

**UNIVERSIDADE DE SÃO PAULO
FACULDADE DE SAÚDE PÚBLICA
DEPARTAMENTO DE EPIDEMIOLOGIA**

DISCIPLINA

HEP 5800 - BIOESTATÍSTICA

EXERCÍCIOS

Elaboração:

***Denise Pimentel Bergamaschi
Gizelton Pereira Alencar
Maria do Rosário Oliveira Dias Latorre
Sabina Léa Davidson Gotlieb***

SÃO PAULO – 2013

1. APURAÇÃO DE DADOS E APRESENTAÇÃO TABULAR

- 1) Com os dados do sistema de referência apresentado na página 3:
- Classifique as variáveis apresentadas quanto à natureza;
 - Cite as variáveis quantitativas que poderiam ser transformadas em qualitativas; apresente as novas modalidades de cada variável;
 - Apure os dados e construa tabelas que representem distribuições de freqüências simples para cada variável (tabelas unidimensionais);
 - Construa uma tabela completa com a distribuição dos recém-nascidos segundo as variáveis sexo e peso ao nascer (em intervalos de classe) e analise;
 - Construa uma tabela completa com a distribuição dos recém-nascidos segundo as variáveis idade da mãe (em classes quinquenais) e número de gestações prévias (em 3 categorias);
 - Repita o item d, usando, para definir as categorias do peso ao nascer, o seguinte critério:
 - baixo peso: menor que 2500g
 - não baixo peso: igual ou maior que 2500g
 - Construa uma tabela completa com a distribuição dos recém-nascidos segundo as variáveis tipo de parto e peso ao nascer, de acordo com o item f.
- 2) Um levantamento foi realizado com o objetivo de estimar a prevalência de verminose entre escolares. Foram observadas 500 crianças, das quais 280 eram do sexo masculino e, entre elas, 196 apresentaram sinais de verminose. Entre as meninas, 50% apresentavam sinais de verminose.
- Construa a distribuição de freqüências com esses dados;
 - No total (n=500), qual a proporção de estudantes com sinais de verminose?
 - Comente os resultados encontrados nos itens a e b.
- 3) A tabela 1.3 apresenta incorreções, do ponto de vista de apresentação tabular. Analise os erros e refaça a tabela.

Tabela 1.3 - Peso ao nascer (g), segundo sexo do recém-nascido. Local X, 1999.

Peso (g)	SEXO			
	Masculino	Feminino		
1500 -- 2000	0	0	15	5,5556
2000 -- 2500	10	10	15	16,666
2500 -- 3000	25	25	15	16,666
3000 -- 3500	40	40	35	38,8888
3500 -- 4500	25	25	20	22,222
TOTAL	100	100	90	99,9984

4) Aponte os erros da tabela 1.4. Refaça a mesma.

Tabela 1.4 - Taxa de mortalidade pela doença D (por 100.000 hab.), Área A, 1999

Grupos Etários	Sexo		Geral
	Masculino	Feminino	
15 a 20	6,9	3,53	5,2
20 a 30			15,7
30 a 40	20,2	10,32	16,8
40 a 49	40,3	32,21	30,7
Total	38,5	28,76	33,2

Fonte: (dados hipotéticos)

5) Tabela 1.5. Número de pacientes segundo tipo de parto e frequência ao pré-natal Maternidade X, 1999.

Tipo de parto	Pré-natal		Total
	Sim	Não	
Normal	120	20	140
Cesáreo	40	20	60
Total	160	40	200

Fonte: SAME, Maternidade X

Em relação à tabela 1.5, o que representam as seguintes relações ou razões:

a) $\frac{160}{200}$

b) $\frac{140}{200}$

c) $\frac{120}{160}$

d) $\frac{120}{140}$

e) $\frac{40}{200}$

f) $\frac{20}{60}$

g) $\frac{40}{160}$

h) $\frac{60}{200}$

Quadro 1 - Relação dos nascidos vivos com dados sobre peso ao nascer (em gramas), duração da gestação (em dias), número de gestações prévias e idade da mãe (em anos). Local A, 1999.

id	Peso (g)	Tipo de parto	Sexo	Duração da gestação (dias)	Número de gestações prévias	Idade da mãe (anos)
1	2.920	Cesáreo	M	ign.	1	20
2	3.560	Normal	M	284	1	25
3	3.920	Cesáreo	M	286	2	23
4	3.050	Normal	F	281	2	30
5	3.020	Normal	F	283	3	30
6	4.200	Cesáreo	M	283	0	17
7	3.000	Cesáreo	F	280	0	18
8	3.050	Normal	F	ign.	0	16
9	3.030	Cesáreo	M	268	3	43
10	4.020	Normal	M	289	2	18
11	3.000	Normal	M	279	0	19
12	3.200	Normal	M	290	1	28
13	4.200	Cesáreo	F	261	0	22
14	3.250	Normal	F	284	0	21
15	1.950	Cesáreo	F	260	3	29
16	3.250	Normal	M	286	0	19
17	3.450	Normal	M	275	0	22
18	3.400	Normal	M	290	1	16
18	4.020	Normal	M	281	2	24
20	2.300	Cesáreo	M	244	0	16
21	2.450	Normal	F	306	2	27
22	3.150	Normal	M	258	0	17
23	1.990	Normal	M	263	0	21
24	2.780	Normal	F	ign.	4	37
25	2.250	Cesáreo	M	261	1	21
26	2.700	Cesáreo	F	245	2	25
27	2.900	Normal	F	275	0	20
28	2.780	Normal	F	266	3	32
29	2.880	Normal	M	254	1	31
30	2.850	Normal	F	269	2	37
31	2.800	Cesáreo	F	276	4	30
32	3.200	Normal	F	290	1	26
33	3.260	Normal	F	276	1	20
34	4.090	Cesáreo	M	279	0	22
35	3.260	Normal	M	274	2	23
36	3.240	Normal	F	279	1	21
37	3.460	Normal	M	278	2	19
38	3.250	Normal	F	ign.	0	18
39	3.450	Normal	M	296	2	23
40	3.610	Normal	F	278	2	25
41	3.610	Normal	F	281	0	15
42	3.190	Normal	M	284	2	44
43	2.630	Normal	F	262	0	22
44	3.200	Normal	M	296	2	32
45	3.640	Normal	F	245	1	28
46	3.710	Normal	M	270	1	32
47	3.510	Normal	F	296	0	19
48	2.710	Cesáreo	F	ign.	1	19
49	3.670	Normal	M	277	2	24
50	3.380	Normal	M	266	2	24

FONTE: Ruiz F. - Conceitos básicos de estatística, demografia e mortalidade. Ministério da Saúde. Brasília, 1976 (modificado)

2. APRESENTAÇÃO GRÁFICA

1) Faça a representação gráfica das tabelas 2.1a, 2.1b, 2.1c, 2.1d, 2.1e, 2.1f.

