão foi 20, qual é a produtividade esperada?

R.:  $\hat{Y} = 9,003 + 1,289 \cdot 20 = 34,785$

d) Para um indivíduo que obteve 28 no teste de produtividade, qual é o resultado no teste de aptidão?

R.:  $28 = 9,003 + 1,289 \cdot X \rightarrow X = 14,74$

FONTE: Adaptado de Magalhães e Lima (2010).

4 Certa empresa, estudando a variação da demanda do seu produto em relação à variação de preço de venda (em unidades monetárias), obteve os seguintes dados:

Preço (u.m.)	Demanda
38	350
42	325
50	297
56	270
59	256
63	246
70	238
80	223
95	215
110	208

## MÉTODOS QUANTITATIVOS

X: preço de venda (em unidades monetárias)

Y: demanda do seu produto

I	$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i \cdot y_i$
1	38	350	1444	122500	13300
2	42	325	1764	105625	13650
3	50	297	2500	88209	14850
4	56	270	3136	72900	15120
5	59	256	3481	65536	15104
6	63	246	3969	60516	15498
7	70	238	4900	56644	16660
8	80	223	6400	49729	17840
9	95	215	9025	46225	20425
10	110	208	12100	43264	22880
TOTAL	663	2628	48719	711148	165327

**a) Determine os coeficientes de correlação e de determinação entre as variáveis.**

R.:

$$r = -0,902$$

$$R^2 = 0,814$$

**b) Encontre a equação da reta ajustada.**

$$R.: \hat{Y} = 386,84 - 1,871 \cdot X$$

**c) Se o preço de venda for 75 u.m., qual é a demanda estimada?**

$$R.: \hat{Y} = 386,84 - 1,871 \cdot 75 = 246,52$$

**d) Se o preço de venda for de 110 u.m., qual é a demanda estimada?**

$$R.: \hat{Y} = 386,84 - 1,871 \cdot 110 = 181,04$$

**e) Qual é o valor máximo de preço possível para que haja demanda?**

R.: Para que haja demanda, Y precisa ser maior ou igual a zero. Supondo Y

$$= 0, \text{ temos } 0 = 386,84 - 1,871 \cdot X \rightarrow X = 206,77$$

Qualquer preço inferior a este valor implicará em haver demanda.

FONTE: Adaptado de Crespo (2009).

## TÓPICO 3

- 1 Você recebeu informações sobre a porcentagem de ações no mercado, preços como porcentagem dos preços médios do competidor, e gastos com propaganda como porcentagem dos preços médios do competidor, para uma linha específica de comida para gatos, nos últimos oito anos. Os dados obtidos seguem a seguir:

Período	Ações no Mercado (%)	Preço Relativo	Propaganda relativa
1	30	89	110
2	31	85	115
3	30,5	86,5	120
4	29	92	114
5	29,5	90	112
6	28,5	102	115
7	28	115	116
8	26	125	114

- a) Encontre a regressão múltipla que melhor estima o comportamento das ações no mercado em relação às outras duas variáveis.

R.:

i	$y_i$	$x_{1i}$	$x_{2i}$	$x_{1i}^2$	$x_{2i}^2$	$x_{1i} \cdot y_i$	$x_{2i} \cdot y_i$	$x_{1i} \cdot x_{2i}$
1	30,0	89,0	110,0	7921,0	12100,0	2670,0	3300,0	9790,0
2	31,0	85,0	115,0	7225,0	13225,0	2635,0	3565,0	9775,0
3	30,5	86,5	120,0	7482,3	14400,0	2638,3	3660,0	10380,0
4	29,0	92,0	114,0	8464,0	12996,0	2668,0	3306,0	10488,0
5	29,5	90,0	112,0	8100,0	12544,0	2655,0	3304,0	10080,0
6	28,5	102,0	115,0	10404,0	13225,0	2907,0	3277,5	11730,0
7	28,0	115,0	116,0	13225,0	13456,0	3220,0	3248,0	13340,0
8	26,0	125,0	114,0	15625,0	12996,0	3250,0	2964,0	14250,0
TOTAL	232,5	784,5	916,0	78446,3	104942,0	22643,3	26624,5	89833,0

