

Lista 1

Exercício 1 - Classificar quanto a natureza, as seguintes variáveis

Variável	Tipo (natureza)
Condição de saúde (doente, não doente)	
Tipo de parto (normal, cesáreo)	
Nível de colesterol sérico (mg/100cc)	
Tempo de um procedimento cirúrgico (minutos)	
Número de praias consideradas poluídas	
Custo do procedimento (reais)	
Peso (g)	
Estado nutricional (desnutrição, eutrofia, sobrepeso, obesidade)	
Consumo de energia (Kcal)	
Realização da refeição café da manhã (sim/não)	
Número de escolares por série	
Realização de atividade física diária (sim/não)	
Tempo assistindo TV/dia (< 2h, 2 a 4h, >4h)	
Porções consumidas por grupo de alimentos	
Percentual de gordura corporal (%)	

Exercício 2 - Os dados a seguir são de peso (kg) de 80 mulheres identificadas pela variável id (identificação).

Id	Peso	Id	peso	Id	Peso	Id	Peso	Id	Peso	Id	Peso
1	65	16	71	31	70	46	75	61	68	76	75
2	65	17	84	32	72	47	79	62	69	77	79
3	58	18	63	33	75	48	79	63	76	78	73
4	59	19	64	34	76	49	82	64	77	79	82
5	67	20	65	35	77	50	83	65	80	80	76
6	68	21	74	36	78	51	65	66	81		
7	74	22	81	37	80	52	68	67	59		
8	81	23	66	38	82	53	75	68	64		
9	66	24	69	39	63	54	76	69	70		
10	61	25	71	40	66	55	78	70	80		
11	64	26	71	41	72	56	78	71	85		
12	65	27	72	42	72	57	81	72	70		
13	67	28	73	43	72	58	85	73	71		
14	68	29	75	44	73	59	66	74	72		
15	70	30	77	45	73	60	68	75	72		

Fonte: Osborn JF. *Statistical Exercises in Medical Research*. John Wiley & Sons Inc., 1979. (adaptado).

- Sorteie uma amostra aleatória de tamanho 20 utilizando a tabela dos números equiprováveis.
- Apresente os valores do peso dos indivíduos sorteados.
- Some os valores e divida pelo tamanho da amostra (número de valores).
- Este valor é o parâmetro, o estimador ou a estimativa do peso médio?

Exercício 3 - Utilize os dados do Exercício 2.

- Sorteie uma amostra sistemática de tamanho 20. Indique o intervalo de amostragem e o começo casual sorteado. Indique o número de identificação de cada elemento da amostra.
- Some os valores e divida pelo tamanho da amostra (número de valores).
- Compare com o peso médio obtido no Exercício 2. Você esperaria o mesmo resultado? Justifique.

d) Qual dos dois valores você diria que representa melhor o conjunto de dados? Justifique.

Exercício 4 - Os dados a seguir são de altura de uma amostra de 351 mulheres idosas selecionadas aleatoriamente de uma comunidade para um estudo de osteoporose.

142	152	154	156	157	158	160	161	163	164	165	169
145	152	154	156	157	158	160	161	163	164	165	169
145	152	154	156	157	158	160	161	163	164	165	169
145	152	154	156	157	158	160	161	163	164	165	169
146	152	155	156	157	158	160	161	163	164	166	169
147	152	155	156	157	158	160	161	163	164	166	169
147	153	155	156	158	158	160	161	163	164	166	169
147	153	155	156	158	158	160	161	163	164	166	170
147	153	155	156	158	159	160	162	163	164	166	170
148	153	155	156	158	159	160	162	163	164	166	170
148	153	155	156	158	159	160	162	163	164	166	170
149	153	155	156	158	159	160	162	163	164	166	170
150	153	155	156	158	159	160	162	163	164	166	170
150	153	155	156	158	159	160	162	163	164	166	170
150	153	155	157	158	159	160	162	163	165	167	170
150	153	155	157	158	159	160	162	163	165	167	170
150	153	155	157	158	159	161	162	163	165	167	170
151	153	155	157	158	159	161	162	163	165	167	171
151	153	155	157	158	159	161	162	163	165	167	171
151	153	155	157	158	159	161	162	163	165	167	171
151	153	155	157	158	159	161	162	163	165	167	173
151	153	155	157	158	159	161	162	163	165	168	173
151	154	155	157	158	159	161	162	163	165	168	173
152	154	155	157	158	159	161	162	163	165	168	174
152	154	156	157	158	160	161	162	163	165	168	176
152	154	156	157	158	160	161	163	163	165	168	177
152	154	156	157	158	160	161	163	164	165	168	178
152	154	156	157	158	160	161	163	164	165	169	178
152	154	156									

Fonte: Hand DJ et alli. *A handbook of small data sets*. Chapman&Hall, 1994.

