Exercícios

Estatística Básica

Bárbara Pasini, Filipe Zabala, Hélio Bittencourt, Rossana Benites, Sérgio Kato

2024-11-18

Capítulo 2 Estatística Descritiva

2.1 (R/F) Classifique as variáveis abaixo (qualitativa nominal/ordinal, quantitativa discreta/contínua).

a) Tempo para fazer um teste.
b) Número de alunos aprovados por turma.
c) Nível socioeconômico.
d) Sexo.
e) Gastos com alimentação.
f) Opinião com relação à pena de morte.
g) Religião.
h) Valor de um imóvel.
i) Conceitos em certa disciplina.
j) Número de geladeiras em casa.
k) Temperaturas da água da piscina em um dia de verão.
l) Número de suicídios em uma cidade no decorrer do ano passado.
m) Concentração de chumbo em uma amostra de água.
n) Lista de editoras de livros.
o) Grau de satisfação dos clientes que frequentam uma rinha de galo.
p) Marcas de amaciantes para roupas.
q) Tempo que um paciente sobrevive após determinado diagnóstico.
r) Participação de mercado (*market share*).
s) Classificação em uma corrida de banheiras.
t) Tempo final de cada corredor.
u) Lista dos nomes das banheiras participantes, tal como Dick Vigarista e Trollface.
v) Distância de Estambul ao Rio de Janeiro.
w) Preço de um automóvel.
x) Nível de dificuldade de uma questão de prova.
y) Número de consultas ao celular durante a aula.

2.2 (R) Os dados a seguir referem-se ao número de livros adquiridos, no ano passado, pelos 40 alunos da Turma A.

4 2 1 0 3 1 2 0 2 1 0 2 1 1 0 4 3 2 3 5 8 0 1 6 5 3 2 1 6 4 3 4 3 2 1 0 2 1 0 3

a) Classifique a variável.
b) Quantos livros foram adquiridos pelos 40 alunos?
c) Organize os dados em uma tabela adequada.
d) Qual a proporção de alunos que adquiriram menos do que 3 livros?
e) Qual a proporção de alunos que adquiriram pelo menos 4 livros?

2.3 (F) Em uma fábrica retirou-se uma amostra de 50 peças de um lote de certo material e contou-se o número de defeitos em cada peça, apresentados na tabela a seguir.

i	**# defeitos**	f_i	fr_i	F_i	Fr_i
1	0	17			
2	1	10			
3	2				
4	3	8			
5	4	5			
6	5	1			
Total	-	**50**		-	-

a) Classifique a variável 'número de defeitos'.
b) Qual a frequência absoluta da classe 3? Interprete.
c) Qual a frequência relativa da classe 3? Interprete.
d) Qual a frequência acumulada da classe 4? Interprete.
e) Qual a frequência acumulada relativa da classe 5? Interprete.
f) Represente os dados utilizando o gráfico que você considerar mais adequado.

2.4 (F) Em 13 de março de 1883, estavam Émile Durkheim e Max Weber no leito de morte de Karl Marx discutindo a respeito de propriedade intelectual. Weber, o mais jovem e disposto da turma, com apenas 19 anos, coletou algumas informações a respeito da Convenção de Paris de 1883, que aconteceria em uma semana. Em suas anotações, estava o número de unidades monetárias que deveria ser paga anualmente por cada país membro do tratado, dependendo da classe à qual o país pertencesse[1]. O valor da unidade iria variar de acordo com a inflação e outros fatores econômicos da época corrente. A tabela abaixo apresenta o resultado dos estudos de Weber.

Classe	**Unidades**	f_i	f_{r_i}	F_i	F_{r_i}
I	25	21			
II	20	26			
III	15	10			
IV	10	9			
V	5	32			
VI	3	38			
VII	1	37			
Total	-	**173**		-	-

a) Qual a frequência simples da classe VI? Interprete.
b) Qual a frequência relativa da classe I? Interprete.
c) Qual a frequência acumulada da classe II? Interprete.
d) Qual a frequência acumulada relativa da classe III? Interprete.
e) Represente os dados utilizando o gráfico que você considerar mais adequado.

2.5 (F) Considere o conjunto de dados $10, -4, 5, 7, 1, 3, 9$.

a) Obtenha o rol.
b) Indique e interprete $x_{(4)}$.

2.6 (Adaptado de (Pagano 2004)) Em uma investigação dos fatores de risco para doenças cardiovasculares, os níveis de cotinina (produto metabólico da nicotina) foram registrados para um grupo de fumantes f) e um grupo de não fumantes (NF) em nanogramas[2] por mililitro (ng/mL). As distribuições de frequência estão na tabela abaixo.

[1] *Paris Convention for the Protection of Industrial Property* (World Intellectual Property Organization 1883).
[2] Um nanograma é um bilionésimo de grama, i.e., $1g = 10^9 ng \therefore 1ng = 10^{-9} g$.

Nível de cotinina (ng/mL)	f_F	fr_F	f_{NF}	fr_{NF}
$0 \vdash 14$	78		3300	
$14 \vdash 50$	133		72	
$50 \vdash 100$	142		23	
$100 \vdash 150$	206		15	
$150 \vdash 200$	197		7	
$200 \vdash 250$	220		8	
$250 \vdash 300$	151		9	
300 +	412		11	
Total	**1539**		**3445**	

Complete a tabela acima e responda:

a) Qual a proporção de fumantes com nível de cotinina até 14 ng/mL? E de não fumantes?
b) Qual a proporção de fumantes que possuem 100 ng/mL ou mais de cotinina?
c) Entre os não fumantes, qual a proporção de pessoas que tem entre 100 e 250 ng/mL de cotinina?
d) Qual o intervalo modal entre os fumantes? E entre os não fumantes?
e) Qual intervalo mediano do nível de cotinina para os fumantes? E para os não fumantes?
f) Represente os dados utilizando o gráfico que você considerar mais adequado.

2.7 (R) Os dados da tabela abaixo se referem aos preços (em reais) de uma amostra de produtos de uma loja de departamento. Determine a proporção de preços inferiores a R$ 300.00.

i	**Preços (R$)**	**# transações**
1	$0 \vdash 100$	8
2	$100 \vdash 200$	10
3	$200 \vdash 300$	22
4	$300 \vdash 400$	25
5	$400 \vdash 500$	35
6	$500 \vdash 600$	10

2.8 (R) A altura de 60 alunos da PUCRS foi registrada em cm.

174 170 156 168 176 178 162 182 172 168 166 156 169 168 162 160 163 168 162 172

168 167 170 153 171 166 168 156 160 172 173 163 170 175 176 182 158 176 161 175

173 163 172 167 170 179 179 170 151 175 152 151 172 173 170 174 167 167 158 174

a) Construa uma distribuição de frequência (fechada no limite inferior) com 8 classes de amplitudes iguais, adotando 150 cm como limite inferior da distribuição.
b) Qual a proporção de alunos com altura mínima de 166 cm? Como você simboliza esta medida?
c) Quantos alunos têm menos de 162 cm?

2.9 (R) O RH de certa empresa fez um levantamento sobre salários mensais (em salários mínimos) a partir de uma amostra de 130 funcionários do setor de produção, obtendo os resultados a seguir.

a) Complete a tabela.
b) Que proporção das pessoas recebe pelo menos 12 salários mínimos?

Salário (sm)	**Freq. Acumulada (F)**
$4 \vdash 8$	50
$8 \vdash 12$	90
$12 \vdash 16$	110
$16 \vdash 20$	130

2.10 (Adaptado de (Picarelli c. 1970)) Considere a renda anual de certa região, conforme dados disponíveis em formatos xlsx[3] e tsv[4].

a) Construa a distribuição de frequências, determinando o nº. de classes de mesma amplitude conforme a fórmula de Sturges.
b) Obtenha os elementos característicos da distribuição decorrente.
c) Interprete o significado dos seguintes elementos característicos:
   c1) frequência simples da 5ª classe.
   c2) frequência relativa da 5ª classe.
   c3) frequência acumulada da 5ª classe.
   c4) frequência relativa acumulada da 5ª classe.
   c5) ponto médio da 5ª classe.
   c6) limite inferior e superior da 5ª classe.
d) Construa os seguintes gráficos:
   d1) histograma de frequências simples.
   d2) histograma de frequências relativas.

2.11 (Adaptado de (Picarelli c. 1970)) Considere o número de dependentes por empregado da empresa XYZ, conforme dados disponíveis em formatos xlsx[5] e tsv[6].

a) Construa a distribuição de frequências.
b) Obtenha os elementos característicos da distribuição decorrente.
c) Interprete o significado dos seguintes elementos característicos: c1) frequência simples do 4^o valor ou ponto. c2) frequência acumulada do 5^o valor ou ponto. c3) frequência relativa do 5^o valor ou ponto. c4) frequência relativa acumulada do 3^o valor ou ponto.
d) Construa os seguintes gráficos: d1) em linha simples, utilizando as frequências simples. d2) em linha simples, utilizando as frequências relativas. d3) histograma de frequências simples. d4) histograma de frequências relativas.

2.12 (Adaptado de (Picarelli c. 1970)) Considere os salários pagos pela empresa ABC, conforme dados disponíveis em formatos xlsx[7] e tsv[8].

a) Construa a distribuição de frequências por classes de valores correspondente, adotando como limite inferior o valor zero e como superior o valor 14, fazendo a amplitude dos intervalos de classe constante e igual a duas unidades monetárias.
b) Calcule os valores dos elementos característicos da distribuição decorrente.
c) Interprete o significado dos seguintes elementos característicos:
   c1) frequência simples da segunda classe.
   c2) frequência relativa da quinta classe.
   c3) frequência acumulada da terceira classe.
   c4) frequência relativa acumulada da quarta classe.
   c5) ponto médio da sexta classe.
   c6) limite inferior e superior da sexta classe.

[3] https://filipezabala.com/data/renda.xlsx
[4] https://filipezabala.com/data/renda.tsv
[5] https://filipezabala.com/data/dependentes.xlsx
[6] https://filipezabala.com/data/dependentes.tsv
[7] https://filipezabala.com/data/salario.xlsx
[8] https://filipezabala.com/data/salario.tsv

d) Construa os seguintes gráficos:
d1) histograma de frequências simples.
d2) histograma de frequências relativas.

2.13 (R) O revisor de um jornal fez durante um mês o levantamento dos erros ortográficos encontrados no editorial do jornal. Os resultados encontrados foram

0 1 0 1 0 0 0 0 2 3 0 1 2 3 4 0 0 0 1 4 1 1 0 0 3 5 1 0 0 1

a) Faça uma distribuição de frequência dos dados.
b) Calcule as medidas de posição.
c) Calcule as medidas de dispersão.

2.14 (R) As informações abaixo indicam o número de acidentes ocorridos com 70 motoristas de uma empresa de ônibus nos últimos 5 anos.

# acidentes	0	1	2	3	4	5	6	7
# motoristas	15	11	20	9	6	5	3	1

a) Obtenha o número total de acidentes ocorridos no período.
b) Determine o número de motoristas com menos de 1 acidente.
c) Determine a proporção de motoristas com pelo menos 3 acidentes.
d) Determine a proporção de motoristas com no máximo 2 acidentes.
e) Obtenha a média, a moda e a mediana de acidentes. Interprete.
f) Obtenha a variância e o desvio padrão de acidentes. Interprete.
g) Obtenha a amplitude interquartílica de acidentes. Interprete.

2.15 (F) Você está dirigindo em uma rodovia e observa que ultrapassa o mesmo número de automóveis que ultrapassam você. Qual medida de tendência central melhor representa a velocidade que você está dirigindo, em comparação com os demais veículos? Explique o seu raciocínio.

2.16 (R) Considere as taxas de juros observadas para dez ações.

Ação	1	2	3	4	5	6	7	8	9	10
Taxa %	2.59	2.64	2.60	2.62	2.55	2.61	2.50	2.63	2.64	2.69

Calcule as seguintes medidas, apresentando o resultado final com 4 casas decimais.

a) a taxa média.
b) a taxa mediana.
c) a taxa modal.
d) o desvio padrão das taxas.
e) o coeficiente de variação das taxas.

2.17 (F) Considerando a Circular nº 2957 do Banco Central, como você calcularia as taxas médias do mercado?

2.18 (R) Em certa empresa trabalham 4 analistas de mercado, 2 supervisores, 1 chefe de seção e 1 gerente. Suas remunerações são, respectivamente, R$1,300.00; R$1,600.00; R$2,750,00 R$5,000.00.

a) Calcule a média e o desvio padrão do salário destes funcionários.
b) Obtenha a mediana e a amplitude interquartílica dos salários.
c) Se cada funcionário em questão recebeu aumento salarial de 10%, qual será o valor do salário médio desses funcionários? E o desvio padrão?

2.19 (R) Um estudo foi realizado por um professor em três turmas, obtendo a média e o desvio padrão das notas de sua disciplina, conforme tabela a seguir. Qual a turma com menor variabilidade? Justifique adequadamente.

Turma	A	B	C
Média	6.5	8.0	8.0
Desvio padrão	2.2	1.7	2.0

2.20 (R/F) Uma amostra de pessoas que trabalham com remuneração diária é formada por subgrupos com as seguintes características:

15 recebem R$ 45,00/h
15 recebem R$ 50,00/h
10 recebem R$ 60,00/h
10 recebem R$ 90,00/h
10 recebem R$ 120,00/h

Obtenha:

a) Média, mediana e moda.
b) Os percentis 25% e 75%.
c) Variância e desvio padrão amostrais.
d) Amplitude interquartílica.
e) Desvio absoluto mediano.

2.21 (R) O tempo de espera em minutos até o atendimento de uma amostra de clientes foi o seguinte:

Cliente	Tempo (min.)
A	15.0
B	10.0
C	5.0
D	15.0
E	10.5
F	15.0
G	21.5

a) Identifique e classifique a variável de interesse.
b) Determine o tempo médio de espera.
c) Determine a moda do tempo espera.

2.22 (R) Um banco quer verificar o tempo necessário para aprovar uma linha de crédito habitacional, medido a partir do momento da solicitação do crédito. Uma amostra de pedidos de crédito foi selecionada, determinando-se o número de dias úteis necessários para aprovar cada pedido conforme lista a seguir.

18 16 17 12 14 16 16 21 18 11 16 18

Um segundo banco apresentou, de acordo com uma amostra, média de 13 dias para aprovação de crédito e desvio padrão 4. Em qual dos bancos existe menor variabilidade no número de dias necessários para aprovação de crédito? Justifique.

2.23 (R/F) Os operários de um setor industrial têm salário médio de 8 salários mínimos (sm) e desvio padrão de 2.3 sm. Um acordo coletivo prevê um aumento de 60%, mais uma parte fixa correspondente a 1.5 sm. Calcule a média e o desvio padrão dos salários após o acordo.

2.24 (F) Foram observadas as quantidades de fotocópias feitas por dois setores de uma empresa no segundo semestre de certo ano, apresentadas na tabela abaixo.

Mês	**jan**	**fev**	**mar**	**abr**	**mai**	**jun**
Setor X	30	15	15	10	39	35

Mês	**jan**	**fev**	**mar**	**abr**	**mai**	**jun**
Setor Y	120	160	15	130	145	300

Sabendo que $\sum_{i=1}^{6} x_i = 144$, $\sum_{i=1}^{6} x_i^2 = 4196$, $\sum_{i=1}^{6} y_i = 870$, $\sum_{i=1}^{6} y_i^2 = 168150$, calcule:

a) A média, mediana e moda do número de cópias de cada setor.
b) A amplitude, a variância e o desvio padrão amostrais do número de cópias de cada setor.
c) Em qual dos setores o número de cópias variou mais? Por quê?
d) Represente os dados utilizando o gráfico que você considerar mais adequado.