Tabela 2.1a. Mortalidade proporcional (%) no sexo feminino segundo os subgrupos de causas violentas. Residentes no Município de São Paulo, 1972.

Causa	%
Acidentes com veículos a motor	52,1
Outros acidentes	23,0
Suicídios	14,2
Homicídios	8,1
Demais causas externas	2,6
Total	100

FONTE: Silveira, MH e Gotlieb, SLD. Acidentes, envenenamentos e violências como causa de morte dos residentes no Município de São Paulo. *Rev. Saúde Públ.*, 10: 45-55, 1976.

Tabela 2.1b. Taxa de mortalidade (por 100.000 hab.) por meningite meningocócica anual e no mês de junho, em São Paulo, de 1968 a 1974.

Ano	Anual	Mês de junho
1968	1,4	0,13
1969	1,5	0,12
1970	2,1	0,22
1971	6,6	0,37
1972	15,6	1,49
1973	26,5	2,24
1974	17,0	6,26

FONTE: *Rev. Saúde Públ.*, 10: 1-16, 1976

Tabela 2.1c. Crianças menores de um ano segundo registro de nascimento e escolaridade do pai (em anos). Salvador, 1978 (nº e %)

Escolaridade do pai	Registro de Nascimento				Total	
	Sim		Não		n	%
	n	%	n	%		
0 -- 4	60	21,0	36	33,3	96	24,4
4 -- 8	95	33,0	44	40,7	139	35,3
8 -- 12	49	17,0	19	17,6	68	17,3
12 -- 16	45	16,0	6	5,6	51	12,9
16 -- 20	35	12,0	3	2,8	38	9,6
20 -- 24	2	1,0	-	-	2	0,5
Total	286	100,0	108	100,0	394	100,0

FONTE: *Rev. Saúde Públ.*, 13: 208, São Paulo, 1979.

Você concorda com a construção da tabela 2.1c? Caso não, como você construiria? Faça o gráfico.

Tabela 2.1d. Número e proporção (%) de internados segundo dia da semana. Hospital X. 1ª semana de agosto de 1999.

Dia da semana	n	%
2ª feira	48	28,4
3ª feira	35	20,7
4ª feira	30	17,8
5ª feira	22	13,0
6ª feira	16	9,5
Sábado	10	5,9
Domingo	8	4,7
Total	169	100,0

FONTE: SAME, Hospital X

Tabela 2.1e. Gestantes segundo gestações prévias. Centro de Saúde X, 1999.

Gestações prévias	n
0	10
1	20
2	30
3	20
4	8
5	2
6	2
Total	92

FONTE: Centro de Saúde X

Tabela 2.1f. Distribuição do número de pacientes segundo idade. Programa A. Centro de Saúde X, primeiro semestre de 1999.

Idade (em anos)	n
10 -- 15	10
15 -- 20	15
20 -- 25	60
25 -- 30	65
30 -- 35	40
35 -- 40	50
40 -- 45	30
Total	270

FONTE: Centro de Saúde X

- 2) Faça um gráfico de frequências acumuladas baseando-se na tabela 2.1f e responda:
- Qual a proporção de pacientes com idade até 30 anos?
 - Qual a proporção de pacientes com idade mínima de 25 anos?
 - Qual é a idade que divide a distribuição em 2 partes iguais, i.é, 50% têm idade abaixo desse valor e 50%, acima?
 - Até qual idade há 70% dos pacientes?
 - Qual a proporção de pacientes com idade entre 15 e 34 anos completos?

3) Faça um gráfico de frequências acumuladas com os dados da tabela 2.3 e responda:

Tabela 2.3 – Número de nascidos vivos segundo peso ao nascer (g). Área X, 1999.

Peso ao nascer (g)	n	%
1.500 -- 2.500	1.200	10,0
2.500 -- 3.000	3.600	30,0
3.000 -- 3.500	4.800	40,0
3.500 -- 4.500	2.400	20,0
Total	12.000	100,0

FONTE: SAME, Maternidade X

- Qual a proporção de nascidos vivos com peso mínimo de 3.500g?
 - Qual a proporção de nascidos vivos com peso menor do que 2.500g?
 - Qual o peso que divide a distribuição em 2 metades iguais, isto é, o peso tal que 50% dos nascidos vivos pesam abaixo desse valor e 50%, pesam acima ?
 - Até que peso ao nascer, se apresentam 70% dos recém-nascidos?
 - Qual a proporção de recém nascidos que pesam de 2.500 a 3.500g?
- 4) Construa um gráfico relativo à questão 1, item d ,da primeira unidade.
- 5) Construa um gráfico relativo à questão 5 da primeira unidade.
- 6) Construa um gráfico relativo às tabelas 2.3 e 2.6.

Tabela 2.6. Número de pacientes da doença D segundo grupos etários (em anos). Local X, 2006

Grupo etário	n
0 - 1	10
1 - 5	55
5 - 10	48
10 - 15	45
15 - 20	40
20 - 30	80
30 - 50	92
Total	370

Fonte: Serviço de Vigilância da SMS de X

3. MEDIDAS DE TENDÊNCIA CENTRAL, DE DISPERSÃO E DE CORRELAÇÃO

- 1) Baseando-se nos dados do Quadro 1 apresentado na página 3, calcule:
- Os pesos médios e medianos dos recém-nascidos, segundo sexo;
 - Pesos médios e medianos dos nascidos com os dados da questão 1 d, (tabela 1.1g);
 - As variâncias, desvios padrão e coeficientes de variação, com os dados da questão 1a e 1b, deste exercício, e interprete as diferenças encontradas.
- 2) Calcule as idades média e mediana e desvio padrão da distribuição apresentada na tabela 3.1.