- Faça uma apuração dos dados e represente-os em uma tabela.
- Interprete os resultados.
- Se entre as 351 mulheres não fossem conhecidas as alturas de 15 delas, como você representaria esses valores?

Exercício 5

Os dados a seguir são relativos ao número de refeições diárias de 50 indivíduos.

2	3	2	1	2	6	5	4	3
1	2	2	1	2	5	6	4	3
2	2	3	2	3	4	2	3	2
3	2	3	3	3	4	3	4	5
3	1	4	3	4	4	3		
3	1	6	4	4	2	4		

Fonte: X.

- Apresente os dados em uma tabela.
- Interprete a dispersão dos dados.

Exercício 6

Os dados a seguir são provenientes do grupo *Western Collaborative Group Study*, criado na Califórnia em 1960-61. Foram estudados 3154 homens de meia idade para investigar a relação entre padrões de comportamento e risco de doença coronariana. Os dados apresentados são de 40 homens para os quais foram medidos os níveis de colesterol (mg/100ml) e realizada uma categorização segundo comportamento. O comportamento de tipo A é caracterizado pela urgência, agressividade e ambição. O de tipo B é relaxado, não competitivo e menos preocupado.

Tipo A: nível de colesterol									
233	291	312	250	246	197	268	224	239	239
254	276	234	181	248	252	202	218	212	325

Tipo B: nível de colesterol									
344	185	263	246	224	212	188	250	148	169
226	175	242	252	153	183	137	202	194	213

- Quais variáveis que estão sendo estudadas? Identifique a natureza de cada variável.
- Apure os dados e apresente a variável nível de colesterol em uma tabela bidimensional, considerando os níveis A e B.
- Classifique a variável nível de colesterol em duas categorias: nível normal (abaixo de 160 mg/100ml) e nível elevado (160 mg/100ml e mais) e faça uma tabela bidimensional cruzando as variáveis: nível de colesterol (normal e alto) e tipo de comportamento (A e B). Interprete os resultados.

Exercício 7

Os dados a seguir são relativos ao peso ao nascer (g) de recém nascidos com síndrome de desconforto idiopático grave. Algumas crianças foram a óbito (*) e outras sobreviveram.

1050*	2500*	1890*	1760	2830
1175*	1030*	1940*	1930	1410
1230*	1100*	2200*	2015	1715
1310*	1185*	2270*	2090	1720
1500*	1225*	2440*	2600	2040
1600*	1262*	2560*	2700	2200
1720*	1295*	2730*	2950	2400
1750*	1300*	1130	2550	3160
1770*	1550*	1575	2570	3400
2275*	1820*	1680	3005	3640

Fonte: Hand DJ *et al.*, 1994.

- Classifique a variável peso ao nascer em duas categorias: baixo peso (abaixo de 2500 g) e não baixo peso (2500 g e mais) e faça uma tabela bidimensional cruzando as variáveis: condição do recém-nascido (sobrevivente ou não sobrevivente) e peso ao nascer (baixo peso e não baixo peso).
- Interprete os resultados.

Exercício 8

A tabela abaixo foi extraída do artigo Tendência secular do peso ao nascer na cidade de São Paulo (1976-1998) de MONTEIRO CA *et al.* (*Rev. Saúde Pública*; 2000:34 (6, supl): 26-40). Comente os resultados apresentados.

Tabela 1 – Distribuição de nascidos vivos segundo intervalos do peso ao nascer nas cidades de São Paulo, SP (Brasil) e Gotemburgo (Suécia).