2.25 (F) Você não sabe a nota da primeira prova (P_1) de três avaliações realizadas no semestre passado. Sabendo que a média das notas das três avaliações foi $\bar{P} = 7.5$ e que $P_2 = 5$ e $P_3 = 9$, qual a nota da P_1?

2.26 (F) Para aprimorar seu chimarrão, você decidiu medir a temperatura da água que estava utilizando durante as duas últimas semanas. As temperaturas (em graus Celsius) foram observadas conforme tabela a seguir.

Semana 1	72.4	84.9	57.5	61.0	87.9	78.1	73.0
Semana 2	76.3	80.0	74.1	67.0	83.2	83.0	58.0

a) Calcule a média e mediana da temperatura da água nos 14 dias.
b) Repita o item a) para semana 1 e para semana 2 separadamente. Parece haver diferença na temperatura de uma semana para outra?
c) Uma embalagem de erva mate aponta a temperatura 75 graus Celsius como ideal, sendo considerado bom o chimarrão com água entre 65 e 85 graus. Acima deste intervalo o mate está quente demais (pelando) e abaixo é considerado frio. Com essa informação, monte uma tabela de frequência para observar quantas vezes nessas duas semanas o chimarrão ficou frio, bom ou pelando para as semanas 1, 2 e durante os 14 dias.
d) A frequência em que o chimarrão estava na temperatura ideal foi diferente nas duas semanas? Comente os resultados, identificando a diferença de uma semana para outra.

2.27 (Adaptado de (Anderson et al. 2007)) Milhões de norte-americanos levantam de manhã e realizam seu trabalho em escritórios residenciais, comunicando-se com a empresa por meios eletrônicos. Coletou-se uma amostra da idade de 20 indivíduos que trabalham em casa. As idades foram as seguintes:

18	54	20	46	25	48	53	27	26	37
40	36	42	25	27	33	28	40	45	25

a) Calcule a média, mediana e moda.
b) Calcule e interprete o primeiro e o terceiro quartis.
c) Se a idade mediana do universo de todos os adultos é 35.5 anos, comente se as pessoas que trabalham em casa tendem a ser mais jovens ou mais velhas que a população de todos os adultos.

2.28 (Adaptado de (Anderson et al. 2007)) Em um teste automobilístico de quilometragem e consumo de gasolina, 13 automóveis foram testados em um percurso de 482.8 quilômetros, em condições de dirigibilidade tanto na cidade (X) quanto na rodovia (Y). Os dados apresentados a seguir foram registrados para o desempenho obtido em termos de quilômetros por galão americano (*US liquid gallon*), equivalente a 3.78 litros.

X	26.07	26.81	25.58	23.17	21.24	24.62	27.03	25.74	25.91	24.62	24.46	24.62	25.74
Y	30.57	32.18	28.96	29.93	30.89	27.35	27.35	28.96	30.57	33.95	31.22	28.96	28.96

Considere $\sum_{i=1}^{13} x_i = 325.61$, $\sum_{i=1}^{13} x_i^2 = 8184.513$, $\sum_{i=1}^{13} y_i = 389.85$, $\sum_{i=1}^{13} y_i^2 = 11732.66$.

a) Calcule a média, mediana e a moda.
b) Faça uma investigação sobre a possibilidade de diferença no consumo de combustível quando de dirige na cidade e na rodovia.

2.29 (Adaptado de (Pagano 2004)) Um estudo foi conduzido comparando mulheres adolescentes bulímicas e não bulímicas, com composição corporal e níveis de atividade física similares. A seguir estão as medidas de consumo calórico, em calorias por quilograma, para amostras de cada grupo.

Bulímicas	15.9	18.9	25.1	16.0	19.6	16.5	21.5	25.6	17.0	17.6	18.1	18.9
Saudáveis	20.7	30.6	22.4	33.2	24.5	33.7	37.1	36.6	26.3	37.4	40.8	37.4

a) Obtenha o consumo calórico médio e mediano por grupo. Eles parecem diferir?
b) Calcule o desvio padrão amostral, a amplitude interquartílica por grupo. Eles parecem diferir?
c) Qual grupo tem maior variabilidade nas medidas? Justifique.

2.30 (F) O Mini-Exame do Estado Mental (MEEM)[9] é um teste rápido para avaliar a função cognitiva. O escore do MEEM pode variar de um mínimo de 0 ponto, indicando o maior grau de comprometimento cognitivo dos indivíduos, até um total máximo de 30 pontos, que corresponde à melhor capacidade cognitiva. A pontuação assume os valores 0, 1, …, 30, de onde calculam-se medidas como média e variância para avaliação dos pacientes.

Foram avaliados dois grupos de pacientes em relação ao MEEM, conforme a tabela abaixo.

i	1	2	3	4	5	6	7	8
MEEM G_1 (x_i)	12	19	12	17	18	12	10	11
MEEM G_2 (y_i)	30	22	27	21	19	18	19	21

Considere $\sum_{i=1}^{8} x_i = 111$, $\sum_{i=1}^{8} x_i^2 = 1627$, $\sum_{i=1}^{8} y_i = 177$, $\sum_{i=1}^{8} y_i^2 = 4041$.

a) Classifique a variável MEEM.
b) Calcule a média, a mediana e a(s) moda(s) do MEEM de cada grupo.
c) Calcule a amplitude e a amplitude interquartílica de cada grupo.
d) Calcule a variância e o desvio padrão amostrais de cada grupo.
e) Em qual dos grupos o MEEM variou mais? Justifique.

2.31 (Adaptado de (Magalhães and Lima 2002)) O Conselho Regional de Odontologia recomenda visitas periódicas ao dentista e, para orientar sua campanha de divulgação, realizou uma pesquisa com 100 crianças. O número médio de visitas no último ano foi 0.5. A mediana e a moda do número de visitas foram iguais a zero. Com base nestes dados, responda V para verdadeiro, F para falso (corrigindo o que estiver errado) e NA para sentenças que não se pode afirmar através das informações fornecidas.

( ) Pelo menos 50 crianças não visitaram o dentista neste ano.
( ) Alguma criança fez mais de três visitas no último ano.
( ) Metade das crianças já foi ao dentista pelo menos uma vez.
( ) Uma criança pode ter ido ao dentista 80 vezes no ano.

2.32 (Adaptado de (Picarelli c. 1970)) Considere os dados de produção do produto A em um certo ano.

Mês	**Toneladas**
janeiro	20
fevereiro	16
março	22
abril	18

[9] https://aps.bvs.br/apps/calculadoras/?page=11

Mês	**Toneladas**
maio	22
junho	24
julho	28
agosto	26
setembro	28
outubro	30
novembro	32
dezembro	34

Obtenha:

a) média aritmética, ilustrando a propriedade de que a soma algébrica dos desvios em relação à média aritmética é nula.
b) mediana e moda.
c) variância e desvio padrão amostrais.
d) coeficiente de variação.
e) amplitude, amplitude interquartílica e desvio absoluto mediano.

2.33 (Adaptado de (Picarelli c. 1970)) Considere os dados de produção de uma mercadoria conforme tabela a seguir.

Ano	**Toneladas**
2017	6
2018	8
2019	7
2020	10
2021	9
2022	12
2023	11

Obtenha:

a) média, mediana e moda.
b) variância e desvio padrão amostrais.
c) coeficiente de variação.
d) amplitude, amplitude interquartílica e desvio absoluto mediano.

2.34 (F) Foram medidas as alturas de 100 alunos de certa disciplina, apresentadas na tabela a seguir.

i	**Altura (cm)**	f_i	f_{r_i}	F_i	F_{r_i}	$\Finv_i$	$\Finv_{r_i}$
1	$140 \vdash 150$	2					
2	$150 \vdash 160$	13					
3	$160 \vdash 170$						
4	$170 \vdash 180$	47					
5	$180 \vdash 190$	8					

a) Classifique a variável ‘altura’.
b) Qual a frequência relativa da classe 3? Interprete.
c) Qual a frequência acumulada da classe 4? Interprete.
d) Qual a frequência acumulada relativa da classe 2? Interprete.
e) Quantos alunos têm pelo menos 1.60m?

f) Represente os dados utilizando o gráfico que você considerar mais adequado.
g) Estime as medidas da questão anterior.

2.35 (Adaptado de (Picarelli c. 1970)) Considere a tabela a seguir.

N^o de Dependentes	**N^o de Empregados**
0	10
1	18
2	30
3	10
4	8
5	4
$\sum$	**80**

Obtenha:

a) média, mediana e moda.
b) variância e desvio padrão amostrais.
c) coeficiente de variação.
d) amplitude, amplitude interquartílica e desvio absoluto mediano.

2.36 (Adaptado de (Picarelli c. 1970)) Considere a tabela a seguir.

Unidades Monetárias	**N^o de Pessoas**
$0 \vdash 2$	9
$2 \vdash 4$	13
$4 \vdash 6$	16
$6 \vdash 8$	45
$8 \vdash 10$	56
$10 \vdash 12$	35
$12 \vdash 14$	14
$14 \vdash 16$	9
$16 \vdash 18$	3

Estime:

a) média, mediana e moda.
b) variância e desvio padrão amostrais.
c) coeficiente de variação.
d) amplitude, amplitude interquartílica e desvio absoluto mediano.

2.37 (F) Considere as variáveis `children` (número de filhos) e `height` (altura) do conjunto de dados disponível em formatos xlsx[10] e tsv[11].

Para cada variável, obtenha:

a) distribuição de frequências.
b) mínimo e máximo.
c) média aritmética.
d) média quadrática.
e) média harmônica.
f) média geométrica.

[10] https://filipezabala.com/data/hospital.xlsx
[11] https://filipezabala.com/data/hospital.tsv

g) compare as médias dos itens c, d, e, f.
h) mediana.
i) moda.
j) primeiro e terceiro quartis.
k) amplitude.
l) variância.
m) desvio padrão.
n) coeficiente de variação.
o) amplitude interquartílica.
p) desvio absoluto mediano.
q) a estimativa do total de filhos em um grupo de 1500 mulheres.

2.38 (Adaptado de (Picarelli c. 1970)) Considere as seguintes informações de três conjuntos de dados.

A	**B**	**C**
$N = 200$	$N = 50$	$\mu = 8$
$\sum f_i x_i = 4000$	$\sum f_i x_i = 500$	$\sum f_i x_i = 3200$
$\sum f_i (x_i - \mu)^2 = 5000$	$\sum f_i x_i^2 = 5450$	$\sum f_i x_i^2 = 32000$

Preencha a tabela com os indicadores solicitados, apontando as distribuições de maior e menor homogeneidade.

Indicador	**A**	**B**	**C**
Média Aritmética			
Variância			
Desvio padrão			
Coef. de variação			

2.39 (Adaptado de (Picarelli c. 1970)) Considere as seguintes informações de quatro conjuntos de dados.

A	**B**	**C**	**D**
$\mu = 8$	$\mu = 100$	$N = 100$	$\mu = 50$
$\sigma^2 = 4$	$\sigma^2 = 121$	$\sum x_i = 5000$	$\sum x_i = 10000$
		$\sum x_i^2 = 256400$	$\sum (x_i - \mu)^2 = 7200$

Preencha a tabela com os indicadores solicitados, apontando as distribuições de maior e menor homogeneidade.

Indicador	**A**	**B**	**C**	**D**
Média Aritmética				
Variância				
Desvio padrão				
Coef. de variação				

2.40 (Adaptado de (Picarelli c. 1970)) Seja que a sequência de variáveis aleatórias $X_1, \ldots, X_n$, independentes e identicamente distribuídas (iid) com $\mu = 10$ e $\sigma^2 = 16$. Determine a média aritmética e a variância de Y.

a) $Y = X + 2$
b) $Y = 2X$
c) $Y = \frac{X}{2} + 4$
d) $Y = \frac{X}{2} - 4$

Questões de concursos

2.41 (Anpec 1991, EST-Q2) Com respeito às distribuições de freqüência pode-se afirmar que:

(0) É sempre verdade que a média está localizada entre a mediana e a moda.
(1) O coeficiente de variação, σ/μ é sempre maior do que um.
(2) A mediana é menos sensível que a média a valores extremos (ou discrepantes).
(3) O coeficiente de correlação é uma medida independente de escala.
(4) A diferença entre o terceiro e o primeiro quartil, chamada de intervalo interquartil, é uma medida de dispersão.

2.42 (Anpec 1992, EST-Q1) Para um conjunto qualquer de dados pode-se afirmar que:

(0) A média geométrica é maior do que a média aritmética que é maior do que a média harmônica.
(1) Se a média e a mediana desse conjunto de dados forem respectivamente 10 e 12 pode-se dizer que essa distribuição apresenta cauda à esquerda.
(2) O coeficiente de variação pode ser perfeitamente substituído pelo desvio padrão.
(3) A variância é uma estatística que independe da unidade de medida.
(4) A média é quem melhor representa um conjunto de dados pois ela é a única medida de tendência central que leva em consideração todas as observações existentes.

2.43 (Prova 1 TRF 2005) Considere a seguinte distribuição das frequências absolutas dos salários mensais, em reais, referentes a 200 trabalhadores de uma indústria.

i	Classes de Salários	f_i
1	$400 \vdash 500$	50
2	$500 \vdash 600$	70
3	$600 \vdash 700$	40
4	$700 \vdash 800$	30
5	$800 \vdash 900$	10

Sobre essa distribuição de salários é correto afirmar que:

a) O salário modal encontra-se na classe de R$ 800 até R$ 900.
b) O salário mediano encontra-se na classe de R$ 600 até R$ 700.
c) O salário modal encontra-se na classe de R$ 600 até R$ 700.
d) O salário modal encontra-se na classe de R$ 700 até R$ 800.
e) O salário mediano encontra-se na classe de R$ 500 até R$ 600.

2.44 (Exame Fundação Médica do Rio Grande do Sul - 2010) Considere uma amostra de 250 pessoas que sofreram acidentes ofídicos (picada de cobra). O resumo dos dados está nas tabelas abaixo. Para cada questão (1 e 2) existe somente uma alternativa correta.

Idade	# pessoas
$6 \vdash 8$	1
$8 \vdash 10$	2
$10 \vdash 12$	7
$12 \vdash 14$	14
$14 \vdash 16$	31
$16 \vdash 18$	44
$18 \vdash 20$	72
$20 \vdash 22$	61
$22 \vdash 24$	18
Total	**250**

Medida	Valor
Moda	19
Média	17.8
Mediana	18
Primeiro quartil	16
Segundo quartil	20
Desvio padrão	3

Questão 1 A frequência relativa de pessoas com idade maior ou igual a 12 anos e menor que 18 anos que sofreram acidentes ofídicos é igual a:

a) 35.6%
b) 38.4%
c) 39.6%
d) 58.8%
e) 64.4%

Questão 2 Analise as alternativas abaixo:

I) Metade das pessoas da amostra apresentou idade menor ou igual a 18 anos.
II) Metade das pessoas da amostra apresentou idade entre 16 e 20 anos.
III) O coeficiente de variação foi de aproximadamente 16.8%.