Tabela 3.1 - Número de pacientes atendidos segundo idade. Centro de Saúde X, 1999.

Idade (em anos)	n de pacientes
0 - 1	480
1 - 5	200
5 - 12	70
12 e mais	250
Total	1.000

Fonte: dados hipotéticos

- 3) Considerando 31 dias no mês de abril de 1999, no hospital X, o número de internações diárias pela doença D foi igual a: 3, 5, 2, 0, 1, 7, 4, 0, 1, 9, 1, 3, 2, 5, 2, 7, 0, 5, 1, 4, 6, 0, 4, 0, 3, 5, 8, 0, 5, 0 e 6 casos. Calcule o número médio e número mediano de internações pela doença D, no hospital e o respectivo desvio padrão.
- 4) Calcule as idades média e mediana e desvio padrão a partir dos casos de câncer no hospital A em 1999.

Tabela 3.2 - Número de casos de câncer segundo idade no Hospital A, em 1999.

Idade	n
0 -- 5	10
5 -- 10	2
10 -- 15	3
15 -- 25	3
25 -- 35	6
35 -- 45	14
45 -- 55	10
55 -- 65	20
65 -- 75	8
75 -- 85	4
Total	80

FONTE: SAME, Hospital A

- 5) Calcule a média, mediana e o desvio padrão do número de gestações prévias (Tabela 3.3)

Tabela 3.3. Número de parturientes segundo o número de gestações prévias. Hospital X, 1999

Número de gestações prévias	Número de parturientes
0	40
1	35
2	20
3	5
4	3
5	2
Total	105

FONTE: SAME, Hospital X

6) Uma população de dez crianças foi avaliada pelo peso e altura. Os resultados estão a seguir.

id	Peso (kg)	Altura (cm)
1	20	108
2	22	110
3	22	113
4	23	115
5	24	120
6	25	124
7	26	122
8	28	128
9	29	125
10	31	135

a) Construa um gráfico de dispersão para a representação das duas variáveis conjuntamente.

b) Calcule a covariância e o coeficiente de correlação entre as variáveis.

7) Analise a correlação entre os indicadores. Construa o gráfico e verifique a magnitude da mesma.

Tabela 3.5 - Taxa de Mortalidade Infantil (por 1.000 n.v.) e Indicador de Swaroop-Uemura (%) segundo países, 1991

País	Swaroop-Uemura (%)	T.M.I. (por 1.000 n.v)
A	93,2	6,9
B	43,6	59,8
C	92,8	8,1
D	80,1	37,5
E	89,4	7,6
F	84,7	10,9
G	61,3	20,2
H	33,3	42,0
I	52,0	50,6
J	51,6	51,2

FONTE: X

T.M.I. – Taxa de Mortalidade Infantil.

OBS: O indicador Swaroop-Uemura (%) apresenta a porcentagem de óbitos de pessoas com 50 anos e mais de uma localidade.

10) Uma população de nascidos vivos foi examinada quanto à duração da gestação e peso ao nascer. Construa o gráfico e verifique a magnitude da correlação entre essas duas variáveis.

Tabela 3.6 - Nascidos vivos segundo peso ao nascer (kg) e duração da gestação (em dias). Local X, em 2005

id	Peso ao nascer	Duração da gestação
1	3,05	281
2	3,02	283
3	3,62	283
4	3,50	286
5	3,54	275
6	2,05	260
7	2,80	245
8	2,73	254
9	2,91	256
10	3,72	292
11	4,14	290
12	2,82	287
13	3,01	277
14	2,88	261
15	4,00	300
16	2,70	254
17	2,63	262
18	2,77	277
19	3,61	306
20	2,40	271

11) Uma população de nascidos vivos foi examinada quanto a uma variável bioquímica X, medida em g/100 ml de sangue. Calcule o valor mediano, médio, variância e desvio padrão (tabela 3.7).

Tabela 3.7. Número de nascidos vivos segundo taxa do elemento X (g/100 ml) no sangue. Local X, 2005

Variável X (g/100 ml)	n
1 3	27
3 6	38
6 10	40
10 15	25
15 25	8
Total	138

Fonte: X

4. ASSOCIAÇÃO

1) Com o objetivo de investigar a associação entre história de bronquite na infância e presença de tosse diurna ou noturna em idades mais velhas, foram observadas 1.319 crianças de 14 anos. Destas, 273 relataram história de bronquite até os 5 anos sendo que 26 tinham tosse diurna ou noturna, aos 14 anos. Estude a associação entre as variáveis (Tabela 4.1).

Tabela 4.1 - Número de crianças segundo história de bronquite aos 5 anos e tosse diurna ou noturna aos 14 anos de idade. Local X, ano Y.

Tosse	Bronquite		Total
	Sim	Não	
Sim	26	44	70
Não	247	1.002	1.249
Total	273	1.046	1.319

Fonte: Holland, WW et al. Long-term consequences of respiratory disease in infancy. *Journal of Epidemiology and Community Health* 1978; 32:256-9.

2) Estude a associação entre as variáveis apresentadas na tabela 4.2.

Tabela 4.2. Pessoas segundo hábito de fumar e presença da doença D. Local X, ano Y

Hábito de fumar	Presença de doença D		Total
	Sim	Não	
Sim	11	79	90
Não	9	201	210
Total	20	280	300

Fonte: (dados hipotéticos).

3) Para as tabelas que se seguem, verifique a associação entre as variáveis, calculando o Qui-Quadrado e o coeficiente de Yule, quando for possível.

a) Tabela 4.3a - Resultados de casos de diagnóstico de pré-natal segundo idade da mãe e a presença de aberração cromossômica.

Idade (anos)	Aberração cromossômica		Total
	Presente	Ausente	
35 -- 40	10	447	457
40 e mais	18	510	528
Total	28	957	985

Fonte: Milunsky e Atkins (1977), in Vieira, S. *Introdução à Bioestatística*, capítulo 12.

b) Tabela 4.3b - Indivíduos segundo a origem étnica e o tipo sanguíneo.

Origem étnica	Tipo sanguíneo				Total
	O	A	B	AB	
Árabe	130	149	29	8	316
Não Árabe	417	292	94	17	820
Total	547	441	123	25	1136

Fonte: Garcia (1977), in Vieira, S. Introdução à Bioestatística, capítulo 12.

c) Tabela 4.3c - Pacientes com câncer de mama segundo a faixa etária por ocasião do diagnóstico e sobrevida após três anos.