$$\left\{ \begin{aligned} S_{11} &= \sum X_1^2 - \frac{(\sum X_1)^2}{n} = 78466,3 - \frac{(784,5)^2}{8} = 1516,2 \\ S_{22} &= \sum X_2^2 - \frac{(\sum X_2)^2}{n} = 104942,0 - \frac{(916,0)^2}{8} = 60,0 \\ S_{12} &= \sum X_1 \cdot X_2 - \frac{(\sum X_1) \cdot (\sum X_2)}{n} = 89833,0 - \frac{784,5 \cdot 916,0}{8} = 7,8 \\ S_{Y1} &= \sum Y \cdot X_1 - \frac{(\sum Y) \cdot (\sum X_1)}{n} = 22643,3 - \frac{(232,5) \cdot (784,5)}{8} = -156,3 \\ S_{Y2} &= \sum Y \cdot X_2 - \frac{(\sum Y) \cdot (\sum X_2)}{n} = 26624,5 - \frac{(232,5) \cdot (916,0)}{8} = 3,3 \end{aligned} \right.$$

$$\left\{ \begin{aligned} S_{Y1} &= \beta_1 \cdot S_{11} + \beta_2 \cdot S_{12} \\ S_{Y2} &= \beta_1 \cdot S_{12} + \beta_2 \cdot S_{22} \end{aligned} \right. \longrightarrow \left\{ \begin{aligned} -156,3 &= 1516,2 \cdot \beta_1 + 7,8 \cdot \beta_2 \\ 3,3 &= 7,8 \cdot \beta_1 + 60,0 \cdot \beta_2 \end{aligned} \right.$$

$$\left\{ \begin{aligned} \beta_1 &= -0,13 \\ \beta_2 &= 0,07 \end{aligned} \right.$$

Encontrando  $\alpha$ :

$$\alpha + \beta_1 \cdot \bar{X}_1 + \beta_2 \cdot \bar{X}_2 = \bar{Y} \longrightarrow \alpha = \frac{232,5}{8} - \left[ (-0,13) \cdot \frac{784,5}{8} + 0,07 \cdot \frac{916,0}{8} \right] = 31,5$$

Equação da regressão:  $\hat{Y} = 31,5 - 0,13 \cdot X_1 + 0,07 \cdot X_2$

**b) Se o preço relativo for de 120 e a propaganda for de 111, qual é a estimativa para a porcentagem de ações no mercado?**

$$R.: \hat{Y} = 31,5 - 0,13 \cdot 120 + 0,07 \cdot 111 = 26,57$$

FONTE: Adaptado de Silver (2000)

## MÉTODOS QUANTITATIVOS

2 Um conjunto de experimentos foi realizado para determinar uma maneira de prever o tempo de cozimento y em junção de vários níveis de largura do forno e da temperatura do gás. Os dados obtidos foram os seguintes:

Experimento	Tempo de cozimento	Largura do forno	Temperatura do gás
1	6,40	1,32	1,15
2	15,05	2,69	3,40
3	18,75	3,56	4,10
4	30,25	4,41	8,75
5	44,85	5,35	14,82
6	48,94	6,20	15,15
7	51,55	7,12	15,32
8	61,50	8,87	18,18
9	100,44	9,80	35,19
10	111,42	10,65	40,40

Estime a equação da regressão múltipla.

R.:

i	$y_i$	$x_{1i}$	$x_{2i}$	$x_{1i}^2$	$x_{2i}^2$	$x_{1i} \cdot y_i$	$x_{2i} \cdot y_i$	$x_{1i} \cdot x_{2i}$
1	6,40	1,32	1,15	1,70	1,30	8,40	7,40	1,50
2	15,05	2,69	3,40	7,20	11,60	40,50	51,20	9,10
3	18,75	3,56	4,10	12,70	16,80	66,80	76,90	14,60
4	30,25	4,41	8,75	19,40	76,60	133,40	264,70	38,60
5	44,85	5,35	14,82	28,60	219,60	239,90	664,70	79,30
6	48,94	6,20	15,15	38,40	229,50	303,40	741,40	93,90
7	51,55	7,12	15,32	50,70	234,70	367,00	789,70	109,10
8	61,50	8,87	18,18	78,70	330,50	545,50	1118,10	161,30
9	100,44	9,80	35,19	96,00	1238,30	984,30	3534,50	344,90
10	111,42	10,65	40,40	113,40	1632,20	1186,60	4501,40	430,30
TOTAL	489,20	60,00	156,50	447,00	3991,10	3875,90	11749,90	1282,50

$$\left\{ \begin{aligned} S_{11} &= \sum X_1^2 - \frac{(\sum X_1)^2}{n} = 87,36 \\ S_{22} &= \sum X_2^2 - \frac{(\sum X_2)^2}{n} = 1543,15 \\ S_{12} &= \sum X_1 \cdot X_2 - \frac{(\sum X_1) \cdot (\sum X_2)}{n} = 344,23 \\ S_{Y1} &= \sum Y \cdot X_1 - \frac{(\sum Y) \cdot (\sum X_1)}{n} = 942,50 \\ S_{Y2} &= \sum Y \cdot X_2 - \frac{(\sum Y) \cdot (\sum X_2)}{n} = 4096,64 \end{aligned} \right.$$