Assinale a melhor opção de resposta.

a) I
b) II
c) III
d) I e II
e) I, II e III

2.45 (CESGRANRIO - 2010 - Petrobrás/Administrador) Uma loja de conveniência localizada em um posto de combustível realizou um levantamento sobre o valor das compras realizadas pelos seus clientes. Para tal tomou uma amostra aleatória de 21 compras, que apresentou o seguinte resultado:

i	Valor	i	Valor	i	Valor
1	R$ 19.40	8	R$ 22.00	15	R$ 18.00
2	R$ 14.00	9	R$ 34.00	16	R$ 29.00
3	R$ 18.30	10	R$ 15.50	17	R$ 34.00
4	R$ 27.20	11	R$ 28.50	18	R$ 15.50
5	R$ 8.70	12	R$ 34.00	19	R$ 13.40
6	R$ 10.30	13	R$ 10.80	20	R$ 17.00
7	R$ 7.20	14	R$ 15.50	21	R$ 19.00

A mediana dessa série de observações é:

a) 15.5
b) 18.0
c) 18.3
d) 28.5
e) 34.0

Capítulo 3 Probabilidade

Fundamentos

3.1 (P) Descreva o espaço amostral associado aos experimentos aleatórios descritos a seguir.

a) Jogue uma moeda e observe a face voltada para cima.
b) Jogue um dado e observe o numero mostrado na face voltada para cima.
c) Jogue uma moeda duas vezes e observe a sequência obtida de caras e coroas.
d) Jogue uma moeda duas vezes e observe o número de caras obtido.
e) Jogue uma moeda três vezes e observe a sequência obtida de caras e coroas.
f) Jogue uma moeda três vezes e observe o número de caras obtido.
g) Conte o número de peças fabricadas em certo processo industrial até que dez peças perfeitas sejam produzidas.
h) Uma caixa contém vinte unidades de certo artigo das quais quatro são defeituosas. Observe o número de peças extraídas de tal caixa até que se obtenha todas as peças defeituosas.
i) Observe o tempo de duração, em horas, de uma lâmpada conectada a uma fonte elétrica.

3.2 (R) Considere a experiência que consiste em pesquisar famílias com três crianças, em relação ao sexo das mesmas, segundo a ordem de nascimento. Enumerar os seguintes eventos descritos a seguir.

a) Ocorrência de dois filhos do sexo masculino.
b) Ocorrência de pelo menos um filho do sexo masculino.
c) Ocorrência de no máximo duas crianças do sexo feminino.
d) Ocorrência de nenhuma criança do sexo feminino.
e) Ocorrência de somente crianças do sexo feminino.

3.3 (R) Considerando dois eventos A e B de um mesmo espaço amostral S, expresse em termos de operações entre eventos:

a) A ocorre mas B não ocorre.
b) Exatamente um dos eventos ocorre.
c) Nenhum dos eventos ocorre.

3.4 (R) Determinar a probabilidade de que a soma dos pontos no lançamento de dois dados seja sete.

3.5 (R) Um grupo de 20 homens e três mulheres concorre a um prêmio, através de um sorteio. Qual a probabilidade de um homem ser sorteado?

3.6 (R) Uma urna contém 5 bolas vermelhas, 3 azuis, 10 amarelas. Qual a probabilidade de sortearmos uma bola e ela não ser vermelha?

3.7 (R) Uma pesquisa de satisfação foi realizada, na qual clientes avaliaram 200 comerciantes de pneus de uma grande cidade. Os resultados estão resumidos na tabela a seguir.

	Bom serviço	Serviço deficiente
Determinada marca	64	16
Qualquer marca	42	78

Selecionado aleatoriamente um desses vendedores de pneus, (isto é, cada vendedor tem a mesma probabilidade de ser selecionado), determine a probabilidade de:

a) escolher um vendedor de determinada marca.
b) escolher um vendedor que presta bons serviços.
c) escolher um vendedor de determinada marca e que presta bons serviços.
d) sabendo-se que o vendedor escolhido é de determinada marca, prestar bons serviços.
e) um vendedor prestar bons serviços, dado que não é vendedor de uma única marca determinada.

3.8 (P) Uma urna contém seis fichas azuis, quatro fichas brancas e cinco fichas cinzas. Supondo a extração de uma ficha, calcule a probabilidade de que saia uma ficha:

a) azul.
b) branca.
c) cinza.
d) azul ou cinza.
e) não azul.

3.9 (P) Supondo a extração sucessiva de três fichas, calcule a probabilidade de que saia:

a) adotado o esquema com reposição, as duas primeiras azuis e a terceira branca.
b) adotado o esquema sem reposição, as duas primeiras azuis e a terceira branca.
c) adotado o esquema sem reposição, as três azuis.
d) adotado o esquema sem reposição, a primeira e a terceira brancas e a segunda azul.
e) adotado o esquema sem reposição, a primeira azul, a segunda branca e a terceira cinza.

3.10 (P) Supondo o lançamento de um dado equilibrado, calcule a probabilidade de se obter:

a) um ponto múltiplo de dois ou de três.
b) um ponto múltiplo de dois e de três.

3.11 (P) Considerando um baralho completo, determine a probabilidade de se obter:

a) supondo a extração de uma carta, ou rei ou espadas.
b) supondo a extração de duas cartas, adotado o esquema com reposição, uma carta vermelha e uma negra.
c) supondo a extração de duas cartas, adotado o esquema com reposição, uma carta de espadas e uma dama.
d) supondo a extração de duas cartas, adotado o esquema sem reposição, duas cartas vermelhas.

3.12 (P) A probabilidade de que certa porta esteja chaveada é igual a 0.8. A chave correspondente a tal porta está em um chaveiro que contém cinco chaves. Se uma pessoa seleciona uma das chaves ao acaso, determine a probabilidade de que a porta seja aberta na primeira tentativa.

3.13 (P) Sabendo que a probabilidade de que um aluno aplicado obtenha aprovação em um teste de Estatística é de 4/5, e que a de um aluno não aplicado é de 1/5, calcule a probabilidade de que:

a) somente o aluno aplicado seja aprovado.
b) somente o aluno não aplicado seja aprovado.
c) ao menos um dos alunos seja aprovado.
d) nenhum aluno seja aprovado.

3.14 (P) A probabilidade de um aluno resolver certo problema é de 1/5 e a de outro aluno é de 5/6. Sabendo que os alunos tentam solucionar o problema independentemente, determine a probabilidade de o problema ser resolvido:

a) somente pelo primeiro aluno.
b) somente pelo segundo aluno.
c) por ambos.
d) por nenhum.
e) por ao menos um dos alunos.

3.15 (P) A probabilidade de um homem estar vivo daqui a 25 anos é de 3/5 e a de sua mulher é de 5/6. Determine a probabilidade de:

a) ambos estarem vivos.
b) somente o homem estar vivo.
c) somente a mulher estar viva.
d) ao menos um estar vivo.
e) ao menos um estar morto.

3.16 (P) A peça “A” de um automóvel é produzida por certa fábrica com 80% de probabilidade de ser perfeita. A peça “B” que na montagem do veículo deve ser ajustada à peça “A”, é produzida por outra fábrica com 90% de probabilidade de ser perfeita. Sabendo que o encaixe de tais peças só é aceitável quando ambas são perfeitas e selecionando-se, ao acaso, uma unidade de cada peça, a probabilidade de se obter um encaixe:

a) aceitável.
b) inaceitável.
c) inaceitável por ser somente a peça “A” defeituosa.
d) inaceitável por ser somente a peça “B” defeituosa.
e) inaceitável por serem as duas peças defeituosas.

3.17 (P) Uma empresa apresenta a probabilidade de erro de data em seus registros igual a 0.04 e de erro por inversão de valores igual a 0.05. Sabendo que o auditor irá apontar a ocorrência de qualquer tipo de incorreção e supondo o exame de 500 documentos, determine o número de documentos que se espera:

a) serem apontados.
b) serem apontados devido somente a erro por inversão de valores.
c) serem apontados devido somente a erro de data.

3.18 (P) Certa máquina apresenta a probabilidade de produzir parafusos com defeito de fenda igual a 0.1 e de parafusos tortos igual a 0.05. Sabendo que o controle de qualidade considera o parafuso inaceitável quando se constata qualquer dos defeitos e supondo o exame de um lote de 5 mil parafusos, determine o número de parafusos que se espera sejam considerados:

a) inaceitáveis.
b) inaceitáveis devido somente a defeito de fenda.
c) inaceitáveis por serem somente tortos.

Variável aleatória discreta

3.19 (P) Construa um espaço amostral correspondente ao lançamento simultâneo de dois dados honestos e calcule a probabilidade de se obter:

a) uma soma de pontos igual a dez.
b) um par de valores iguais.
c) uma soma de pontos maior do que dez.
d) um par de valores com a primeira componente menor ou igual a três ou com a segunda componente menor ou igual a dois.
e) uma soma de pontos igual a sete ou dez.

3.20 (P) Seja uma urna que contém quatro fichas azuis, seis fichas brancas e dez fichas cinzas. Considerando X a variável aleatória que assume o valor zero, um e dois quando ocorre, respectivamente, a extração de ficha azul, branca e cinza, determine:

a) a distribuição de probabilidade associada a X.
b) valor esperado, variância, desvio padrão e coeficiente de variação.
c) $P(X \leq 1)$.
d) $P(X > 0)$.
e) $P(0 < X \leq 2)$.

3.21 (P) Seja uma urna que contém duas fichas azuis, quatro fichas brancas, seis fichas cinzas e oito fichas verdes. Considerando X a variável aleatória que assume o valor zero, um, dois e três quando ocorre, respectivamente, a extração de ficha azul, branca, cinza e verde, determine:

a) a distribuição de probabilidade associada a X.
b) valor esperado, variância, desvio padrão e coeficiente de variação.
c) a $P(X \leq 1)$.
d) a $P(X > 1)$.
e) a $P(1 < X \leq 3)$.

3.22 (P) Seja uma urna que contém três fichas azuis, cinco fichas brancas e duas fichas cinzas. Considerando X a variável aleatória que assume o valor zero, um e dois quando ocorre, respectivamente, a extração de uma ficha azul, branca e cinza, determine:

a) a distribuição de probabilidade associada a X.
b) valor esperado, variância, desvio padrão e coeficiente de variação.
c) $P(X \leq 1)$.
d) $P(X > 0)$.
e) $P(0 < X \leq 2)$.

3.23 (P) Considere a distribuição de probabilidade da variável aleatória X a seguir.

x	$p(x)$
0	0.15
1	0.20
2	0.35
3	0.15
4	0.10
5	0.05
$\sum$	**1.00**

Determine:

a) valor esperado, variância, desvio padrão e coeficiente de variação.
b) $P(X \leq 3)$.
c) $P(X > 2)$.
d) $P(1 < X \leq 4)$.
e) $P(0 < X \leq 5)$.

3.24 (P) Considerando a distribuição de probabilidade da variável aleatória X conforme tabela a seguir, determine:

x	$f(x)$
0	0.05
1	0.30
2	0.35
3	0.20
4	0.10
$\sum$	**1.00**

a) valor esperado, variância, desvio padrão e coeficiente de variação.
b) $P(X \leq 1)$.
c) $P(X > 2)$.
d) $P(1 < 3)$.
e) $P(1 < X \leq 3)$.

Distribuição binomial

3.25 (P) Supondo o experimento aleatório consistente no duplo lançamento de uma moeda honesta e considerando X como sendo a variável "número de vezes que ocorre a face cara em dois lançamentos", calcule:

a) $P(X = 0)$.
b) $P(X = 1)$.
c) $P(X = 2)$.

3.26 (P) A probabilidade de exemplar defeituoso com que opera certo processo produtivo é de 0.10. Considerando X a variável "número de unidades defeituosas em uma amostra ocasional de quatro unidades", determine:

a) a distribuição de probabilidades associada a X.
b) valor esperado, variância, desvio padrão e coeficiente de variação.
c) $P(X \leq 3)$.
d) $P(X > 2)$.
e) $P(1 < X \leq 3)$.
f) $P(1 \leq X < 3)$.

3.27 (P) Sendo X a variável aleatória 'número de vezes que ocorre a face cara no triplo lançamento de uma moeda perfeita' (ou no lançamento simultâneo de três moedas em iguais condições), determine:

a) a distribuição de probabilidades associada a X.
b) o valor esperado, a variância e o coeficiente de variação.
c) $P(X > 2)$.
d) $P(X \leq 3)$.
e) $P(0 < X \leq 2)$.
f) $P(1 \leq X < 3)$.

3.28 (P) Sendo X a variável aleatória 'número de vezes que ocorre o evento número par no triplo lançamento de um dado perfeito' (ou no lançamento simultâneo de três dados em iguais condições), determine:

a) a distribuição de probabilidade associada a X.
b) o valor esperado, a variância e o coeficiente de variação.
c) $P(X \leq 2)$.
d) $P(X > 1)$.
e) $P(0 < X \leq 2)$.
f) $P(1 \leq X < 3)$.

3.29 (P) Sendo X a variável aleatória número de vezes que ocorre o evento 'número múltiplo de dois ou de três no triplo lançamento de um dado perfeito' (ou no lançamento simultâneo de três dados em iguais condições), determine:

a) a distribuição de probabilidade associada a X.
b) o valor esperado, a variância e o coeficiente de variação.
c) $P(X \leq 1)$.
d) $P(X > 1)$.
e) $P(0 < X \leq 2)$.
f) $P(1 \leq X < 3)$.

3.30 (P) A probabilidade de exemplar defeituoso com que opera certo processo produtivo e de 0.20. Considerando X a variável aleatória 'número de unidades defeituosas em uma amostra ocasional de cinco unidades', determine:

a) a distribuição de probabilidade associada a X.
b) o valor esperado, a variância e o coeficiente de variação.
c) $P(X \leq 4)$.
d) $P(X > 2)$.
e) $P(1 < X \leq 4)$.
f) $P(2 \leq X < 4)$.

3.31 (P) Sabendo que certo processo industrial produz, em média, 20% de unidades defeituosas e considerando uma amostra ocasional de quatro unidades, determine a probabilidade de se obter:

a) nenhuma unidade defeituosa.
b) uma unidade defeituosa.
c) mais de uma unidade defeituosa.

d) ao menos uma unidade defeituosa.
e) zero ou uma unidade defeituosa.

3.32 (P) Sabendo que certo processo industrial produz, em média, 10% de unidades defeituosas e considerando uma amostra ocasional de cinco unidades, determine a probabilidade de se obter:

a) nenhuma unidade defeituosa.
b) uma unidade defeituosa.
c) mais de uma unidade defeituosa.
d) ao menos uma unidade defeituosa.
e) zero ou uma unidade defeituosa.

3.33 (P) Sabendo que a probabilidade de se obter aprovação em certo teste de Estatística é igual a 0.7 e considerando um grupo de cinco estudantes, determine a probabilidade de:

a) nenhum ser aprovado.
b) apenas um ser aprovado.
c) ao menos um ser aprovado.
d) dois serem aprovados.
e) no máximo dois serem aprovados.

Distribuição Poisson

3.34 (P) Sabendo que certo processo industrial produz, em média, 4% de unidades defeituosas e considerando uma amostra ocasional de cem unidades, determine a probabilidade de se obter:

a) nenhuma unidade defeituosa.
b) uma unidade defeituosa.
c) mais de uma unidade defeituosa.
d) ao menos uma unidade defeituosa.
e) duas unidades defeituosas.
f) mais de duas unidades defeituosas.
g) ao menos duas unidades defeituosas.
h) uma ou duas unidades defeituosas.

3.35 (P) Sabendo que em certo processo industrial, em média, duas máquinas necessitam de conserto, determine a probabilidade de que o número de máquinas que necessitem de conserto, num dia qualquer, seja:

a) igual a zero.
b) igual a um.
c) maior ou igual a um.
d) maior do que um.
e) igual a dois.
f) maior do que dois.
g) menor do que dois.

3.36 (P) Sabendo que a central telefônica de certa cidade pode fazer, no máximo, 10 atendimentos por minuto e que a média de chamadas é de 180 por hora, determine a probabilidade da central telefônica receber num determinado minuto:

a) nenhuma solicitação de atendimento.
b) uma solicitação de atendimento.
c) mais de uma solicitação de atendimento.
d) ao menos uma solicitação de atendimento.
e) mais solicitações de atendimento do que pode suportar.