Peso ao nascer (gramas)	São Paulo -1998		Gotemburgo -1972/73	
	N	%	N	%
500 – 1.000	940	0,43	29	0,24
1.001 – 1.500	1.650	0,76	45	0,37
1.501 – 2.000	3.519	1,62	113	0,93
2.001 – 2.500	13.187	6,09	368	3,02
2.501 – 3.000	53.882	24,87	1.689	13,85
3.001 – 3.500	90.151	41,60	4.351	35,69
3.501 – 4.000	44.045	20,33	3.980	32,64
4.001 – 4.500	8.225	3,80	1.356	11,12
4.501 – 5.000	967	0,45	250	2,05
5.001 – 6.000	116	0,05	11	0,09
Total	216.682	100,00	12.192	100,00

Exercício 9

Apresente os dados da tabela abaixo em um gráfico e interprete o mesmo. Utilize as informações de produção de leite em mil litros/ano.

Tabela 2

Municípios brasileiros com maior densidade de produção, volume produzido e produtividade por animal, 2012

Município	Produção de leite		
	Mil litros/ano	Litros/vaca/ano	Litros/Km ² /ano
Westfalia - RS	16.996	4.721	266.805
Coronel Freitas - SC	59.400	3.960	253.676
São João do Oeste - SC	40.440	3.370	247.113
Casca - RS	65.200	5.400	239.935
Anchieta - SC	54.000	3.000	236.241
Cunhataí - SC	12.300	2.847	225.643
Boa Vista do Buricá - RS	23.536	5.105	216.459
Estrela - RS	37.950	4.929	206.051
Palmitos - SC	70.600	2.521	201.317
São Carlos - SC	32.000	2.735	201.273
Carambeí - PR	129.600	5.891	199.483
Nova Erechim - SC	12.831	3.554	199.239
Teutônia - RS	35.550	4.444	198.415
Tunápolis - SC	25.200	3.150	189.603
Guarujá do Sul - SC	18.630	2.700	185.281
TOTAL / MÉDIA	634.233	3.889	217.769

Fonte: IBGE/PPM 2013.

Exercício 10

Os dados a seguir são relativos ao número de refeições diárias de 50 indivíduos, utilizados no exercício 5.

2	3	2	1	2	6	5	4	3
1	2	2	1	2	5	6	4	3
2	2	3	2	3	4	2	3	2
3	2	3	3	3	4	3	4	5
3	1	4	3	4	4	3		
3	1	6	4	4	2	4		

- Apresente os dados em um gráfico.
- Interprete o gráfico.

Exercício 11

Os dados são referentes a produção (kg) de carne de peixes e de carne vermelha e de carneiro por pessoa, no mundo, no período de 1950 a 2000.

Ano	Pesca oceânica (kg)	Carne vermelha e de carneiro (kg)
1950	7,9	9,0
1960	12,0	10,0
1970	16,1	12,0
1980	15,5	11,9
1990	16,3	12,0
2000	15,0	11,7

Fonte: State of the World, 2001. The Worldwatch Institute.

- Apresente os dados em um gráfico.
- Interprete os resultados.

Exercício 12

Os dados são referentes à distribuição de pacientes segundo taxa de albumina no sangue (g/dL).

Taxa de albumina (g/dL)	Nº	%
4,40 -4,60	6	10,0
4,60 -4,80	11	18,3
4,80 -5,00	14	23,3
5,00 -5,20	18	30,0
5,20 -5,40	8	13,3
5,40 -5,60	2	3,3
5,60 -5,80	0	-
5,80 -6,00	1	1,7
Total	60	100

Fonte: Soares JF, Siqueira AL. COOPMED, 2002.

- Apresente os dados em um histograma.
- Interprete os resultados.

Exercício 13

Os dados a seguir são da altura (cm) de uma amostra de mulheres de Bangladesh.

Altura (cm)	número
137,0 --140,0	71
140,0 --143,0	137
143,0 --145,0	154
145,0 --147,0	199
147,0 --150,0	279
150,0 --153,0	221
153,0 --155,0	94
155,0 --157,0	51
157,0 --160,0	37
Total	1243

Fonte: Hand DJ et al, 1994 (adaptado).

- Represente os dados acima, graficamente em um histograma.
- Interprete os resultados.

Exercício 14

Os dados a seguir são referentes à distribuição de usuárias do Serviço de Saúde X segundo idade (anos). Município de São Paulo, 2009.