Idade (anos)	Sobrevida		Total
	Sim	Não	
< de 50	11	6	17
50 - 70	18	8	26
70 e mais	15	9	24
Total	44	23	67

Fonte: Morrison (1973), In Vieira, S. Introdução à Bioestatística, capítulo 12.

4) Estude a presença de associação entre as variáveis: condição da moradia e tipo de gestação. Os dados são apresentados na tabela 4.4.

Tabela 4.4. Número de gestantes segundo condição da moradia e tempo de gestação. Local W, ano Y

Condição da moradia	Gestação		Total
	Pré-termo	A termo	
Proprietária	50	849	899
Institucional	29	229	258
Aluguel	11	164	175
Mora com pais	6	66	72
Outro	3	36	39
Total	99	1344	1443

Fonte: In Bland, M. An introduction to medical statistics. pg 225; 1997

5) Estude a associação entre as variáveis: grupo etário e tipo de parto. Os dados são apresentados na tabela 4.5.

Tabela 4.5. Número e proporção (%) de partos realizados segundo tipo de parto e grupo etário da mãe (adolescente). Local X, 2005

Tipo de parto	10 a 14		15 a 19		Total	
	n	%	n	%	n	%
Vaginal	48	80,0	452	83,7	500	83,3
Cesário	12	20,0	88	16,3	100	16,7
Total	60	100,0	540	100,0	600	100,0

Fonte X

5. AMOSTRAGEM

- 1) Dê dois exemplos de amostragem probabilística e de não probabilística. Justifique a sua escolha.
- 2) Interessado em estimar a proporção de casas com ligação de água dentro do domicílio, de uma comunidade, um pesquisador tomou como amostra todos os domicílios das 10 principais ruas daquela área (a comunidade apresentava-se em 30 ruas). A amostra é probabilística? Justifique.
- 3) Sorteie uma amostra de tamanho $n = 21$ usando o processo sistemático, com base no rol anexo.
- 4) Construa distribuição de freqüências da amostra de domicílios segundo variável tipo de construção.
- 5) Construa distribuição de freqüências da amostra de domicílios segundo número de pessoas.
- 6) Construa uma distribuição de freqüências da amostra de domicílios por renda em salários mínimos.

Relação dos domicílios em 3 regiões de certa comunidade com as variáveis tipo de construção do domicílio, número de pessoas no domicílio e renda em salários mínimos.

Região 1							
Casa n	Tipo de construção	Número de pessoas	Renda	Casa n	Tipo de construção	Número de pessoas	Renda
1	A	3	6,5	41	C	3	3,5
2	B	2	3,6	42	A	5	4,8
3	C	5	4,9	43	B	3	5,9
4	A	4	3,9	44	A	3	5,7
5	C	5	4,0	45	C	3	3,5
6	B	4	4,4	46	A	4	3,8
7	B	4	3,6	47	A	2	3,7
8	A	5	4,0	48	A	4	3,9
9	C	3	6,6	49	B	2	4,6
10	B	5	3,1	50	A	7	5,8
11	C	3	3,5				
12	A	5	5,0				
13	C	4	3,9				
14	A	2	4,7				
15	B	5	5,4				
16	C	6	4,0				
17	B	2	5,4				
18	A	5	4,0				
19	A	4	4,5				
20	B	3	4,6				
21	C	6	3,4				
22	A	4	3,2				
23	C	3	5,4				
24	B	3	3,1				
25	B	5	3,4				
26	C	2	4,6				
27	A	2	5,0				
28	B	7	5,9				
29	C	2	4,5				
30	A	4	3,6				
31	A	4	4,3				
32	B	3	3,8				
33	C	3	3,7				
34	C	3	3,7				
35	B	4	6,4				
36	A	5	5,0				
37	C	3	5,1				
38	A	2	3,1				
39	B	5	4,2				
40	B	3	6,2				

Região 2							
Casa n	Tipo de construção	Número de pessoas	Renda	Casa n	Tipo de construção	Número de pessoas	Renda
51	A	5	7,3	102	B	4	9,8
52	B	4	8,2	103	C	3	10,1
53	B	4	8,3	104	C	7	5,4
54	C	4	7,5	105	C	4	6,7
55	A	3	6,2	106	B	4	10,7
56	C	4	8,9	107	C	3	7,1
57	B	3	8,1	108	A	3	8,0
58	B	3	10,8	109	B	4	3,0
59	B	3	5,9	110	A	4	7,7
60	B	5	11,6	111	A	4	7,7
61	A	3	6,8	112	B	5	6,8
62	C	4	8,9	113	B	4	8,8
63	B	4	6,3	114	C	4	8,8
64	B	3	9,0	115	A	4	9,1
65	C	4	9,4	116	C	3	8,7
66	A	4	5,6	117	B	4	5,3
67	C	4	8,3	118	B	4	10,2
68	A	4	5,6	119	B	4	9,4
69	B	4	9,8	120	A	4	5,7
70	B	4	6,4	121	A	3	17,4
71	C	5	7,6	122	B	7	10,4
72	B	4	6,1	123	C	6	16,9
73	B	4	4,8	124	C	8	18,1
74	A	5	8,2	125	A	8	11,1
75	A	3	11,2	126	C	5	9,9
76	B	4	2,2	127	C	4	17,2
77	B	3	10,3	128	B	3	7,6
78	C	4	10,4	129	B	8	14,5
79	B	3	6,3	130	A	4	15,7
80	B	4	8,6	131	C	5	10,7
81	A	4	8,1	132	C	3	12,6
82	C	4	8,2	133	B	3	10,8
83	A	3	8,4	134	C	2	12,7
84	C	4	9,4	135	A	3	6,9
85	B	4	8,7	136	C	7	13,7
86	B	4	6,8	137	C	2	8,3
87	C	3	4,9	138	A	5	7,0
88	A	4	8,7	139	B	6	9,9
89	B	4	6,9	140	C	6	10,7
90	B	3	5,4	141	C	3	11,7
91	B	4	11,1	142	B	3	13,8
92	B	3	10,0	143	A	4	8,2
93	C	3	9,5	144	C	3	12,6
94	A	3	9,6	145	A	6	15,1
95	C	4	6,6	146	B	3	13,9
96	B	4	7,4	147	C	6	16,6
97	B	3	7,3	148	C	4	11,1
98	A	3	8,1	149	B	7	16,1
99	A	4	5,4	150	A	4	11,6
100	A	3	8,6				
101	B	3	5,8				