3.37 (P) Sabendo que em certo processo industrial a média mensal de acidentes pessoais é igual a 0.5, determine a probabilidade de que, ao longo de quatro meses, verifique-se:

a) nenhum acidente.

b) um acidente.
c) ao menos um acidente.
d) mais de um acidente.
e) dois acidentes.
f) no máximo dois acidentes.

Variável aleatória contínua

3.38 (P) Sendo X uma v.a.c. com função densidade de probabilidade $f(x) = \frac{1}{2} - \frac{x}{8}$ e que assume valores no intervalo $[0, 4]$, determinar:

a) $P(X > 3)$.
b) $P(X < 1)$.
c) $P(1 < X < 3)$.

3.39 (P) Sendo X uma v.a.c. com função densidade de probabilidade definida por $f(x) = \frac{1}{6}$, $0 < x < 6$, calcular:

a) $P(X < 3)$.
b) $P(4 < X < 6)$.
c) $P(2 < X < 5)$.

3.40 (P) Sendo X uma v.a.c. com distribuição continua de probabilidade dada por $f(x) = 1 - \frac{x}{2}$, $[0, 2]$, calcular:

a) $P(0.5 < X < 1.5)$.
b) $P(X < 0.5)$.
c) $P(1.5 < X < 2)$.

3.41 (P) Considerando que o tempo, dado em horas, gasto com a recuperação de certo tipo de chapa é uma v.a.c., com função densidade de probabilidade $f(x) = kx(1 - x)$, $x \in [0, 1]$, determinar:

a) k.
b) a probabilidade de que se gaste menos do que 1/4 de hora com a recuperação de uma chapa.
c) a probabilidade de que se gaste mais do que 3/4 de hora com a recuperação de uma chapa.
d) $P(0.5 < X < 0.8)$.
e) a média aritmética.
f) a variância.
g) o desvio padrão.
h) o coeficiente de variação.

3.42 (P) Sendo X uma v.a.c. com função densidade de probabilidade representada por $f(x) = kx$ e que assume valores no intervalo $[0, 2]$, determinar:

a) k.
b) $P(X > 1)$.
c) $P(X < 0.5)$.
d) $P(0.5 < X < 1)$.
e) a média aritmética.
f) a variância.
g) o desvio padrão.
h) coeficiente de variação.

Distribuição normal

3.43 (P) Sendo Z uma variável aleatória contínua com distribuição $\mathcal{N}(0, 1)$, determine:

a) $P(0 < Z < 1)$.
b) $P(0 < Z < 1.35)$.
c) $P(-2.33 < Z < 1.96)$.

d) $P(0.65 < Z < 1.96)$.
e) $P(-2.33 < Z < -0.65)$.
f) $P(-1.96 < Z < 1.96)$.
g) $P(Z < -2.33 \text{ ou } Z > 1.96)$.
h) $P(Z < -1.96 \text{ ou } Z > 2.33)$.

3.44 (P) Sendo Z uma variável aleatória contínua com distribuição $\mathcal{N}(0,1)$, determine os valores de z que satisfazem as seguintes condições:

a) $P(0 < Z < z) = 0.4772$.
b) $P(z < Z < 0) = 0.4452$.
c) $P(1.20 < Z < z) = 0.1019$.
d) $P(-1.20 < Z < z) = 0.7061$.
e) $P(-2.21 < Z < z) = 0.3236$.
f) $P(Z < -1.96 \text{ ou } Z > z) = 0.0478$.
g) $P(Z < -1.00 \text{ ou } Z > z) = 0.1837$.

3.45 (R) Seja Z uma variável aleatória contínua distribuída normalmente com média zero e desvio padrão 1, isto é, $Z \sim \mathcal{N}(0,1)$. Determine:

a) $P(Z \leq -1)$
b) $P(Z \leq -2.89)$
c) $P(Z \geq 2)$
d) $P(-1 \leq Z \leq 2.03)$
e) $P(Z > -1.3)$
f) $P(Z < 0)$
g) $P(-1 \leq Z \leq -0,61)$
h) $P(Z > 1.6)$
i) $P(Z < -1.74)$
j) $P(1.31 \leq Z \leq 2.41)$
k) $P(Z < 2.35)$
l) $P(Z > 4.36)$
m) $P(Z < -6.32)$
n) $P(Z > 4.21)$

3.46 (R) Seja Z uma v.a.c. normalmente distribuída com média 0 e desvio padrão 1. Determine o valor de z tal que:

a) $P(Z \leq z) = 0.0495$
b) $P(Z \leq z) = 0.9474$
c) $P(Z \geq z) = 0.0618$
d) $P(Z \geq z) = 0.8212$

3.47 (R?) Sejam z_1 e z_2, simétricos, dois particulares valores de Z. Determine-os tais que:

a) $P(z_1 \leq Z \leq z_2) = 0.9216$
b) $P(z_1 \leq Z \leq z_2) = 0.8858$

3.48 (F) Considere a variável aleatória Z com distribuição normal padrão anotada por $Z \sim \mathcal{N}(0,1)$. Obtenha as seguintes quantidades:

a) $Pr(Z \leq -1.96)$.
b) $Pr(-1 \leq Z \leq 1)$.
c) $Pr(Z > 1.64)$.
d) q tal que $Pr(Z \leq q) = 0.025$.
e) q tal que $Pr(-q \leq Z \leq q) = 0.6826894921$.
f) q tal que $Pr(Z > q) = 0.05$.

3.49 (P) Sendo X uma variável aleatória contínua com distribuição $\mathcal{N}(10, 2)$, determine os valores de x que satisfazem as seguintes condições:

a) $P(10 < X < x) = 0.4772$.
b) $P(8 < X < x) = 0.3413$.
c) $P(6 < X < x) = 0.1359$.
d) $P(11 < X < x) = 0.2417$.
e) $P(X > x) = 0.1251$.
f) $P(X > x) = 0.8869$.
g) $P(X > x) = 0.1020$.
h) $P(X < x) = 0.1075$.
i) $P(X < x) = 0.7996$.
j) $P(X < x) = 0.1112$.

3.50 (P) Sabendo que a massa (em quilogramas) dos estudantes de certa universidade está distribuida segundo uma $\mathcal{N}(80, 5)$, determine a proporção de alunos que pesam:

a) entre 76 e 85 quilogramas.
b) entre 82 e 86 quilogramas.
c) entre 76 e 78 quilogramas.
d) mais do que 82 quilogramas.
e) mais do que 78 quilogramas.
f) menos do que 82 quilogramas.
g) menos do que 78 quilogramas.
h) menos do que 75 ou mais do que 90 quilogramas.

3.51 (P) Sabendo que a estatura, dada em centímetros, dos 2000 estudantes de certa escola está distribuída segundo uma $\mathcal{N}(170, 10)$, determine o número de alunos com estatura:

a) entre 162 e 174 centímetros.
b) entre 174 e 180 centímetros.
c) entre 150 e 168 centímetros.
d) maior do que 174 centímetros.
e) maior do que 165 centimetros.
f) menor do que 160 centímetros.
g) menor do que 180 centímetros.
h) menor do que 160 ou maior do que 180 centímetros.

3.52 (P) Considerando que o peso, dado em gramas, de determinado artigo produzido por uma fábrica seja $\mathcal{N}(20, 4)$, determine:

a) a probabilidade de uma unidade, selecionada ao acaso, pesar entre 16 e 22 gramas.
b) a probabilidade de uma unidade, selecionada ao acaso, pesar entre 22 e 25 gramas.
c) a probabilidade de uma unidade, selecionada ao acaso, pesar mais de 23 gramas.
d) a probabilidade de uma unidade, selecionada ao acaso, pesar menos de 16 gramas.
e) o intervalo, centrado na média, dado em gramas, para o qual corresponda a probabilidade de 0,899 de que o peso de uma unidade selecionada ao acaso nele esteja contida.
f) o peso, dado em gramas, abaixo do qual se espera encontrar 25.78% das unidades produzidas.
g) o peso, dado em gramas, acima do qual se espera encontrar 34.46% das unidades produzidas.
h) supondo a extração de uma amostra aleatória de 500 unidades, quantas se espera que pesem entre 15 e 21 gramas.
i) supondo a extração de uma amostra aleatória de 1000 unidades, quantas se espera que pesem mais de 24 gramas.
j) o peso médio das unidades produzidas pela máquina do processo industrial mencionado, que deve ser adotado, para que apenas 2.28% das unidades pesem menos do que 12 gramas.
k) a probabilidade de uma unidade, selecionada ao acaso, pesar menos do que 18 ou mais do que 24 gramas.

l) a probabilidade de uma unidade, selecionada ao acaso, pesar menos do que 16 ou mais do que 24 gramas.

3.53 (P) Sabendo que os quocientes de liquidez normal das empresas de uma determinada região estão distribuídos segundo uma $\mathcal{N}(1.10, 0.08)$, determine:

a) a probabilidade de que uma empresa, selecionada ao acaso, possua um quociente de liquidez entre 1.02 e 1.26.
b) a probabilidade de que uma empresa, selecionada ao acaso, possua um quociente de liquidez entre 1.18 e 1.26.
c) a probabilidade de que uma empresa, selecionada ao acaso, possua um quociente de liquidez maior do que 1.30.
d) a probabilidade de que uma empresa, selecionada ao acaso, possua um quociente de liquidez menor do que 1.26.
e) o intervalo, centrado na média, dado em termos de quociente de liquidez, para o qual corresponda a probabilidade de 0,8664 de que o quociente de liquidez de uma empresa selecionada ao acaso nele esteja contido.
f) o quociente de liquidez normal abaixo do qual se espera encontrar 10.56% das empresas.
g) o quociente de liquidez normal acima do qual se espera encontrar 1.22% das empresas.
h) sabendo que na região existem 5000 empresas e que a rede bancária aceita o desconto de títulos somente das que possuem um quociente de liquidez normal igual ou superior a 1.04, qual o numero provável das empresas que utilizam a mencionada operação.
i) sabendo que a Secretaria da Receita Federal resolveu efetuar um exame detalhado das empresas cujos quocientes de liquidez normal se afastam do quociente médio duas e duas e meia unidades de desvio padrão, respectivamente, para menos e para mais, qual o número de inspeções a serem realizadas.

3.54 (R) Seja X uma v.a.c. normalmente distribuída com média 300 e desvio padrão 2. Calcule a probabilidade de X assumir valores:

a) menores que 302.48
b) maiores que 298.14.
c) entre 297.6 e 303.86.

3.55 (R) O número de pedidos de compra de certo produto que uma empresa recebe por semana distribui-se normalmente, com média 125 e desvio padrão de 25. Se em uma dada semana o estoque disponível é de 150 unidades, determine:

a) a probabilidade de que todos os pedidos sejam atendidos.
b) qual deveria ser o estoque para se tivesse 99% de probabilidade de que todos os pedidos fossem atendidos.

3.56 (R) Uma máquina automática que enche garrafas de refrigerantes está regulada para que o volume médio de líquido em cada garrafa seja de 1000 cm^3, com desvio padrão de 10 cm^3. Pode-se admitir que a distribuição da variável seja normal.

a) Qual a proporção de garrafas em que o volume de líquido é menor que 990 cm^3?
b) Qual a proporção de garrafas em que o volume do líquido não se desvia da média em mais do que dois desvios padrões?
c) O que acontecerá com a proporção do item (b) se a máquina for regulada de forma que a média seja 1200 cm^3 e o desvio padrão 20 cm^3?

3.57 (R) Suponha que o peso das pessoas de certa comunidade tenha distribuição normal com média 60 kg e desvio padrão 10 kg. Determine a porcentagem das pessoas que pesam:

a) 55kg ou mais.
b) 58 kg ou menos.
c) entre 52 e 70 kg.

3.58 (F) A idade de pessoas que pedem financiamentos de automóveis (X) segue uma distribuição normal com média 27 anos e desvio padrão 4 anos, denotada por $X \sim \mathcal{N}(27, 4)$.

a) Em um grupo de pessoas que pedem financiamento, qual a probabilidade de uma pessoa ter mais de 33 anos?
b) Neste mesmo grupo, qual a probabilidade de uma pessoa ter idade entre 32 e 40 anos?
c) Se $P(X < x) = 0.6217$, qual o valor de x?
d) Interprete o valor de x no contexto do problema.

3.59 (R) Sabe-se que os graus atribuídos por certo professor a seus alunos tem distribuição normal com média 5 e desvio padrão 2. O professor atribuiu conceitos da seguinte forma:

A: grau maior ou igual a 8.
B: grau maior ou igual a 6 e inferior a 8.
C: grau maior ou igual a 4 e inferior a 6.
D: grau inferior a 4.

Determine a proporção de alunos com conceito A, B, C, e D.

3.60 (R) O gerente de um banco tem seu domicílio no bairro A. Ele deixa sua casa às 8h e 45min dirigindo-se ao emprego e iniciando seu trabalho às 9h. A duração dessa viagem tem média de 13 min e desvio padrão 3 min. Considerando que o tempo de duração da viagem tem distribuição normal, determine a probabilidade de o gerente chegar atrasado ao banco.

3.61 (R) As vendas diárias de um restaurante têm distribuição normal com média igual a 53 unidades monetárias e desvio padrão igual a 12 u.m.:

a) Qual a probabilidade de as vendas excederem 70 u.m. em determinado dia?
b) Esse restaurante deve vender no mínimo 30 u.m. por dia, para não ter prejuízo. Qual a probabilidade de que em certo dia haja prejuízo?

3.62 (R) Suponhamos que o tempo de estudo formal de adultos de certo país apresente distribuição normal com média de 11 anos e desvio padrão de 2 anos. Determine:

a) a probabilidade de que um adulto, escolhido aleatoriamente, tenha entre 9 e 14 anos de estudo formal.
b) a probabilidade de que um adulto tenha mais de 18 anos de estudo formal.
c) o número de adultos que se espera que tenham menos de 7 anos de estudo formal, considerando uma amostra de 500 adultos.

3.63 (F) As geladeiras produzidas por uma fábrica possuem um determinado tempo de vida até o primeiro estrago. Estudos apontam que este tempo segue distribuição normal com média 1.45 ano e desvio padrão igual a 0.15 ano.

a) A fábrica oferece garantia de 1 ano. Qual a probabilidade de uma geladeira estragar neste período?
b) Qual a probabilidade de uma geladeira estragar fora da garantia?
c) Qual a probabilidade de uma geladeira falhar entre o primeiro e o segundo ano de uso?
d) Qual a probabilidade de uma geladeira durar mais de 2 anos sem apresentar falhas?
e) Se a fábrica produziu 80 mil geladeiras, quantas pessoas devem acionar a garantia?

3.64 O atendimento dos caixas de um determinado banco fica sobrecarregado entre o primeiro e o décimo dia do mês. Neste período, o tempo de espera do caixa convencional (X) tem distribuição normal com média de 23 minutos e desvio padrão de 4 minutos. Para o caixa prioritário (Y), este tempo distribui-se com média de 15 minutos e desvio padrão igual a 3.

a) No caixa convencional, qual a probabilidade de você esperar mais de 20 minutos para ser atendido? E no caixa prioritário?
b) Você leva em torno de meia hora para ler o caderno de esportes do jornal. Qual a probabilidade de você terminar a leitura enquanto espera na fila do caixa? Faça as contas para ambos os caixas e compare.
c) Uma vovó de 90 anos chegou no banco. Qual a probabilidade de ela esperar entre 20 e 25 minutos para ser atendida?
d) Há uma grande placa indicando que o tempo de espera máximo é de 12 minutos para os clientes preferenciais e 18 minutos para os demais clientes. Com que frequência as pessoas esperam mais do que este tempo para serem atendidas?

e) Você foi chamado para corrigir este tempo máximo. A orientação é que apenas 10% dos clientes sejam atendidos em um tempo maior que o indicado. Qual deveria ser o novo tempo para o caixa preferencial? E para o caixa convencional?