Idade (anos)	n	%
15 -- 20	14	19,5
20 -- 25	24	33,3
25 -- 30	16	22,2
30 -- 35	9	12,5
35 -- 40	8	11,1
40 --45	1	1,4
Total	72	100

Fonte: Dados hipotéticos.

- Apresente a variável em um polígono de frequências simples.
- Interprete os resultados.

Exercício 15

Distribuição de homens segundo nível de glicose no sangue (mg%).

Nível de glicose no sangue (mg%)	n
50 -- 100	13
100 -- 150	45
150 -- 200	28
200 -- 250	10
250 -- 300	3
300 -- 450	1

Fonte: X.

- Apresente os dados acima graficamente utilizando o polígono de frequências simples.
- Interprete os resultados.

Exercício 16

Os dados a seguir são medidas de circunferência do tórax (polegadas) de 5732 soldados escoceses apresentados pelo matemático belga *Adolphe Quetelet* (1796–1874).

Medida (polegada)	número	%	% acumulada
33,0 – 34,0	3		
34,0 – 35,0	19		
35,0 – 36,0	81		
36,0 – 37,0	189		
37,0 – 38,0	409		
38,0 – 39,0	753		
39,0 – 40,0	1062		
40,0 – 41,0	1082		
41,0 – 42,0	935		
42,0 – 43,0	646		
43,0 – 44,0	313		
44,0 – 45,0	168		
45,0 – 46,0	50		
46,0 – 47,0	18		
47,0 – 48,0	3		
48,0 – 49,0	1		
Total	5732		

Fonte: Daly F et al. Elements of Statistics, 1999.

- Represente os dados em um polígono de frequências acumuladas.
- Utilizando o gráfico, identifique o valor da circunferência de tórax que deixa 25% dos indivíduos abaixo.
- Qual o valor de circunferência do tórax que divide a distribuição em 2 partes iguais, isto é, qual é o valor da variável que deixa 50% das observações abaixo dele?
- Qual a proporção de soldados com circunferência do tórax entre 40 a 42 polegadas?
- Qual é o valor de circunferência do tórax que deixa 95% dos soldados abaixo dele?

Exercício 17

A tabela apresenta dados de classificação de pessoas segundo doença coronariana (CHD) segundo hábito de consumo de café para uma coorte de 1718 homens com idade 40-55 anos.

Distribuição de homens segundo presença de doença coronariana (CHD) e consumo de café.

condição	Alto consumo de café (≥ 100 xícaras/mês)		Moderado consumo de café (<100 xícaras/mês)		Total	
	n	%	n	%	n	%
Com CHD	38	4,8	39	4,2	77	4,5
Sem CHD	752	95,2	889	95,8	1641	95,5
Total	790	100	928	100	1718	100

- Apresente os dados graficamente.
- Interprete os resultados.

Exercício 18

Utilize os dados da tabela e apresente-os graficamente.

Distribuição de escolares de 7 a 10 anos segundo peso e sexo. Duas escolas do Município de São Paulo, 2005.

<i>Peso (kg)</i>	<i>Sexo</i>	
	<i>Masculino</i>	<i>Feminino</i>
15,0 -- 25,0	52	68
25,0 -- 35,0	146	132
35,0 -- 45,0	59	53
45,0 -- 55,0	11	18
55,0 -- 65,0	10	2
65,0 -- 75,0	3	1
75,0 -- 85,0	0	0
85,0 -- 95,0	0	1
Total	281	275

Fonte: Koga CR. Estado nutricional de escolares de 7 a 10 anos de idade: diagnóstico e comparação de métodos. São Paulo; 2005. [Dissertação de Mestrado-Faculdade de Saúde Pública da Universidade de São Paulo/USP].

Exercício 19

Utilize os dados da tabela e apresente-os graficamente.

Distribuição de percentual (%) de escolares segundo estatura (cm) e sexo.

<i>Estatura (cm)</i>	<i>Sexo</i>	
	<i>Masculino</i>	<i>Feminino</i>
105,0 – 119,9	3	16
120,0 – 124,9	36	31
125,0 – 129,9	61	74
130,0 – 134,9	57	41
135,0 – 139,9	52	43
140,0 – 144,9	38	30
145,0 – 149,9	22	26
150,0 – 159,9	12	14
Total	281	275

Fonte: Koga CR. Estado nutricional de escolares de 7 a 10 anos de idade: diagnóstico e comparação de métodos. São Paulo; 2005. [Dissertação de Mestrado-Faculdade de Saúde Pública da Universidade de São Paulo/USP].