Região 3							
Casa n	Tipo de construção	Número de pessoas	Renda	Casa n	Tipo de construção	Número de pessoas	Renda
151	C	3	13,8	180	C	5	11,8
152	C	2	10,9	181	C	3	10,5
153	A	2	8,9	182	C	3	10,6
154	C	6	6,4	183	C	3	15,4
155	B	4	17,8	184	C	5	12,2
156	C	6	11,4	185	A	6	10,9
157	C	4	15,1	186	B	2	6,4
158	A	5	15,0	187	C	7	6,1
159	B	4	15,1	188	C	5	15,5
160	C	3	17,2	189	A	7	19,8
161	B	2	15,3	190	C	5	16,8
162	C	4	12,3	191	B	2	12,2
163	A	4	15,3	192	C	2	17,3
164	C	5	8,2	193	B	3	10,3
165	C	8	10,1	194	A	5	16,0
166	B	2	15,9	195	C	3	11,9
167	A	7	11,6	196	C	5	9,6
168	C	6	17,3	197	B	4	17,3
169	A	3	12,2	198	C	2	4,9
170	A	3	6,2	199	A	3	1,0
171	C	3	8,6	200	C	5	13,2
172	C	3	12,9	201	C	2	8,4
173	B	4	11,7	202	A	2	13,5
174	B	3	6,8	203	B	3	13,1
175	A	3	10,7	204	C	4	15,2
176	C	7	6,7	205	C	6	13,6
177	C	5	13,1	206	B	10	15,5
178	B	3	7,9	207	A	6	7,4
179	B	3	9,1	208	C	3	14,6
				209	A	5	8,8
				210	B	6	8,9

6. DISTRIBUIÇÃO BINOMIAL

- 1) Certa doença apresenta taxa de fatalidade de 60%. Em 10 pacientes com esta doença, estabeleça a distribuição de probabilidades da variável $X = \text{n}^\circ$ de óbitos. Nesta distribuição calcule:
- a probabilidade de que todos os pacientes morram;
 - a probabilidade de que nenhum paciente morra;
 - a probabilidade de que 6 pacientes morram;
 - a probabilidade de que pelo menos (no mínimo) 2 pacientes morram;
 - a probabilidade de que pelo menos (no mínimo) 3 pacientes sobrevivam;
 - a probabilidade de que no máximo 4 pacientes morram;
 - a probabilidade de que no máximo 4 pacientes sobrevivam;
 - o número esperado de óbitos;
 - a variabilidade a que estará sujeito o número de óbitos.
- 2) Analisando prontuários de várias maternidades, estimou-se que a probabilidade de um parto ser do tipo cesáreo é de 10% ($p = 0,10$). Em 20 partos de uma maternidade X , calcule:
- a probabilidade de 5 partos serem do tipo cesáreo;
 - a probabilidade de que todos os partos sejam normais;
 - a probabilidade de que no mínimo haja um parto normal;
 - a probabilidade de que no máximo haja um parto tipo cesáreo;
 - o número médio de partos do tipo cesáreo;
 - a variabilidade a que estará sujeito o número de partos do tipo cesáreo.

- 3) Uma doença tem coeficiente de incidência de 10 casos por mil habitantes ao ano. Num ano, admitindo que fossem sorteados 150 habitantes, a probabilidade de que fossem diagnosticados 5 casos seria de:

a) $\frac{150!}{5!145!} \times 0,10^5 \times 0,90^{145}$ b) $5 \times 0,01$ c) $(0,01)^5$
 d) $\frac{150!}{5!145!} \times 0,01^5 \times 0,99^{145}$

- 4) A doença D tem taxa de letalidade de 40% ($p = 0,40$). Para 15 casos desta doença, construa a distribuição de probabilidades da variável $X = n^\circ$ de óbitos por D. Nesta distribuição, calcule:

- a) a probabilidade de que todos os pacientes morram;
 b) a probabilidade de que nenhum paciente morra;
 c) a probabilidade de que dois pacientes morram;
 d) a probabilidade de que no mínimo 2 pacientes morram;
 e) a probabilidade de que no máximo 2 pacientes morram;
 f) a probabilidade de que no mínimo 2 pacientes sobrevivam;
 g) a probabilidade de que no máximo 2 pacientes sobrevivam;
 h) o número esperado de óbitos.

- 5) Em uma área, a doença D tem taxa de incidência estimada de 2 por 1.000 habitantes por ano. Em um dado momento, foram sorteadas 30 pessoas, qual a probabilidade de que 2 pessoas estejam doentes?

a) 0,067
 b) 0,000004
 c) $\frac{30!}{2!(30-2)!} \times (0,02)^2 \times (0,98)^{28} = 0,09883$
 d) $\frac{30!}{28!(30-28)!} \times (0,002)^{28} \times (0,998)^2 = 0,00000$
 e) $\frac{30!}{2!(30-2)!} \times (0,002)^2 \times (0,998)^{28} = 0,00165$

- 6) A eficiência estimada de uma vacina é igual a 60%. Admitindo que fossem sorteadas 30 pessoas vacinadas, qual a probabilidade de que seis pessoas continuassem suscetíveis à doença?

a) $3,68\% = 6 \times 0,6$
 b) $20,0\% = 6:30$
 c) $0,001\% = \frac{30!}{6!24!} \times (0,6)^6 \times (0,4)^{24}$
 d) $1,152\% = \frac{30!}{24!6!} \times (0,6)^{24} \times (0,4)^6$
 e) $4,67\% = (0,6)^6$

7. DISTRIBUIÇÃO NORMAL

- 1) Supondo-se que, em indivíduos sadios, a concentração de albumina no sangue tenha distribuição normal com média igual a 4g/100cc e desvio-padrão de 0,6g/100cc, então, em uma população de indivíduos sadios, calcule:

- a) a probabilidade de se ter indivíduos com concentração de albumina menor do que 3g/100cc;
 b) a probabilidade de se ter indivíduos com concentração de albumina maior do que 4,9/100cc;