3.65 Em um concurso estão inscritas 1000 pessoas para 150 vagas. As notas das provas seguiram distribuição normal com média 6.2 e desvio padrão igual a 1, anotado por $X \sim \mathcal{N}(6.2, 1)$.

a) Selecionando ao acaso um candidato, qual a probabilidade de ele ter tirado menos que 5 na prova?
b) Qual a probabilidade de um candidato ter notas entre 5 e 6?
c) Aproximadamente quantas pessoas tiraram notas entre 5 e 6?
d) Qual a nota mínima para obter a aprovação?

3.66 O lucro líquido de uma loja (X) segue uma distribuição normal com média 15000 reais e desvio padrão de 5000 reais, denotado por $X \sim \mathcal{N}(15000, 5000)$.

a) Qual a probabilidade de o lucro líquido ser maior que 20 mil reais?
b) Qual a probabilidade de o lucro líquido estar entre 13 mil e 22 mil reais?
c) Qual a probabilidade de a loja dar prejuízo?

Questões de concursos

3.67 (Anpec 1991, EST-Q3) Em uma universidade, 30% dos homens e 20% das mulheres estudam matemática. Além disso, 45% dos estudantes são mulheres. Se um estudante é escolhido aleatoriamente:

(0) A probabilidade dele estudar matemática é 0.255.
(1) Se o estudante escolhido estuda matemática, então a probabilidade de que seja uma mulher é 0.09.
(2) Se o estudante escolhido estuda matemática, então a probabilidade de que seja um homem é de 0.647.
(3) A probabilidade dele não estudar matemática é de 0.65.
(4) Não é possível obter as probabilidades acima.

3.68 (Anpec 1991, EST-Q4) Com respeito às distribuições teóricas de probabilidade pode-se afirmar que:

(0) A distribuição de Bernoulli corresponde ao experimento com dois valores possíveis sendo que um deles está associado à ocorrência de sucesso e o outro à de fracasso.
(1) A distribuição binomial corresponde ao número de sucessos em n repetições independentes de experimentos de Bernoulli.
(2) A distribuição hipergeométrica corresponde ao experimento de extração aleatória de uma população dividida segundo dois atributos.
(3) A distribuição geométrica corresponde ao número de vezes que se deve replicar experimentos de Bernoulli, supostos independentes, até que ocorra pela primeira vez um sucesso.
(4) É correto utilizar a distribuição de Poisson para aproximar a binomial qualquer que seja o tamanho de n (tamanho da amostra) e p (probabilidade de sucesso).

3.69 (Anpec 1991, EST-Q9) Em relação às distribuições de probabilidade pode-se afirmar que:

(0) As medidas de localização, média, moda e mediana coincidem na distribuição normal.
(1) A soma de qui-quadrados independentes tem distribuição qui-quadrado.
(2) Uma variável aleatória com distribuição F de Snedecor é proporcional ao quociente de qui-quadrados independentes.
(3) A distribuição t de Student é caracterizada por um único parâmetro que representa os seus graus de liberdade.
(4) A distribuição qui-quadrado é simétrica.

3.70 (Anpec 1991, EST-Q10) A demanda mensal de produtos de uma empresa é normalmente distribuída com média de 10 unidades e variância de 64. Se as demandas dos meses se comportam independentemente pode-se afirmar que:

(0) A probabilidade de a demanda em abril ser superior a 16 unidades é maior do que 20%.
(1) A probabilidade de a demanda no 2º trimestre ser inferior a 50 unidades é maior do que 70%.
(2) A probabilidade de a demanda no 1º semestre ficar entre 80 e 100 unidades é menor do que 10%.

(3) A probabilidade de a demanda anual ser maior ou igual a 200 unidades é aproximadamente 15%.
(4) Não se pode calcular nenhuma das probabilidades acima.

3.71 (Anpec 1991, EST-Q12) Sejam X e Y variáveis aleatórias independentes tais que: $E(X) = 3$, $E(Y) = 2$, $E(X^2) = 10$ e $E(Y^2) = 7$. Pode-se afirmar que:

(0) $E(X, Y) = 6$.
(1) $Var(X + Y) = 4$.
(2) $Var(Y - 3X) = 6$.
(3) O coeficiente de correlação entre X e Y é igual a 1/9.
(4) $E\{[X - E(X)][Y - E(Y)]\}$ não pode ser calculado.

3.72 (Anpec 1991, EST-Q13) O preço X de um produto é considerado uma variável aleatória com função densidade de probabilidade dada por $f(x) = kx^3$, $1 \leq x \leq 3$ onde k é uma constante positiva. Pode-se afirmar que:

(0) O preço médio deste produto é aproximadamente 2.42.
(1) A probabilidade do preço ser menor que 1.5 e menor do que 70%.
(2) A probabilidade do preço ser menor que 1.2 é maior do que 0.5%.
(3) O valor de k é 1/80.
(4) A variância do preço deste produto é 9.1.

3.73 (Anpec 1992, EST-Q2) Com relação à Teoria da Probabilidade pode-se afirmar que:

(0) O espaço amostral de um experimento é o conjunto de resultados possíveis deste experimento.
(1) O evento é um resultado possível do experimento.
(2) Se A e B são eventos independentes, então $P(A|B) = Pa)$.
(3) Se A e B são eventos mutuamente exclusivos, então eles são independentes.
(4) A definição clássica de Probabilidade pressupõe que todos os resultados de um experimento são igualmente prováveis.

3.74 (Anpec 1992, EST-Q3) Em uma universidade, 30% dos homens e 20% das mulheres estudam matemática. Além disso, 45% dos estudantes são mulheres. Se um estudante é escolhido aleatoriamente está estudando matemática, qual é a probabilidade de que esse estudante seja mulher? (use duas casas decimais e multiplique o resultado por 100).

3.75 (Anpec 1992, EST-Q4) Sejam X e Y duas variáveis aleatórias contínuas, então:

(0) Se elas forem independentes, $E(XY) = E(X)E(Y)$.
(1) Se elas forem independentes, $Cov(XY) = 0$.
(2) Se elas forem independentes, $V(X/Y) = V(X)/V(Y)$.
(3) O coeficiente de correlação linear entre as variáveis X e Y é dado por $Cov(X, Y)/\sqrt{V(X)V(Y)}$.
(4) Se o coeficiente de correlação for nulo, isto indica que X e Y são independentes.

3.76 (Anpec 1992, EST-Q5) Qual o menor valor que k pode assumir, de acordo com o Teorema de Chebyshev, para que $P(\mu - k\sigma \leq x \leq \mu + k\sigma) \geq 0.06$, isto é, a probabilidade de x estar entre $\mu - k\sigma$ e $\mu + k\sigma$ seja maior ou igual a 0.06?

3.77 (Anpec 1992, EST-Q7) As distribuições Normal, t, F e χ^2 são muito utilizadas em Econometria. Sobre elas pode-se afirmar que:

(0) Somente a Normal é simétrica.
(1) Todas têm como parâmetro o número de graus de liberdade.
(2) Todas são limitadas na reta real.
(3) A variável χ^2 resulta da soma de variáveis aleatórias independentes que têm distribuição Normal padronizada.
(4) A F é, por definição, a razão de duas variáveis qui-quadrado.

3.78 (Anpec 1993, EST-13) Com relação às distribuições t, Qui-quadrado e F pode-se afirmar que:

(0) A soma de normais com média 0 e variância 1 dá sempre uma Qui-quadrado.

(1) Uma F com m e n graus de liberdade resulta da divisão de uma Qui-quadrado com m graus de liberdade por uma Qui-quadrado com n graus de liberdade, sendo ambas independentes.
(2) A Qui-quadrado com m graus de liberdade resulta da soma de m normais independentes, de média zero, ao quadrado.
(3) A distribuição t de n graus de liberdade coincide com a de F com 1 e n graus de liberdade.
(4) A raiz quadrada de uma Qui-quadrado com 2 graus de liberdade dividida por dois distribui-se com uma distribuição normal com média 0 e variância 1.

3.79 (Anpec 1993, EST-Q10) Quanto à desigualdade de Chebyshev, supondo-se que uma variável aleatória tenha média e variância finita, ela assegura que a probabilidade:

(0) da variável assumir um valor maior ou igual a n é menor ou igual à variância mais a média divididos por n^2.
(1) da variável ultrapassar a média de um valor maior ou igual n vezes o desvio padrão é menor ou igual ao inverso de n^2.
(2) da variável ultrapassar a média de um valor maior ou igual a n é menor ou igual ao inverso de n^2.
(3) da variável ultrapassar a média de um valor maior ou igual a n é menor ou igual ao segundo momento dividido por n^2.

3.80 (Anpec 1993, EST-Q1) Em relação às distribuições de probabilidade pode-se afirmar que:

(0) A distribuição F é uma razão entre dois t de Student independentes.
(1) A distribuição binomial é uma distribuição definida por dois parâmetros.
(2) A distribuição de Poisson descreve o comportamento de variáveis aleatórias discretas.
(3) A variável t de Student quando elevada ao quadrado é sempre igual a uma variável F.
(4) A distribuição qui-quadrado tende para a distribuição normal quando o tamanho da amostra tende para o infinito.

3.81 (Anpec 1993, EST-Q12) Com relação à esperança, o segundo momento (não centrado) e a variância de uma variável aleatória, pode-se dizer que:

(0) O segundo momento é sempre menor que a esperança ao quadrado.
(1) A variância pode ser menor ou maior do que o segundo momento.
(2) Se a esperança é zero, o segundo momento é maior do que a variância.
(3) Se a variância é zero, o segundo momento é igual à esperança.
(4) O quadrado da esperança nunca pode ser igual à variância.

3.82 (Anpec 1993, EST-Q4) Sejam X, Y e Z variáveis aleatórias quaisquer. Então:

(0) $Var(X+Y+Z) = Var(X) + Var(Y) + Var(Z)$.
(1) $Var(2X+4) = 4Var(X) + 16$.
(2) $E(X+Y) = E(X) + E(Y)$.
(3) $Cov(X,Y) = E(XY) - E(X).E(Y)$.
(4) $E(X^2) = [E(X)]^2$.

3.83 (Anpec 1993, EST-Q6) Suponha duas caixas de bombons. Na caixa A temos dois bombons com recheio e quatro sem recheio. Na caixa B temos três bombons com recheio e três sem recheio. Um bombom retirado de uma das caixas não tem recheio. Qual a probabilidade que tenha vindo da caixa B? (Multiplique o resultado por 7).

3.84 (Anpec 1993, EST-Q8) Se a função de distribuição de uma variável aleatória X é dada por:

$$F(x) = \begin{cases} 0 & \text{se } x < 0 \\ \frac{1}{2} & \text{se } x \in [0,1) \\ \frac{x}{2} & \text{se } x \in [1,2] \\ 1 & \text{se } x > 2 \end{cases}$$

pode-se dizer que:

(0) X é uma variável aleatória contínua.

(1) A probabilidade de X assumir um valor no intervalo $[1/2, 1]$ é $1/4$.
(2) X assume valores uniformemente no intervalo $[1, 2]$.
(3) A esperança de X é $3/2$.

3.85 (Anpec 1993, EST-Q9) O Teorema Central do Limite (TCL), resultado maior dentre os teoremas da teoria da probabilidade, nas versões estudadas na graduação, estabelece condições que asseguram a convergência de somas de variáveis aleatórias, convenientemente padronizadas, à distribuição normal.

(0) Esta convergência se dá em probabilidade, ou seja, é a mesma da Lei Fraca dos Grandes Números.
(1) As variáveis aleatórias utilizadas na composição da soma não precisam ser independentes.
(2) Se as variáveis aleatórias, atendendo às hipóteses do TCL, tem a mesma média μ e variância σ^2, o teorema garante que a soma das n primeiras, subtraídas de $n\mu$, e dividida por $n\sigma$, converge para uma distribuição normal com média 0 e variância 1 ($N(0,1)$).
(3) A convergência de uma distribuição binomial (n, p), onde n é o número de ensaios de Bernoulli e p é a probabilidade de sucesso em cada um, quando n aumenta, pode ser provada como um simples caso particular do TCL.

3.86 (Anpec 1993, EST-Q15) A variável aleatória Z guarda com a variável aleatória X a relação $Z = 5+5X+U$ onde U é uma variável aleatória, independente de X. Pode-se afirmar que:

(0) Z tem correlação 1 com X.
(1) Qualquer que seja o valor do termo constante na relação acima, a correlação de Z com X não se altera.
(2) A covariância de Z com X é de 25 vezes a variância de X.
(3) Se os desvios padrão de X e de U forem idênticos e iguais a 2, a variância de Z valerá 104.
(4) A correlação de U com Z é independente dos coeficientes da relação acima.

3.87 (Anpec 2023, EST-Q4) Um estudo sobre a eficácia de um teste para gripe foi realizado com 753 pessoas e foram obtidos os seguintes resultados:

	Situação	real	
Resultado do Teste	Doente	Não Doente	Total
Positivo (Doente)	344	108	452
Negativo (Não doente)	84	217	301
Total	428	325	753

Julgue as afirmativas como verdadeiras ou falsas:

(0) Asensibilidade do teste é de 80.37%.
(1) A especificidade do teste é de 33%, aproximadamente.
(2) A taxa de falso positivo é de 33.23%.
(3) A taxa de falso negativo é inferior a 20%.
(4) A probabilidade de o teste apresentar resultado correto é superior a 75%.

3.88 (Anpec 2023, EST-Q5) Sendo X uma variável aleatória com distribuição geométrica com parâmetro p, é correto afirmar:

(0) $E(X) = \frac{1}{p}$.
(1) $Var(X) = \frac{1}{p^2}$.
(2) A probabilidade de $X = 3$ é dada por: $P(X = 3) = p(1-p)^2$.
(3) A probabilidade de $X > 5$ é dada por: $P(X > 5) = p(1-p)^4$.
(4) $P(X \geq 8|X > 4) = P(X > 4)$.

3.89 (Anpec 2024, EST-Q1) Supondo que a taxa de desemprego de determinado país corresponda a 10% da população economicamente ativa (PEA), verifique se as afirmativas abaixo sobre esse país são corretas:

(0) Em uma amostra aleatória de 10 indivíduos da PEA, a probabilidade de que nenhum desses 10 indivíduos esteja desempregado é $10 \times (0.9)^{10}$.

(1) Em uma amostra aleatória de 10 indivíduos da PEA, a probabilidade de encontrar exatamente dois desempregados é $45 \times (0.1)^2 \times (0.9)^8$.
(2) Em uma amostra aleatória de 5 indivíduos da PEA, a probabilidade de 2 ou mais desses indivíduos estarem desempregados é $1 - (0.9)^5 - 0.5 \times (0.9)^4$.
(3) A probabilidade de encontrar 10 desempregados em uma amostra aleatória de 10 indivíduos da PEA é igual a probabilidade de encontrar 5 desempregados em uma amostra aleatória de 5 indivíduos da PEA.
(4) Em uma amostra aleatória de 3 indivíduos da PEA, a probabilidade de encontrar exatamente dois desempregados é igual a 0.027.