Exercício 20

Os dados a seguir são referentes à mortalidade por câncer de esôfago, segundo sexo, no município de São Paulo no período de 1968-1998.

Coeficientes de mortalidade por câncer de esôfago (por 100.000 hab.).
Município de São Paulo, 1968-1998.

Ano	Masculino	Feminino
1968	8,81	2,00
1973	12,38	2,61
1978	10,93	1,98
1983	9,41	2,00
1988	8,60	1,67
1993	8,33	1,27
1998	8,37	1,12

Fonte: Incidência de câncer no Município de São Paulo, 1997-1998. Registro de Câncer de São Paulo. FSP/USP.

- a) Represente os coeficientes de mortalidade por câncer de esôfago para o sexo masculino e feminino em um único gráfico, utilizando escala aritmética.
- b) Represente os coeficientes de mortalidade por câncer de esôfago para o sexo masculino e feminino em um único gráfico utilizando escala logarítmica.
- c) Comente os gráficos dos itens a e b. Qual a melhor representação para os dados?

Exercício 21

Apresente a variável peso ao nascer graficamente utilizando a variável definida em duas categorias, conforme tabela abaixo.

Distribuição de recém-nascidos acometidos de síndrome de desconforto idiopático grave segundo peso ao nascer (g). Austrália, 1993.

Peso(g)	Nº	%
Baixo peso (<2500 g)	37	74,0
Não baixo peso (2500 g e mais)	13	26,0
Total	50	100

Fonte: Hand DJ et al. *A handbook of small data sets*. Chapman&Hall, 1994.

Exercício 22

Apresentar a variável comprimento ao nascer em um histograma.

Distribuição de recém-nascidos segundo comprimento ao nascer (cm). Hospital X, 2009.

Comprimento (cm)	n	%
40 --43	1	1,2
43 --46	45	55,6
46 --49	25	30,9
49 --52	4	5
52 --55	3	3,7
55 --58	1	1,2
58 --61	2	2,5
Total	81	100

Fonte: dados hipotéticos.

Exercício 23

Apresentar a variável idade da mãe em um polígono de frequências.

Distribuição de mães segundo idade (anos). Centro de Saúde X, 2009.

Idade (anos)	n	%
15 -- 20	14	19,5
20 -- 25	24	33,3
25 -- 30	16	22,2
30 -- 35	9	12,5
35 -- 40	8	11,1
40 --45	1	1,4
Total	72	100

Fonte: dados hipotéticos.

Exercício 24

Defeitos do tubo neural são má formações congênitas que surgem durante o desenvolvimento fetal. É conhecida como spina bífida. Estes dados são de um estudo realizado no país de Gales – Reino Unido, para investigar possível associação entre defeito do tubo neural e dieta materna. O estudo é do tipo caso-controle: mães que tinham tido bebês com spina bífida (casos) e suas irmãs que não tinham tido (controles) foram avaliadas segundo suas dietas e classificadas em boa, razoável e ruim.

Distribuição de recém-nascidos casos (acometidos de spina bífida) e controles segundo dieta da mãe.

Dieta	Casos		Controles		Total	
	n	%	n	%	n	%
Boa	34		43		77	
Razoável	110		48		158	
Pobre	100		32		132	
Total	244		123		367	

Fonte: Hand DJ ET al., 1994.

- Calcular percentuais tomando-se como 100% o grupo (caso, controle) e interprete os resultados.
- Apresentar os dados em um gráfico.

Exercício 25

Represente os dados da tabela em um polígono de frequências e interprete os resultados. Trata-se de condenados por embriaguez em Londres, 1970.

Idade	Homens		Mulheres	
	Número	%	Número	%
0 --30	185	20,5	4	9,1
30 -- 40	207	22,9	13	29,5
40 -- 50	260	28,8	10	22,7
50 --60	180	19,9	7	15,9
60 --80	71	7,9	10	22,7
total	903	100	44	100

Fonte: Hand DJ et alli. *A handbook of small data sets*. Chapman&Hall, 1994.