- c) a probabilidade de se ter pessoas com concentração de albumina entre 3,2g/100cc e 5,2/100cc;
 d) a concentração de albumina que é ultrapassada por 5% da população;
 e) a menor concentração de albumina que não é alcançada por 10% da população.
- 2) Em uma curva normal reduzida, a área compreendida entre -1,64 e zero vale 45%. Nestas condições, a área à direita do valor +1,64 valerá:
 a) 45% b) Infinito c) 50% d) 100% e) 5%
- 3) Em uma curva normal reduzida, a área compreendida entre -1,96 e zero vale 47,5%. Nestas condições, a área à direita do valor +1,96 valerá:
 a) 45% b) 47,5% c) 95% d) 100% e) 2,5%
- 4) Supondo que a concentração de fósforo no sangue em indivíduos sadios tenha distribuição normal com média igual a 3mg/100cc de soro e desvio-padrão igual a 0,5 mg/100cc, então em uma população de indivíduos sadios, calcule:
 a) a maior concentração de fósforo que é ultrapassada por 32% da população;
 b) a concentração de fósforo que não é alcançada por 10% (10% da população têm valor menor);
 c) a probabilidade de indivíduos com concentração de fósforo menor do que 3,5 mg/100cc;
 d) a probabilidade de se ter indivíduos com concentração de fósforo maior do que 4 mg/100cc;
 e) a probabilidade de se ter indivíduos com concentração de fósforo entre 3,5 e 4 mg/100cc de soro;
 f) a probabilidade de se ter indivíduos com concentração de fósforo exatamente igual a 3 mg/100cc.
- 5) Supondo que a concentração de ácido úrico no plasma em indivíduos sadios tem distribuição normal com média igual a 3,8mg% e desvio-padrão de 0,6mg%, então, em uma população de indivíduos sadios, calcule:
 a) a probabilidade de se ter indivíduos com concentração de ácido úrico menor do que 3,2 mg%;
 b) a probabilidade de se ter indivíduos com concentração de ácido úrico maior do que 5,18 mg%;
 c) a probabilidade de indivíduos com concentração de ácido úrico entre 2,82mg% e 4,78mg%;
 d) a concentração de ácido úrico que 10% da população têm no mínimo tal valor;
 e) a concentração de ácido úrico que não é alcançada por 15% da população;
 f) a probabilidade de indivíduos com concentração de ácido úrico exatamente igual a 3,8 mg%.
- 6) Após um grande levantamento populacional, foi detectado que uma dada variável apresentava uma distribuição normal com média igual a 65 mg e desvio padrão igual a 24 mg. Se dessa população fosse sorteado um número extremamente grande de amostras de tamanho $n = 64$ e nessas amostras fossem calculadas as médias amostrais, esperar-se-ia encontrar (aponte a alternativa correta).
 a) 4,75% das amostras com média abaixo de 70mg;
 b) 4,75% das amostras com média entre 65 e 70mg;
 c) 1,67% das amostras com média igual a 70mg;
 d) 95,25% das amostras com média abaixo de 70mg;
 e) 45,25% das amostras com média abaixo de 70mg.
- 7) Em uma curva normal reduzida, a área compreendida entre -1,0 e +2,0 vale 81,85%. Nestas condições, cada uma das áreas valerá, respectivamente:
 a) 31,74% e 4,56%;
 b) 40,93% e 40,93%;
 c) 68,26% e 95,44%;
 d) 72,67% e 45,49%;
 e) nenhum dos valores apontados.

8. TESTE DE HIPÓTESES

- 1) Com o objetivo de testar uma hipótese H_0 contra a hipótese alternativa H_a , um pesquisador fixou as probabilidades de erros de 1ª e 2ª espécies, respectivamente, em 5% e 10%. Realizado o teste, imaginem-se 2 situações diferentes:
 Em **A**: O pesquisador rejeitou a hipótese de nulidade
 Em **B**: O pesquisador não rejeitou a hipótese de nulidade.

Para cada situação (A e B) assinale a alternativa correta e justifique.

- a) O pesquisador certamente estará cometendo um erro cuja probabilidade de ocorrência é igual a 5%;
- b) O pesquisador certamente estará cometendo um erro cuja probabilidade de ocorrência é igual a 10%;
- c) Se o pesquisador estiver cometendo um erro, a probabilidade de ocorrência associada a este erro é de 15%;
- d) Se o pesquisador estiver cometendo um erro, a probabilidade de ocorrência associada a este erro é de 10%;
- e) Se o pesquisador estiver cometendo um erro, a probabilidade de ocorrência associada a este erro é de 5%.

Teste de uma proporção populacional

- 2) Certa enfermidade, quando tratada pela terapia usual, apresenta 50% de curas, isto é, $E_p=50\%$. Uma nova terapia foi proposta com o intuito de elevar esta proporção de curas. Com o objetivo de testar a nova terapia contra a usual, 15 pacientes (com características semelhantes) foram tratados pela nova terapia. Nestas condições:
- a) Formule as hipóteses correspondentes ao teste.
 - b) Adotando a seguinte regra de decisão: "rejeitar H_0 se na amostra de 15 pacientes houver 12, 13, 14, ou 15 curados", calcule:
 - b1) a probabilidade de rejeitar H_0 se H_0 fosse verdadeira;
 - b2) a probabilidade de aceitar H_0 se H_1 : $E_n = 60\%$ fosse verdadeira;
 - b3) a probabilidade de aceitar H_0 se H_1 : $E_n = 70\%$ fosse verdadeira;
 - b4) a probabilidade de aceitar H_0 se H_1 : $E_n = 80\%$ fosse verdadeira.
 - c) Calcule as probabilidades indicadas no item (b), adotando a seguinte regra de decisão: "rejeitar H_0 se na amostra de 15 pacientes houver 13 ou mais curados".
- 3) Em certa área, os relatórios hospitalares informaram um total de 20.000 nascimentos dos quais 18.000 foram partos normais, isto é, uma proporção de $p=90\%$. Uma nova maternidade foi instalada nesta área e durante um mês ocorreram 20 partos, dos quais 16 foram normais, isto é, uma proporção de 80%. Nestas condições, trabalhando em um nível de significância de 5% (ou mais próximo de 5%) diga se concorda com as autoridades sanitárias que concluíram que na nova maternidade a proporção de partos normais é significativamente menor do que 90%.
- 4) Em certa comunidade durante um período de vários anos, a meningite meningocócica tem apresentado uma taxa de letalidade de 20% para o grupo etário de 20 |-- 45 anos. Em 1999 nessa localidade, a meningite meningocócica se manifestou em 15 indivíduos desse grupo etário. Dos 15 casos (os quais eram semelhantes aos usualmente encontrados na referida localidade), a investigação de rotina verificou que 4 morreram. A fim de saber se a letalidade por meningite meningocócica aumentou em 1999, faça um teste de hipóteses adotando um nível de significância de 1% (ou aproximadamente de 1%).
- 5) A resistência ao resfriado comum em uma dada indústria, durante o inverno, é de $p=0,60$. Foi proposto um tratamento preventivo com a finalidade de aumentar a resistência ao resfriado para $p=0,70$. Então:
- a) formule as hipóteses.