3.90 (Anpec 2024, EST-Q1) Em um concurso para professor do departamento de economia de determinada universidade, 12 candidatos foram aprovados e serão contratados ao longo de três anos. No primeiro ano, serão contratados 4 desses 12 candidatos. Os 12 candidatos estão divididos em 3 áreas: 6 em Microeconomia, 4 em Macroeconomia e 2 em Econometria. Os candidatos não estão classificados por nota, todos foram aprovados em igualdade de condições, e a discussão no departamento de economia é definir um critério para escolher os 4 primeiros candidatos a ingressar. Decide-se, então, fazer um sorteio aleatório para escolher esses 4 candidatos entre os 12 aprovados. Defina as variáveis aleatórias X, Y e Z da seguinte forma:

$X =$ número de candidatos na área de Microeconomia contratados no primeiro ano. $Y =$ número de candidatos na área de Macroeconomia contratados no primeiro ano. $Z =$ número de candidatos na área de Econometria contratados no primeiro ano. São corretas as afirmativas:

(0) A probabilidade de ter exatamente 2 candidatos da área de Microeconomia e 1 de Macroeconomia entre os 4 contratados no primeiro ano é igual a $frac833$.
(1) A probabilidade de ter exatamente 2 candidatos da área de Microeconomia entre os 4 contratados no primeiro ano é igual a $frac511$.
(2) A probabilidade de ter pelo menos um candidato da área de Econometria entre os 4 contratados no primeiro ano é igual a $\frac{12}{33}$.
(3) $E(X) = 1$.
(4) A probabilidade de ter exatamente 1 candidato da área de Microeconomia e 1 de Macroeconomia entre os 4 contratados no primeiro ano é igual a $\frac{8}{165}$.

Capítulo 4 Amostragem

Veja a tabela de dígitos aleatórios da USP ao final deste material.

4.1 (R) Uma escola deseja selecionar alunos para uma pesquisa sobre o uso de tecnologia. A escola possui 400 alunos. Eles criam uma lista de todos os alunos e selecionam um aluno a cada 10 da lista. Que tipo de amostragem é essa?

a) Aleatória Simples
b) Estratificada
c) Sistemática
d) Conglomerados
e) Aleatória Simples com substituição

4.2 (R) Uma universidade quer pesquisar a opinião dos estudantes sobre a qualidade do ensino. Ela divide os estudantes em grupos com base nos cursos (por exemplo, Medicina, Engenharia, Administração) e, em seguida, seleciona aleatoriamente um número proporcional de alunos de cada grupo. Qual é o tipo de amostragem?

a) Sistemática
b) Aleatória Simples
c) Estratificada
d) Conglomerados
e) Intencional

4.3 (R) Uma organização de saúde quer examinar a prevalência de diabetes em uma cidade. Eles possuem uma lista completa de 10.000 residentes e selecionam 500 para realizar o exame de diabetes de forma aleatória. Que técnica de amostragem foi utilizada?

a) Conglomerados
b) Estratificada
c) Sistemática
d) Aleatória Simples
e) Conveniência

4.4 (R) Uma organização governamental de estudantes está interessada em avaliar a proporção de estudantes que está a favor de mudança no tempo de duração dos cursos. Uma lista de nomes e endereços dos 645 estudantes arrolados durante o corrente semestre está disponível no escritório do oficial de registro. Usando números aleatórios de três dígitos na linha 10 da tabela de dígitos aleatórios, e movendo-se da esquerda para a direita, identifique os primeiros 10 estudantes que seriam selecionados usando a AAS.

4.5 (R) O County and City Data Book, publicado pelo Departamento de Recenseamento, lista informações dos 5139 condados de todos os Estados Unidos. Use números aleatórios de quatro dígitos a partir da última coluna da Tabela para identificar os números correspondentes a cinco condados selecionados para a amostras. Ignore os primeiros dígitos e comece com os números aleatórios de quatro dígitos 9.945, 8.364, 5.702 e assim por diante.

4.6 (R) Suponha que queremos identificar uma amostra aleatória simples de 8 dos 372 médicos em uma determinada cidade. Os nomes dos médicos estão disponíveis na organização médica local. Use a oitava coluna de números aleatórios de cinco dígitos na Tabela para identificar os 8 médicos para a amostra. Ignore os primeiros dois dígitos aleatórios em cada um dos grupos de números aleatórios de cinco dígitos. Este processo começa com o número aleatório 108 e continua de cima para baixo na coluna de números aleatórios.

Questões de concursos

4.7 (Anpec 1993, EST-Q2) Dada uma população de 10 elementos, se dela se tirar, com reposição, todas as amostras aleatórias de tamanho três possíveis, pode-se afirmar que:

(0) Ter-se-á 720 amostras.
(1) O primeiro elemento de qualquer uma das amostras é um estimador não viesado da média da população.
(2) O primeiro elemento de qualquer uma das amostras é um estimador consistente da média da população.

(3) A média das médias amostrais é igual à média da população dividida por n (tamanho da amostra).
(4) A média amostral é um estimador eficiente da média da população.

4.8 (Anpec 1993, EST-Q14) Dada uma população finita, de tamanho N, e uma amostra aleatória de tamanho n,

(0) a média amostral é um estimador não viesado da média da população somente se a amostragem for feita com reposição.
(1) a variância da média amostral será igual à variância da população sobre n somente se a amostragem for feita com reposição.
(2) δ^2 multiplicado por $\frac{N}{N-1}$ é sempre um estimador não viesado da variância da população.
(3) intervalos de confiança para a média da população podem ser obtidos, na maioria dos casos, i.e. n maior que 40, com o auxílio da distribuição normal.

Capítulo 6 Inferência Clássica

6.1 (R) Um pesquisador deseja estimar a altura média de uma população de estudantes. Ele seleciona uma amostra de 50 estudantes e encontra uma média amostral de 1,75 metros. Qual é a melhor estimativa pontual para a média da população?

a) 1.60 metros
b) 1.65 metros
c) 1.70 metros
d) 1.75 metros
e) 1.80 metros

6.2 (R) Em um estudo sobre o peso dos alunos de uma escola, uma amostra de 100 alunos apresentou um desvio padrão amostral de 4 kg. Qual é a melhor estimativa pontual para o desvio padrão da população?

a) 2 kg
b) 4 kg
c) 16 kg
d) 20 kg
e) 400 kg

6.3 (R) Os seguintes dados são de uma amostra aleatória simples (AAS): 5, 8, 10, 7, 10, 14.

a) Qual é a estimativa por ponto da média da população?
b) Qual é a estimativa por ponto do desvio padrão da população?

6.4 (R) Um levantamento para uma amostra de 150 indivíduos, produziu 75 respostas *Sim*, 55 respostas *Não* e 20 respostas *Sem opinião.*

a) Qual é a estimativa por ponto da proporção na população que respondeu *Sim*?
b) Qual é a estimativa por ponto da proporção na população que respondeu *Não*?

6.5 (F) O instituto de pesquisa OPINAS avaliou o cenário eleitoral em certa região do Brasil em uma amostra de 500 eleitores, constatando que o candidato A tem 45% das intenções de votos enquanto seu concorrente, o candidato B, tem 37%.

a) Construa o intervalo de confiança 95% para a proporção de votos do candidato A.
b) Construa o intervalo de confiança 95% para a proporção de votos do candidato B.
c) As margens de erro são iguais?

6.6 (F) Um grande conglomerado, com centenas de empresas, quer entender melhor sobre o processo de separação do lixo em suas unidades. Para isso, selecionou 100 empresas do conglomerado e verificou que 82 delas faziam a separação do lixo.

a) Qual a estimativa por ponto da proporção de empresas do conglomerado que separam o lixo?
b) Construa um intervalo de confiança de 95% para a proporção de empresas do conglomerado que separam o lixo.

6.7 (F) A pesquisa de intenção de votos do Datafolha com 3281 eleitores nos dias 14 e 15 de outubro apontou 47% de intenção de votos para a candidata Dilma Rousseff.

a) Encontre os intervalos de 84% e 95% de confiança para a verdadeira proporção de votos da candidata.
b) Quais as margens de erro dos intervalos de confiança do item anterior?

6.8 (F) Em uma amostra aleatória de 85 rolamentos de automóveis de certa fábrica, 10 apresentaram defeitos de fabricação.

a) Forneça a estimativa por ponto da verdadeira proporção de rolamentos defeituosos na fábrica.
b) Construa um intervalo de 91% de confiança para a verdadeira proporção de rolamentos defeituosos na fábrica.

6.9 (F) Em uma turma de Estatística Básica com $N = 57$, 23 alunos tiraram nota igual ou superior a 7 na P1, de um total $n = 41$ que realizaram a prova.

a) Qual a estimativa por ponto da proporção de pessoas que tiraram pelo menos a média na P1?
b) Sabendo que há 57 pessoas matriculadas e supondo que todas elas tivessem feito a prova, quantas pessoas espera-se que tirassem pelo menos 7 na P1?
c) Faça um intervalo de confiança de 85% para proporção de pessoas que tiraram pelo menos a média na P1.

6.10 (F) Em uma amostra de 5 empresas brasileiras de importação de rolamentos, constatou-se que elas gastaram R$65,000,000.00 (sessenta e cinco milhões de reais) em compra de rolamentos da China.

a) Qual a estimativa por ponto do gasto médio das importações de rolamentos de empresas do ramo no Brasil?
b) Sabendo que o desvio padrão amostral de R$1,500,000.00, encontre um intervalo de 90% de confiança para o gasto médio das importações de rolamentos de empresas do ramo no Brasil.

6.11 (R) Numa amostra aleatória de 200 eleitores, perguntou-se em quem votariam na próxima eleição para governo do estado. Sabendo que o mais votado recebeu 45% das intenções de voto, determine um intervalo de confiança das intenções de voto para esse candidato, com $1 - \alpha = 0.95$.

6.12 (R) Numa pesquisa de mercado bem conduzida, 57 de 150 pessoas afirmaram que seriam compradores de certo produto a ser lançado. Essa amostra é suficiente para estimar a proporção real de futuros compradores, com uma precisão de 0.07 e uma confiabilidade de 90%? Se a amostra for suficiente, determine a estimativa para a proporção de compradores.

6.13 (R) As regras da legislação eleitoral atualmente em vigor estabelecem que os cargos majoritários sejam disputados em dois turnos, exceto se um dos candidatos obtiver no 1º turno, mais de 50% dos votos válidos. Para as prefeituras, tal regra só é válida para as cidades com mais de 200 000 habitantes. Em uma dessas cidades, 396 pessoas de uma amostra de 720 demonstraram a intenção de votar no candidato da situação. Podemos dizer que existe uma confiança de 97% que ali a eleição será definida no primeiro turno?

6.14 (R) Uma amostra aleatória de 400 domicílios mostra-nos que 25% deles são casas de aluguel. Qual é o intervalo de confiança que podemos razoavelmente supor que seja o da proporção de casas de aluguel? Use uma confiança de 98%.

6.15 (F) (Adaptado de (Anderson et al. 2007)) Uma rádio do estado anunciou que 90% dos hotéis da Serra Gaúcha estariam lotados no final de semana do dia dos pais. A estação aconselhou os ouvintes a fazerem reserva antecipada para se hospedar na Serra nestes dias. No sábado à noite uma amostra de 58 hotéis revelou que 49 diziam “sem vagas”. Qual é a sua reação à afirmação da rádio, depois de ver a evidência da amostra? Use 5% de nível de significância.

6.16 (R) O coordenador de um curso de pós-graduação em economia está interessado em determinar a proporção de alunos matriculados que tem acesso a um computador pessoal fora da universidade. Determine uma estimativa, com 99% de confiança, se uma amostra de 150 alunos revelou que 105 têm acesso a um computador pessoal fora da universidade.

6.17 (R) O gerente de um banco em uma cidade pequena gostaria de determinar a proporção de seus correntistas que recebem salários semanalmente. Determine uma estimativa, com 90% de confiança, da proporção, na população, de correntistas do banco que são pagos por semana, se uma amostra de 145 apresentou 29 que recebiam por semana.

6.18 (R) Uma empresa de televisão a cabo deseja calcular a proporção de clientes que comprariam um guia de programação de tevê a cabo. A empresa gostaria de ter 95% de confiança de que sua estimativa esteja correta, com uma margem de erro de, no máximo, 0.05. Experiências do passado, em áreas similares, indicam que 30% dos clientes comprariam o guia de programação. Qual o tamanho mínimo necessário da amostra?

6.19 (F) O uso de drogas para a expansão da consciência foi relatado no livro As Portas da Percepção (1954) de Aldous Huxley, que inspirou o nome da lendária banda The Doors. Seguindo esta linha, o psiquiatra Stanislav Grof cunhou o termo “psicologia transpessoal”, falando dos estados ampliados de consciência.

Em 1956, iniciou uma pesquisa no Departamento de Psiquiatria da Faculdade de Medicina de Praga, onde trabalhava, e em 1967, em um instituto dos Estados Unidos. A pedido de um laboratório da Suíça testou a substância alucinógena LSD em voluntários, com o intuito de descobrir o potencial que a droga oferecia no tratamento de distúrbios psíquicos. Após, desenvolveu a técnica de respiração holotrópica, com o mesmo intuito de provocar alterações de consciência uma vez que o LSD é ilegal há muitas décadas.

a) Se em 30 dos 40 voluntários foram verificadas ondas do cérebro alteradas via eletroencefalograma, calcule a estimativa por ponto da verdadeira proporção de pessoas que apresentariam esta alteração neurológica.
b) Calcule o intervalo de 95% de confiança para a verdadeira proporção de pessoas que apresentariam a alteração neurológica citada no item anterior.

6.20 (F) Em um lugar chamado *Imaginationland*, por algum motivo são conhecidos os desvios padrão de todas as variáveis já estudadas, mas não são conhecidas as médias. Foi então realizado um estudo para estimar a média de importação de idéias entre Imaginationland e Realand. Em 20 transações observaram um total de 35 idéias.

a) Qual a estimativa por ponto do número médio de idéias por transação?
b) Sabendo que o desvio padrão é de 2 idéias, encontre um intervalo de 90% de confiança para o número médio de idéias por transação.

6.21 (R) Uma amostra aleatória simples de 50 itens resultou em uma média amostral de 32 e um desvio padrão da população de 6.

a) Forneça um intervalo de confiança de 90% para a média da população.
b) Forneça um intervalo de confiança de 95% para a média da população.
c) Forneça um intervalo de confiança de 99% para a média da população.

6.22 (R) Informa-se que um intervalo de confiança de 95% para uma média da população é de 122 a 130. Se a média é 126 e o desvio padrão é 16.07, que tamanho de amostra foi usado nesse estudo?

6.23 (R) Os seguintes dados foram coletados para uma amostra de uma população normal: 10, 8, 12, 15, 13, 11, 6, 5.

a) Qual é a estimativa pontual da média da população?
b) Qual é a estimativa pontual do desvio padrão da população?
c) Qual é o intervalo de confiança de 95% para a média da população?

6.24 (R) O Secretário de Saúde do Império Romano propôs-se a melhorar o atendimento médico à plebe. Como não há dinheiro para contratar mais médicos, ele decidiu tornar o atendimento mais eficiente. Para estimar o tempo médio gasto em cada consulta ele sorteou 64 pacientes de um hospital público aleatoriamente escolhido: essa amostra indicou que o tempo médio de atendimento era de 10 minutos, com um desvio padrão de 3 minutos. Com base nisso, determine o tempo médio de atendimento a um nível de confiança de 90%.

6.25 (R) Uma pesquisa em 17 cinemas de São Paulo, indicou que o ingresso custava, em média, US$ 5.50, com um desvio padrão de US$ 0.50. Com base nesses resultados, determine:

a) a estimativa do preço médio dos ingressos de cinema em São Paulo, em nível de confiança de 95% para a estimativa.
b) o erro máximo associado a essa estimativa.

6.26 (R) A *Good Times* mediu o tempo de duração de 50 filmes. O tempo médio obtido foi 61.8 min, com um desvio padrão de 3.5 min. Determine, com nível de confiança de 95%:

a) o erro máximo do tempo médio.
b) a estimativa intervalar.
c) o tamanho da amostra para que o erro fique limitado a 2 minutos.