$$\left\{ \begin{aligned} S_{Y1} &= \beta_1 \cdot S_{11} + \beta_2 \cdot S_{12} \\ S_{Y2} &= \beta_1 \cdot S_{12} + \beta_2 \cdot S_{22} \end{aligned} \right. \longrightarrow \left\{ \begin{aligned} \beta_1 &= 2,71 \\ \beta_2 &= 2,05 \end{aligned} \right.$$

$$\alpha + \beta_1 \cdot \bar{X}_1 + \beta_2 \cdot \bar{X}_2 = \bar{Y} \longrightarrow \alpha = 0,58$$

Equação da regressão:  $\hat{Y} = 0,58 + 2,71 \cdot X_1 + 2,05 \cdot X_2$

FONTE: Adaptado de Silver (2000)

**3 Um estudo foi conduzido para determinar se o peso de um animal pode ser previsto depois de um período com base no seu peso inicial e na quantidade de ração que ele comeu. Os dados obtidos, em quilogramas, estão na tabela a seguir:**

## MÉTODOS QUANTITATIVOS

Animal	Peso final	Peso inicial	Peso ração
1	95	42	272
2	77	33	226
3	80	33	259
4	100	45	292
5	97	39	311
6	70	36	183
7	50	32	173
8	80	41	236
9	92	40	230
10	84	38	235

a) Encontre a equação do plano que melhor ajuda os dados obtidos.

R.:

i	$y_i$	$x_{1i}$	$x_{2i}$	$x_{1i}^2$	$x_{2i}^2$	$x_{1i} \cdot y_i$	$x_{2i} \cdot y_i$	$x_{1i} \cdot x_{2i}$
1	95	42	272	1764,0	73984,0	3990,0	25840,0	11424,0
2	77	33	226	1089,0	51076,0	2541,0	17402,0	7458,0
3	80	33	259	1089,0	67081,0	2640,0	20720,0	8547,0
4	100	45	292	2025,0	85264,0	4500,0	29200,0	13140,0
5	97	39	311	1521,0	96721,0	3783,0	30167,0	12129,0
6	70	36	183	1296,0	33489,0	2520,0	12810,0	6588,0
7	50	32	173	1024,0	29929,0	1600,0	8650,0	5536,0
8	80	41	236	1681,0	55696,0	3280,0	18880,0	9676,0
9	92	40	230	1600,0	52900,0	3680,0	21160,0	9200,0
10	84	38	235	1444,0	55225,0	3192,0	19740,0	8930,0
TOTAL	825,0	379,0	2417,0	14533,0	601365,0	31726,0	204569,0	92628,0

$$\left\{ \begin{array}{l} S_{11} = \sum X_1^2 - \frac{(\sum X_1)^2}{n} = 168,90 \\ S_{22} = \sum X_2^2 - \frac{(\sum X_2)^2}{n} = 17176,10 \\ S_{12} = \sum X_1 \cdot X_2 - \frac{(\sum X_1) \cdot (\sum X_2)}{n} = 1023,70 \\ S_{Y1} = \sum Y \cdot X_1 - \frac{(\sum Y) \cdot (\sum X_1)}{n} = 458,50 \\ S_{Y2} = \sum Y \cdot X_2 - \frac{(\sum Y) \cdot (\sum X_2)}{n} = 5166,50 \end{array} \right.$$

$$\left\{ \begin{array}{l} S_{Y1} = \beta_1 \cdot S_{11} + \beta_2 \cdot S_{12} \\ S_{Y2} = \beta_1 \cdot S_{12} + \beta_2 \cdot S_{22} \end{array} \right. \longrightarrow \left\{ \begin{array}{l} \beta_1 = 1,40 \\ \beta_2 = 0,22 \end{array} \right.$$

$$\alpha + \beta_1 \cdot \bar{X}_1 + \beta_2 \cdot \bar{X}_2 = \bar{Y} \longrightarrow \alpha = -23,0$$

Equação da regressão:  $\hat{Y} = -23 + 1,40 \cdot X_1 + 0,22 \cdot X_2$

**b) Dê uma previsão para o peso final do animal que tinha como peso inicial 35 kg e foi alimentado com 250 kg de ração.**

$$R.: \hat{Y} = -23 + 1,40 \cdot 35 + 0,22 \cdot 250 = 80,3$$



**UNIASSELVI**

Construa sua própria história.