b) fixando $\alpha = 0,05$ (ou valor mais próximo) e admitindo ter sido sorteada uma amostra de tamanho $n=20$, observou-se que 4 operários ficaram resfriados. Nestas condições, qual é a conclusão quanto à eficiência do medicamento?

6) Para se estimar a letalidade da doença B, acompanhou-se uma amostra de 30 doentes durante um ano. Após esse período, cinco deles haviam morrido. Testar a hipótese de que essa letalidade é igual a 20%. Fixe o erro de 1ª espécie em um $\alpha=10\%$.

7) Certa comunidade apresentou em um período de vários anos coeficiente de incidência da doença X de 12 por 10.000 hab.. Em 1999, a incidência foi de 70 casos e a população estimada foi igual a 50.000 habitantes. Nestas condições, ao nível de significância de 1% (ou mais próximo) diga se concorda com as autoridades sanitárias que consideraram a situação dentro do esperado.

8) Desejando-se conhecer o coeficiente de prevalência de determinada doença na cidade A, selecionou-se uma amostra aleatória de 500 pessoas. Nesta amostra detectaram-se 20 doentes. Teste a hipótese de que a prevalência é semelhante à descrita na literatura de 10%. (Fixando $\alpha=5\%$)

9) Estima-se que o medicamento A provoque efeitos colaterais em 55% dos pacientes. Deseja-se testar se uma nova droga B tem menor efeito colateral do que A. Para tanto, foram tratados 50 pacientes com a nova droga B e 20 deles apresentaram efeitos colaterais. Há diferença entre as proporções de pacientes com efeitos colaterais nos dois medicamentos? ($\alpha=1\%$)

10) Sabe-se que na cidade Y, 40% dos homens são obesos. Estudou-se uma amostra de 200 mulheres desta mesma cidade e verificou-se que havia 50 obesas. A prevalência de obesidade entre mulheres é igual à dos homens? ($\alpha=10\%$)

Teste de uma média populacional

11) Admite-se que a quantidade de carne ingerida por pessoa por semana (com renda familiar menor do que 3 salários mínimos e tamanho da família de 5 membros), na região Sudeste, possui distribuição normal com média 600g e desvio padrão 100g. Deseja-se saber se no sub-distrito de Pirituba o consumo médio é menor do que o da região sudeste.

Para isto foi conduzida uma pesquisa, com nível de significância de 5%, cujos valores amostrais de consumo são apresentados a seguir. Elabore de forma completa o teste de hipótese.

Consumo médio semanal (em gramas): 300; 400; 350; 450; 100; 220; 150; 500; 900; 800; 600; 150; 50; 170; 370; 220;

12) A concentração de CO_2 plasmático (vol/100cc) em recém-nascidos sadios tem distribuição normal com média igual a 55 vol/100cc e desvio padrão de 7 vol/100cc. Um pesquisador está interessado em saber se em recém-nascidos portadores de fibrose pulmonar a concentração média de CO_2 plasmático é diferente de 55 vol/100cc. Para tanto ele fez as seguintes especificações:

1) se a concentração média de CO_2 plasmático para portadores de fibrose pulmonar fosse 55 vol/100cc ele estaria disposto a dizer que ela é diferente de 55 vol/100cc em apenas 1% das vezes;

2) se a diferença entre a concentração média para os sadios e para os portadores de fibrose pulmonar fosse de um desvio padrão, ele estaria disposto a dizer que as médias são iguais em apenas 20% das vezes.

Nestas condições:

a) Qual o tamanho mínimo da amostra de recém-nascidos portadores de fibrose pulmonar?

b) Supondo que na amostra, cujo tamanho foi determinado acima, a concentração média seja igual a 58 vol/100cc, qual será a conclusão?

13) Supondo agora, que o pesquisador tivesse motivos para supor que a variabilidade da concentração de CO_2 plasmático para os portadores de fibrose pulmonar fosse diferente de 7

vol/100cc e que na amostra observou desvio padrão (s) de 9vol/100cc, com os valores de α , β e d, realize o teste.

14) Em indivíduos sadios, o consumo renal médio de oxigênio tem distribuição normal com média e desvio padrão iguais a 12 e 1,5 cc/minuto, respectivamente. Um pesquisador interessado em saber se indivíduos com insuficiência cardíaca tinham consumo maior, fixou $\alpha = 5\%$ e $\beta = 10\%$, se a m_E fosse igual a 13,0 cc/minuto. Nestas condições:

- 1) Formule as hipóteses
- 2) Estime o tamanho mínimo da amostra

Supondo que, para o tamanho da amostra obtido em 2, o consumo médio de oxigênio tivesse sido de cc/minuto, qual seria a conclusão?

15) Com dados da tabela A, diga se o nível de proteína médio da população de onde esses pacientes foram retiradas é estatisticamente menor ($\alpha=5\%$) do que o nível médio de proteína no plasma da população de indivíduos sadios, cuja média é 7,0g/100cc.

Tabela A- Distribuição dos pacientes com endocardite sub-aguda bacteriana segundo nível de proteína no plasma sanguíneo (g/100cc). Local X, ano Y.

Proteína no plasma	Número de pacientes
4,1	1
4,6	2
5,3	3
5,7	4
6,0	3
6,8	2
7,6	1
Total	16

16) Em uma amostra aleatória de 50 alunos estimou-se altura média de 165 cm (desvio-padrão de 15 cm). Testar a hipótese de que a média da população de onde esta amostra foi retirada é igual padrão 170 cm (desvio-padrão conhecido = 20 cm) ($\alpha=5\%$).