6.27 (R) A Industrial ABC S/A, fabricante de lâmpadas elétricas, desejando conhecer o tempo médio de duração de seu produto, selecionou uma amostra aleatória de 10 unidades, apurando os seguintes valores, em horas: 220 – 249 - 236 – 224 – 251 – 218 – 232 – 254 – 238 – 278. Determine um intervalo de 95% de confiança para a estimação desejada.

6.28 (F) Um artigo do jornal *Materials Engineering* (1989, Vol. II, No. 4, pp. 275–281) descreve o resultado de testes de tensão em 22 ligas U-700. A carga de rompimento foi medida em megapascais (MPa), e a amostra apresentou média de 13.71 MPa e desvio padrão de 3.55 MPa.

a) Quais são as estimativas por ponto da média e variância populacionais?
b) Construa um intervalo de 98% de confiança para a verdadeira média populacional.
c) Obtenha um IC para σ^2 e outro para σ com confiança 95%. Considere $\chi^2_{0.025,21} \approx 10.28$ e $\chi^2_{0.975,21} \approx 35.48$.

6.29 (F) Um artigo de 1993 do *Transactions of the American Fisheries Society* apresentou o resultado de um estudo na investigação da contaminação por mercúrio na região da Flórida (EUA). Uma amostra de 53 peixes foi observada, de onde se calculou uma concentração média de mercúrio no tecido muscular de 0.5250 ppm e um desvio padrão de 0.3486 ppm. Encontre o intervalo de confiança de 95% para a média de mercúrio no tecido muscular dos peixes da Flórida, sabendo que $t_{52,0.025} \approx 2.007$.

6.30 (F) Para uma população normal com variância conhecida, responda:

a) Qual o nível de confiança para o intervalo $\overline{x} \pm 2.14\frac{\sigma}{\sqrt{n}}$?
b) Quais os valores de z que levam a um intervalo de 94% de confiança?

6.31 (F) (Adaptado de (Magalhães and Lima 2002)) Uma variável aleatória tem distribuição normal com desvio padrão igual a 12. Estamos testando se a média é igual ou diferente de 20. Para isso coletamos uma amostra de tamanho 100, obtendo uma média amostral de 17.4.

a) Formule as hipóteses.
b) Obtenha a região crítica e dê a conclusão para os seguintes níveis de significância: 1%, 5% e 10%.
c) Construa um intervalo com 95% de confiança para a média. Interprete.

6.32 (R) Um comprador potencial deseja estimar o valor médio das compras em uma loja de brinquedos num aeroporto. Com base em dados de outros aeroportos similares sabe-se que o desvio padrão de tais vendas é de US$ 0.80. Qual o tamanho mínimo que deveria ter a amostra aleatória se ele deseja estimar a média das vendas com uma precisão (erro) de US$ 0.25 e com uma confiança de 99%?

6.33 (R) O diretor de recursos humanos da Informatex necessita estimar o número médio de horas de treinamento para seus vendedores. Determine o tamanho mínimo de amostra necessário para tanto, admitindo um fator de erro de mais ou menos quatro horas e um desvio padrão populacional de 12 horas:

a) a um nível de confiança de 95%.
b) a um nível de confiança de 99%.
c) Compare os resultados de (a) e (b). Comente.

6.34 (R) A distribuição dos diâmetros dos parafusos produzidos por certa máquina tem distribuição normal. Uma amostra de 5 parafusos apresentou os seguintes diâmetros(em mm): 25.4 – 25.2 – 25.3 – 25.0 – 25.4. Construa um intervalo de 98% de confiança para o diâmetro médio de todos os parafusos produzidos por essa máquina.

6.35 (R) O desempenho dos vendedores de uma organização comercial foi medido através de escores de 0 a 10. Uma amostra de 200 vendedores apresentou os escores conforme tabela a seguir. Estime, com $1 - \alpha = 0.9$, a proporção de vendedores com escore mínimo 6.

Escore	# vendedores
$0 \vdash 2$	28
$2 \vdash 4$	43
$4 \vdash 6$	49
$6 \vdash 8$	41
$8 \vdash 10$	39

Escore	# vendedores

6.36 (R) Um pesquisador de mercado de uma grande empresa de eletroeletrônicos gostaria de estudar os hábitos, relativos a assistir televisão, dos moradores de certa cidade. Uma amostra de 40 moradores forneceu um tempo médio assistido por semana de 15.3 horas, desvio padrão de 3.8 horas e 27 moradores que assistem ao noticiário noturno, no mínimo 3 noites por semana.

a) Estime, com 90% de confiança, o tempo médio que se assiste TV, por semana, nesta cidade.
b) Estime, com 96% de confiança, a proporção de pessoas que assistem ao noticiário noturno, pelo menos três vezes por semana.

Se o pesquisador quiser conduzir uma pesquisa semelhante, em outra cidade

c) Qual o tamanho mínimo da amostra, se desejar 95% de confiança de não errar mais de 2 horas e partir do pressuposto de que o desvio padrão populacional é de 5 horas?
d) Qual o tamanho mínimo da amostra, se desejar 95% de confiança e um erro de 0.05 ao estimar a proporção de pessoas que assistem ao noticiário noturno, pelo menos 3 vezes por semana? Utilize como estimativa da proporção, o valor obtido na pesquisa acima.

6.37 (F) (Adaptado de (Pagano 2004)) A distribuição da pressão sanguínea diastólica na população de mulheres diabéticas segue distribuição Normal com média desconhecida. Os médicos desejam saber se esta média é a mesma da população de mulheres sem diabetes, que é 74.4 mmHg.

a) Construa as hipóteses.
b) Uma amostra de 10 mulheres diabéticas foi selecionada. A amostra apresentou média $\bar{x}_{10} = 84$ mmHg e desvio padrão $s_{10} = 9.1$ mmHg. Faça o teste bilateral para testar as hipóteses do item (a), com 5% de nível de significância.
c) Calcule o valor p do teste.
d) A conclusão teria sido a mesma se tivéssemos escolhido um nível de significância de 1%?
e) Construa um intervalo com 90% de confiança e outro com 95%. Compare. O que acontece quando aumentamos a confiança do intervalo?

6.38 (F) (Adaptado de (Anderson et al. 2007)) Na Western University, a média histórica da pontuação nos exames para obtenção de bolsas de estudo é 900. Uma amostra de tamanho $n = 200$ foi observada, de onde se calculou $\bar{x} = 935$. Presume-se ainda que o desvio padrão da população é conhecido e igual a $\sigma = 180$.

a) O vice-reitor deseja avaliar se a média histórica se modificou. Estabeleça as hipóteses.
b) Construa o intervalo de confiança de 95% sob H_0 para testar as hipóteses estabelecidas no item (a). Qual a sua conclusão?

6.39 (F) Uma empresa que fornece serviços de digitação (antiga datilografia) afirma que seus digitadores cometem, em média, não mais do que 3 erros de digitação por página. Uma amostra aleatória de 25 páginas digitadas pela empresa foi selecionada, e neste grupo foi contabilizado um total de 80 erros e um desvio padrão de 0.4 erro por página.

a) Quais as estimativas por ponto da média e do desvio padrão populacionais?
b) Defina as hipóteses.
c) Utilizando nível de significância de 5%, você aceita ou rejeita a hipótese apresentada pela empresa? Apresente o desenvolvimento, a decisão estatística (DE) e a conclusão experimental (CE).
d) Se um livro possui 150 páginas, qual o total de erros de digitação estimado?

6.40 (F) Uma fábrica que embala certo produto afirma que o conteúdo de suas embalagens contém em média 500g. Um consumidor com bastante tempo livre resolveu fazer o teste, comprando 25 embalagens do produto e pesando-as. A média dos pesos (massas, para ser fisicamente mais preciso) foi de 492g, e o desvio padrão de 30g.

a) Se o objetivo do órgão fiscalizador é decidir se a fábrica deve ou não ser multada, quais as hipóteses mais adequadas neste caso?

b) Utilizando nível de significância de 5%, a fábrica deve ser multada? Apresente o desenvolvimento, a distribuição utilizada, a estatística do teste e o valor crítico.

6.41 (F) A fábrica de automóveis WMB afirma que seus carros têm um rendimento médio de 10.4 km/L. Uma locadora, especializada em modelos da WMB observou uma amostra de 25 carros, obtendo uma média de 9.8 km/L e um desvio padrão de 2.3 km/L. Com $\alpha = 5\%$ pode-se admitir que os carros apresentam rendimento médio igual ao declarado pela fábrica?

6.42 (R) Uma operação de linha de produção foi programada para colocar 500g de detergente em pó em cada caixa de papelão. Uma amostra de caixas de papelão é periodicamente selecionada e pesada para determinar se está dentro do padrão. Se os dados da amostra levarem à conclusão de excesso de enchimento ou falta de enchimento, a linha de produção será paralisada e calibrada para se obter o enchimento apropriado das caixas.

a) Formule as hipóteses nula e alternativa que auxiliarão a decisão de paralisar a produção.
b) Comente a conclusão quando H0 não pode ser rejeitada.
c) Comente a conclusão quando H0 pode ser rejeitada.

6.43 (R) Os estadunidenses gastam uma média de 8.6 minutos por dia lendo jornais (USA Today, 10 de abril de 2015). Um pesquisador acredita que os indivíduos nas posições de gerência gastam mais do que o tempo médio nacional por dia lendo jornais. Uma amostra dos indivíduos em posições de gerenciamento será selecionada pelo pesquisador. Os dados sobre os tempos de leitura dos jornais serão usados para testar as seguintes hipóteses nula e alternativa: $H_0 : \mu \leq 8.6$ *vs* $H_1 : \mu > 8.6$.

a) Qual é o erro do tipo I? Quais são as consequências de se cometer esse erro?
b) Qual é o erro do tipo II? Quais são as consequências de se cometer esse erro?

6.44 (R) O pessoal de vendas da Carpetland tem tido vendas de US$ 8 000 por semana, em média. Paulo Ferreira, o vice-presidente da empresa, propôs um plano de compensação com novos incentivos de vendas. Paulo espera que os resultados de um período experimental de vendas lhe possibilitem concluir que o plano de compensação aumenta as vendas médias por vendedor.

a) Desenvolva as apropriadas hipóteses nula e alternativa. Qual é o erro do tipo I nessa situação? Quais são as consequências de se cometer esse erro?
b) Qual é o erro tipo II nessa situação? Quais são as consequências de se cometer esse erro?

6.45 (R) A floricultura de Jane se especializou em jardinagem com projetos-padrão para as áreas residenciais. O custo de mão de obra associado a uma determinada proposta de jardinagem está baseado no número de plantações de árvores, arbustos, etc, a serem usados no projeto. Para estimativas de custos, os gerentes usam 2h como tempo de mão de obra para se plantar uma árvore de tamanho médio com desvio padrão de 0.5 hora. Os tempos de uma amostra de 10 plantações durante o mês passado são apresentados a seguir (em horas): 1.9 1.7 2.8 2.4 2.6 2.5 2.8 3.2 1.6 2.5. Usando $\alpha = 0.05$, teste se o tempo médio de plantações de uma árvore excede 2 horas. Qual é sua conclusão e que recomendações você consideraria fazer aos gerentes?

6.46 (R) Um fabricante de conservas anuncia que o conteúdo líquido das latas de seu produto é, em média, de 2000 gramas, com desvio padrão de 40 gramas. A fiscalização de pesos e medidas investigou uma amostra de 64 latas, verificando média de 1990 g. Fixado o nível de significância em 5%, deverá o fabricante ser multado por efetuar a venda abaixo do especificado?

6.47 (R) Um fabricante de lâmpadas afirma que a duração média de seu produto é de 500 horas. Acreditando que a média seja inferior à anunciada, um comprador selecionou 10 lâmpadas, que acusou média de 490 h, com desvio padrão de 12 h. Teste, com $\alpha = 0.05$.

6.48 (R) A indústria ABC S/A, fabricante de certo equipamento eletrônico, substituiu certo componente importado pelo similar nacional. Um comprador da referida indústria supõe que tal substituição tenha diminuído a duração do produto que antes era anunciada como sendo, em média, 200 horas. Para julgar sua suposição, o comprador testou uma amostra de 10 unidades, verificando média de 197 horas, com desvio padrão de 6.32 horas. Com $\alpha = 0.05$, estabeleça a conclusão alcançada pelo comprador.

6.49 (R) Uma cadeia de lanchonetes se propõe a instalar uma nova filial se, pelo local, passarem mais de 200 carros por hora, em certo período do dia. Em 20 horas escolhidas ao acaso, passaram pelo local, no período de interesse, em média 208.5 carros com desvio padrão de 30 carros. Com $\alpha = 0.05$, a nova filial deve ser instalada?

6.50 (R) Uma companhia de seguros está disposta a iniciar uma campanha de colocação de apólices no mercado se verificar que a quantia média segurada por família da região alvo é inferior a 10000 u.m. Uma amostra casual de 20 famílias da referida região acusou média de 9600 u.m., com desvio padrão de 1000 u.m. Usando 5% de significância, decida sobre se a campanha deve ou não iniciar, admitindo normalidade para a população.

6.51 (R) Certo fabricante de parafusos anuncia que 90% do seu produto não apresenta qualquer tipo de defeito. Um comprador acredita que a percentagem de parafusos perfeitos é diferente da anunciada pelo fabricante. Para verificar tal hipótese, examinou 400 parafusos, verificando que 344 eram perfeitos. Com $\alpha = 0.02$, realize o teste correspondente.

6.52 (R) Certa organização médica afirma que um novo medicamento é de qualidade superior ao até então existente, que é 80% eficaz na cura de determinada doença. Examinada uma amostra de 300 pessoas que sofriam da doença, constatou-se que 249 ficaram curadas com o novo medicamento. Com $\alpha = 0.05$, teste a afirmação da organização.

6.53 (R) Uma agência de viagens tem um tradicional plano de férias que é oferecido a todos os possíveis clientes que procuram a agência. O índice de respostas positivas é historicamente 20%. Este ano, uma amostra de 50 potenciais clientes mostrou que 15 adquiriam o plano de férias. Teste, $\alpha = 0.06$, a hipótese de que a proporção de respostas positivas tenha aumentado este ano.

6.54 (R) Um banco de investimento calcula que 20% dos empréstimos são concedidos a pequenas empresas. Com a finalidade de aumentar esta percentagem, o governo resolveu subsidiar os juros para este tipo de empresa. Após algum tempo de vigência desta política, o banco amostrou ao acaso 40 projetos de financiamento e verificou que 12 deles se destinavam a pequenas empresas. Teste ao nível de 1% a hipótese de que a política tenha sido bem sucedida.

6.55 (R) Um consumidor de certo produto denunciou seu fabricante afirmando que este coloca no mercado uma quantidade de unidades defeituosas que supera 20% da quantidade total. Uma investigação foi conduzida com uma amostra aleatória de 50 unidades, das quais 28% acusaram defeito. Você diria que a investigação fundamenta a denúncia? Use 10% de significância.

6.56 (R) Os produtores de um programa de televisão acham que devem modificá-lo caso sua assistência regular seja inferior a um quarto dos possuidores de aparelhos de TV. Uma pesquisa foi feita em 400 domicílios, selecionados aleatoriamente, mostrando que em 80 o programa era assistido regularmente. Qual a decisão dos produtores, com 3% de significância?

6.57 (F) Um engenheiro aeroespacial deseja estimar a proporção de clientes satisfeitos com os produtos fabricado sob seu comando. Para isso realizou uma pesquisa durante duas semanas, obtendo um total de 96 respondentes. Destes, 84 declararam estarem satisfeitos com os produtos fornecidos. Se as hipóteses são $H_0 : \pi \geq 0.9$ *versus* $H_0 : \pi < 0.9$ e o valor-p é $p = 0.2171$, qual sua decisão considerando $\alpha = 0.10$? Forneça a decisão estatística e a conclusão experimental.