17) Idem ao exercício 16, porém supondo que a variância da altura na população padrão seja desconhecida. ($\alpha=5\%$)

18) Desejando-se conhecer a média de consumo de carne em uma determinada população, selecionou-se uma amostra aleatória de 100 pessoas. Os resultados mostraram que, em média, os indivíduos consumiam 1.000 g/mês (desvio-padrão de 625 g). Teste a hipótese de que o consumo médio dessa população está de acordo com o esperado, que é 1.200 g/mês. ($\alpha=10\%$)

19) Deseja-se saber se o número médio de atendimentos diários no Posto de Saúde Z é menor do que o número médio diário dos postos da rede municipal (igual a 40 atendimentos). Coletou-se, então, a informação dos últimos 20 dias de atendimento e verificou-se que a média diária foi de 30 pacientes e desvio-padrão de 10 pacientes. Verifique se o número médio diário de atendimentos no Posto de Saúde Z é igual à média diária dos postos da rede municipal ($\alpha=2\%$)

Teste de associação

20) Faça o teste de associação, em um nível de significância de 5%, utilizando os resultados obtidos na 4ª unidade.

21) Os dados a seguir referem-se ao trabalho "Análise de sobrevivência em pacientes com lúpus eritematoso sistêmico". (LATORRE, L.C. - Tese apresentada à Faculdade de Saúde Pública da

USP para a obtenção do título de Doutor - 1997). Faça o teste de associação para as tabelas a seguir:

Tabela 21.a Número de pacientes com LES, segundo ocorrência de óbito e presença de envolvimento cardiopulmonar. Clínica de Reumatologia do Hospital Heliópolis. São Paulo - 1978 a 1995.

Envolvimento cardiopulmonar	Óbito		Total
	Sim	Não	
Não	16	147	163
Pulmão	10	38	48
Coração	3	15	18
Pulmão+coração	8	16	24
Total	37	216	253

Tabela 21.b - Número de pacientes com LES, segundo ocorrência de óbito e presença de hipertensão arterial (HA). Clínica de Reumatologia do Hospital Heliópolis. São Paulo - 1978 a 1995.

HÁ	Óbito		Total
	Sim	Não	
Sim	10	62	72
Não	27	154	181
Total	37	216	253

9. INTERVALO DE CONFIANÇA

- 1) Em uma amostra de 16 gestantes com diagnóstico clínico de pré-eclâmpsia, a taxa média de ácido úrico no plasma foi de 5,3 mg. Em gestantes normais a variabilidade a que está sujeita a taxa de ácido úrico no plasma é de 0,60 mg.
 - a) Estime, com 95% de confiança, a taxa média de ácido úrico no plasma da população de gestantes com diagnóstico de pré-eclâmpsia.
 - b) Que pressuposições foram necessárias para a estimação do item a?
- 2) Em certa área, baseando-se na amostra de 100 recém-nascidos, a letalidade por diarreia do recém-nascido no verão e outono foi de 40%. Nestas condições, estime com 99% de confiança, a verdadeira letalidade por diarreia.
- 3) Desejando-se estimar a eficiência de uma droga, uma amostra de 100 pacientes foi sorteada. Supondo-se que tivessem sido observados:
 - a) 10 curados
 - b) 90 curados
 Quais seriam, respectivamente, os intervalos com 95% de confiança para a verdadeira eficiência da droga utilizada.
- 4) Com o intuito de estudar o conteúdo de ácido láctico no sangue de indivíduos com demência precoce, uma amostra de 16 pacientes foi sorteada e os resultados foram: média (\bar{x}) = 13mg/100 cc e desvio padrão (s) = 4,6 mg/100 cc. Estime, através de um intervalo de 98% de confiança, a taxa média de ácido láctico no universo dos pacientes com demência precoce.
- 5) Com a finalidade de estudar o efeito da aplicação de hormônios gonadotróficos, no tratamento de gestantes diabéticas, 60 delas receberam tal tratamento e, destas, 15 gestantes tiveram, como produto da concepção, um natimorto. Estabeleça, a partir deste resultado, o

intervalo com 95% de confiança para a verdadeira proporção de nascidos vivos, para medir a eficiência do tratamento.

6) Em uma amostra aleatória de 50 alunos, a altura média foi de 165cm e (desvio-padrão de 15cm). Construir o intervalo de confiança para estimar a média da altura na população da qual foi retirada a amostra, com 90% de confiança.

7) Deseja-se conhecer a média de consumo de carne na população; então, selecionou-se amostra aleatória de 100 pessoas. Como resultado, os indivíduos consumiam em média 1.000g/mês (desvio-padrão de 625g). Obtenha o intervalo de para o consumo médio populacional com 95% de confiança.

8). Em um estudo sobre o número de atendimentos em um hospital, encontrou-se, durante um ano, o número médio de atendimentos mês igual a 500 pacientes (desvio-padrão de 100 pacientes). Obtenha o intervalo de confiança para o número médio de atendimentos populacional com 98% de confiança.

9) Desejando-se conhecer a prevalência de determinada doença na cidade A, selecionou-se uma amostra aleatória de 500 pessoas. Nesta amostra foram detectados 20 doentes. Apresente a prevalência e o respectivo intervalo de confiança (95%).

10) Para se determinar a letalidade da doença B, observou-se uma amostra de 30 doentes durante um ano. Ocorreram cinco mortes. Estime a letalidade da doença B e calcule o intervalo de confiança (90%).

11) Desejando-se estimar a proporção de obesos em um população, coletou-se uma amostra de 700 pessoas, sendo 350 homens. Destes, 130 homens e 70 mulheres eram obesos. Estime a proporção de obesos para cada um dos sexos e calcule os respectivos intervalos de confiança (80%).

12) Os neonatologistas da Maternidade A estão interessados em conhecer a prevalência de recém-nascidos com baixo peso ao nascer. Para tanto estudaram todos os nascimentos ocorridos durante o mês de janeiro. Observaram que 200 dos 1.200 nascimentos pesaram menos de 2500 g. Estimar a prevalência de baixo peso ao nascer e calcular o intervalo de 95% de confiança.