Questões de concursos

6.58 (F) (FEPESE - 2010 - SEFAZ-SC - Analista Financeiro) Assinale a alternativa coreta:

Uma pesquisa de opinião eleitoral foi conduzida através de amostragem casual, indicando que certo candidato a cargo majoritário é indicado como o preferido por uma proporção de 30% dos eleitores, com uma margem de erro de 2.5%, para uma confiança de 95%. Isso significa que:

(a) Há 5% de probabilidade de que o intervalo entre 27.5% e 32.5% contenha a proporção populacional de eleitores que preferem o candidato.

(b) Há 5% de probabilidade de que o intervalo entre 29.25% e 30.75% contenha a proporção populacional de eleitores que preferem o candidato.
(c) Há 95% de probabilidade de que o intervalo entre 27.5% e 32.5% contenha a proporção populacional de eleitores que preferem o candidato.
(d) Há 95% de probabilidade de que o intervalo entre 29.5% e 30.5% contenha a proporção populacional de eleitores que preferem o candidato.
(e) Há 95% de probabilidade de que o intervalo entre 29.25% e 30.75% contenha a proporção populacional de eleitores que preferem o candidato.

6.59 (Anpec 1991, EST-Q7) Com respeito aos testes de hipóteses pode-se afirmar que:

(0) O nível de significância de um teste é a probabilidade de cometer erro do tipo I, isto é, a probabilidade de aceitar a hipótese nula, quando ela é falsa.
(1) O valor $1-\beta$ é o poder do teste, onde β é a probabilidade do erro do tipo II, isto é, a probabilidade de rejeitar a hipótese nula quando ela é verdadeira.
(2) Se $x_1, \ldots, x_n$ é uma amostra aleatória de uma população normal com média μ e variância conhecida σ^2, para testar $H_0 : \mu = \mu_0$ contra $H_1 : \mu \neq \mu_0$ usa-se a distribuição t de Student.
(3) Dada uma população de indivíduos de tamanho n, deseja-se verificar se a população de empregados é de 0.5. Esta verificação pode ser feita através do teste de hipóteses: $H_0 : p = 1/2$ contra $H_1 : p \neq 1/2$ usando-se, para tanto, a distribuição normal como aproximação da binomial.
(4) Uma empresa afirma que 60% dos seus empregados são ligados à produção. Para verificar esta afirmativa, o sindicato decide usar uma amostra de 200 trabalhadores e observa que 105 deles estão ligados à produção. Ao nível de significância de 5% pode-se afirmar que o número de empregados ligados à produção é inferior a 60%.

6.60 (Anpec 1991, EST-Q11) Considere uma amostra aleatória $y_1, \ldots, y_n$ de uma variável normal Y de média μ e variância σ^2. Definem-se os seguintes estimadores:

$$\hat{y} = \frac{\sum y_i}{n}, s^2 = \frac{\sum (y_i - \bar{y})^2}{n}$$

Pode-se então afirmar que:

(0) $E(s^2) = \sigma^2$.
(1) $Var(\hat{y}) = 2s^2/n$.
(2) $\hat{y}$ é um estimador não tendencioso de μ e de variância mínima.
(3) ns^2/σ^2 tem distribuição qui-quadrado com $n-1$ graus de liberdade.
(4) $\hat{y}$ tem distribuição normal.

6.61 (Anpec 1992, EST-Q9) Considere $x_1, x_2, \ldots, x_n$ uma amostra aleatória extraída de uma população que tem distribuição Normal com média μ e variância σ^2. Pode-se dizer que:

(0) $\bar{x} = \frac{\sum x_i}{n}$ é um estimador não-viesado em μ.
(1) $s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$ é um estimador não-viesado de σ^2.
(2) $\bar{x}$ tem distribuição Normal com média μ e variância unitária.
(3) s^2 tem distribuição qui-quadrado com $n-1$ graus de liberdade.
(4) $P[-\bar{x} < \mu < \bar{x}] = 68\%$.

6.62 (Anpec 1992, EST-Q10) Suponha que X e Y sejam variáveis aleatórias independentes, com mesma média e variâncias $V(X) = 0.5$ e $V(Y) = 2$. O valor da média populacional é desconhecido e se propõe estimá-lo a partir do estimador B:

$$B = aX + (1-a)Y$$

Qual é o valor de a que produz o estimador mais eficiente? Multiplique o resultado por 5.

6.63 (Anpec 1992, EST-Q10) O intervalo de confiança permite avaliar a precisão de um estimador. Sobre ele é possível afirmar que:

(0) O nível de confiança indica a probabilidade do parâmetro populacional estar dentro do intervalo estabelecido.

(1) O tamanho do intervalo varia inversamente com o tamanho da amostra.
(2) Dado um tamanho de amostra, quanto maior o nível de confiança, menor o erro amostral permitido.
(3) O intervalo com 100% de confiança para a variância σ^2 estende-se de $-\infty$ a $+\infty$.

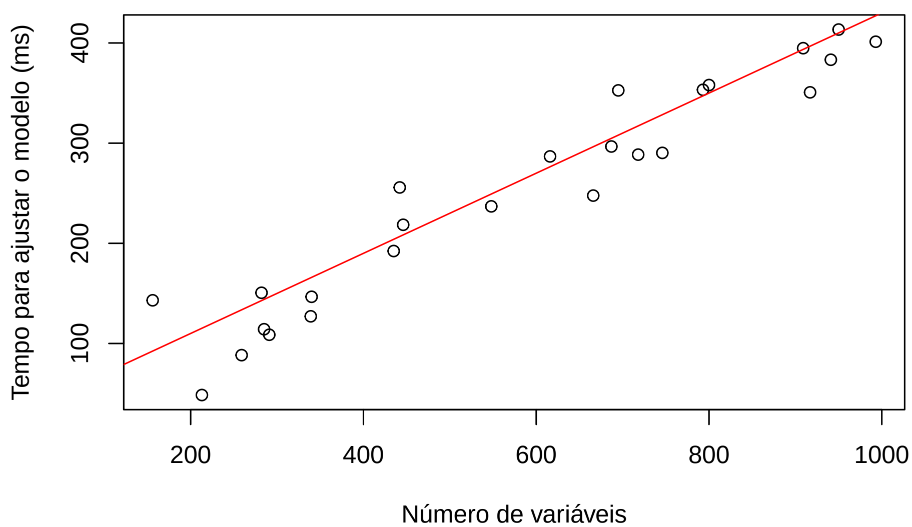
6.64 (Anpec 1992, EST-Q11) Economistas afirmam que o salário médio anual de advogado é maior do que o salário médio anual de economista.

(0) Para testar a afirmação dos economistas é necessário apenas a hipótese de que as populações originais sejam normais.
(1) A estatística do teste tem distribuição normal.
(2) A hipótese alternativa deverá ser $H_a : \mu_A \neq \mu_E$.
(3) Rejeitar a hipótese nula implica em aceitar a veracidade da afirmativa.
(4) A hipótese nula deverá ser $H_0 : \mu_A = \mu_E$.

6.65 (Anpec 2024, EST-Q6) Em uma empresa que fabrica determinado utensílio doméstico, metade das peças produzidas apresentam algum defeito, gerando muitas reclamações dos clientes. Um novo processo de produção foi adotado para reduzir a taxa de defeitos. Para testar esse novo processo, é retirada uma amostra aleatória de 5 itens. Definindo Y como a soma do número de itens produzidos sem defeito nessa amostra, o teste proposto consiste em rejeitar H_0 caso Y seja maior ou igual a 4, e não rejeitar H_0 caso contrário. Supondo que Y tenha distribuição binomial com parâmetro $p = \frac{1}{2}$, obtenha o nível de significância desse teste. Multiplique o nível de significância por 100 e assinale a parte inteira.

Capítulo 7 Correlação e Regressão

7.1 (F) Um analista de dados implementou um modelo com diferentes números de variáveis (X) e observou o tempo de ajuste deste modelo em milissegundos (Y). Observou uma amostra de tamanho $n = 25$ e obteve a equação na forma $y = 0.4x + 30$.



a) Se $r = 0.96$, teste $H_0 : \rho = 0$ *versus* $H_1 : \rho \neq 0$ com $\alpha = 5\%$. Forneça a decisão estatística e a conclusão experimental.
b) Interprete a inclinação no contexto do problema.
c) Se forem inseridas 330 variáveis, qual o tempo estimado para ajustar o modelo?
d) Se um modelo levou 200 milissegundos para ser ajustado, qual a estimativa do número de variáveis considerado no ajuste?

7.2 (R) Um agricultor plantou um pé de feijão no quintal de sua casa, anotando semanalmente sua altura. Com tais dados (abaixo), determine:

a) O coeficiente de correlação linear de Pearson;
b) A equação da reta de regressão que define o crescimento do pé de feijão;

c) A época em que o dito pé de feijão terá 60 cm de altura;
d) A altura que o pé de feijão tinha com 3.5 semanas de vida.
e) Faça o teste de significância, com $\alpha = 5\%$, para a correlação.

```
idade_semana <- c(1,2,3,4,5,6,7,8,9)
altura_cm <- c(5,12,16,22,34,38,41,45,50)
```

7.3 (R) A tabela a seguir apresenta o tempo médio de treino de um grupo de 11 atletas, com seus respectivos pontos em um teste de aptidão física.

```
tempo <- c(5,3,4,1,6,8,10,9,7,2,11)
pontos <- c(30,25,27,15,38,45,51,47,41,19,55)
```

a) Represente estes dados graficamente.
b) Calcule o coeficiente de correlação linear de Pearson.
c) Determine a equação para estimar o número de pontos em função do tempo de treino.
d) Encontre o número de pontos para um atleta que não treine e para um que treine 8h.
e) Faça o teste de significância ($\alpha = 5\%$), para a correlação.

Questões de concursos

7.4 (Anpec 1991, EST-Q8) A capacidade de produção instalada (Y), em toneladas, de uma firma, pode ser função da potência instalada (X), em 1000kW, ou da área construída (Z) em $100m^2$. Dados: $\sum X = 38$, $\sum Y = 80$, $\sum Z = 100$, $\sum X^2 = 182$ $\sum Y^2 = 736$, $\sum Z^2 = 1048$, $\sum XY = 361$, $\sum Yz = 848$

Sendo $n = 10$, pode-se afirmar que:

(0) Ao fazer uma regressão da capacidade de produção instalada em função da potência instalada ($Y = \alpha + \beta X$) obtém-se como estimativas de α e β, 2.24 e 1.52, respectivamente.
(1) O R^2 da regressão em (0) é de 0,90.
(2) Ao fazer uma regressão da capacidade de produção instalada em função da área construída ($X = \alpha + \beta Z$) obtém-se como estimativas de α e β, -2.00 e 1.00, respectivamente.
(3) O R^2 da regressão em (2) é de 0.92.
(4) Pode-se dizer que é melhor (porque tem maior correlação) usar a potência instalada (X), do que a área construída (Z) para estimar a capacidade de produção

7.5 (Anpec 1991, EST-Q14) O lucro Y de uma empresa em função do tempo, X, tem distribuição normal com média seguindo a estrutura linear de regressão $\alpha + \beta X$ e variância, por pressuposição, constante. Ajustando-se esse modelo por mínimos quadrados a uma série de 5 anos obtiveram-se os somatórios abaixo: $\sum X = 15$, $\sum Y = 38$, $\sum X^2 = 55$, $\sum Y^2 = 362$, $\sum XY = 141$. Pode-se afirmar que, exceto por erro de arredondamento,

(0) A estimativa do termo constante é -0.5.
(1) O quadrado do coeficiente de correlação múltipla R^2 é 80%.
(2) A estimativa do coeficiente de X é 1.4.
(3) A estimativa da variância do erro estocástico é 0.10.
(4) A soma dos quadrados explicada pelo modelo é 85.6.

7.6 (Anpec 1991, EST-Q15) Ainda em relação à questão anterior (7.2) pode-se concluir que, exceto por erro de arredondamento:

(0) O erro padrão da estimativa de a é igual a 0.77.
(1) O erro padrão da estimativa de b é igual a 0.10.
(2) A previsão do lucro para $X = 10$ é 26.5 milhões de reais.
(3) O coeficiente do tempo é altamente significativo produzindo um valor observado da estatística t de Student de 27.
(4) A estatística F de adequação do modelo é 729.

7.7 (Anpec 1992, EST-Q13) O custo total, C, de uma indústria e sua produção, X, têm uma relação linear do tipo $C = \beta_0 + \beta_1 X$. Para se ajustar esse modelo por mínimos quadrados ordinários é preciso assumir certas hipóteses como:

(0) A variável independente X seja aleatória.
(1) Os erros tenham média zero.
(2) A variância dos erros não seja constante.
(3) Os erros sigam uma distribuição normal.
(4) A variável independente X seja independente do temo erro.

7.8 (Anpec 1992, EST-Q14) Considere o modelo $Y_i = \alpha + \beta X_i + u_i$, estimado pelo método dos Mínimos Quadrados Ordinários. Pode-se dizer que:

(0) Se o coeficiente β estimado for igual a um ($\hat{\beta}_1$), a correlação linear entre X e Y será perfeita.
(1) A significância de β será testada tanto por meio de teste t quanto do teste F.
(2) Se o R^2 estimado for menor que 10%, o modelo deverá ser abandonado.
(3) Se for introduzida mais uma variável explicativa no modelo, o R^2 certamente não diminuirá.
(4) Se o coeficiente α estimado for inferior à unidade ($\hat{\alpha} < 1$), a reta passará pela origem.

7.9 (Anpec 1993, EST-Q3) Suponha que se tenha usado dados de 12 plantações para estimar a função de produção:

$$Y = 2.10 + 0.32X$$

$$(0.3) \quad (0.08)$$

em que Y é medido em toneladas de café por hectare de X em centenas de quilo de fertilizante por hectare. O erro padrão das estimativas s_{b0} e s_{b1} são dados entre parênteses. Pode-se afirmar que:

(0) Ao nível de 5% ambas estimativas são significantes.
(1) Se o desvio padrão da variável X (s_X) é 1 e o desvio padrão da variável Y (s_Y) é 1 então o coeficiente de correlação entre X e Y, r, é 0.32.
(2) Para fazer a análise de variância dessa regressão precisa-se conhecer apenas a variação explicada pela regressão um vez que os graus de liberdade já são conhecidos.
(3) Se o café custar $15 por tonelada e o fertilizante $3 por 100 quilos, não vale a pena ao fazendeiro usar mais fertilizante para aumentar a produção, pois o custo marginal excede a receita marginal.
(4) Um aumento de 100 quilos de fertilizante provoca um aumento de 2.42 toneladas na produção de café.

Referências

Anderson, D. R., D. J. Sweeney, T. A. Williams, and L. S. de Castro Paiva. 2007. *Estatística Aplicada à Administração e Economia*. Cengage Learning. https://www.cengage.com.br/livro/estatistica-aplicada-a-administracao-e-economia-4/.

Magalhães, M. N., and A. C. P. Lima. 2002. *Noções de Probabilidade e Estatística*. Editora da Universidade de São Paulo. https://www.edusp.com.br/livros/nocoes-de-probabilidade-e-estatistica/.

Pagano, M. 2004. *Princípios de Bioestatística*. Pioneira Thomson Learning. https://www.cengage.com.br/livro/ebook-principios-de-bioestatistica/.

Picarelli, José Rodrigues. c. 1970. *Exercícios de Estatística*. Não publicado. https://filipezabala.com/exer/picarelli_original.pdf).

World Intellectual Property Organization, WIPO -. 1883. "Paris Convention for the Protection of Industrial Property." https://www.unido.org/sites/default/files/2014-04/Paris_Convention_0.pdf.