



Departamento de Estatística - DEs
5170: Probabilidade e Estatística
1ª Lista de Exercícios

Exercício 1. *A má estatística pode muito bem resultar de formular perguntas de maneira errada ou para a pessoa errada. Explique por que os seguintes casos podem levar a dados ruins.*

- (a) *Para determinar o que a pessoa comum gasta com um relógio de pulso, uma pesquisadora entrevista somente pessoas que usam relógios de marca Rolex.*
- (b) *Para prever o resultado de uma eleição, um pesquisador entrevista pessoas que estão saindo do edifício que abriga a sede de um partido político.*
- (c) *Para estudar os padrões de gastos de indivíduos, faz-se uma pesquisa durante as três primeiras semanas de Dezembro.*
- (d) *Para determinar a renda anual média de seus alunos que se formaram há 10 anos, a secretaria de uma faculdade mandou questionários, em 2002, a todos os formandos da turma de 1992.*

Exercício 2. *O New York Times publicou um artigo que incluía as afirmativas: “Finalmente, o chocolate se coloca em seu lugar de direito na pirâmide da alimentação, em algum lugar da nobre vizinhança do vinho tinto, das frutas, dos vegetais e do chá verde. Vários estudos, relatados no Journal of Nutrition, mostraram que, após comerem chocolate, os sujeitos de teste apresentaram níveis aumentados de antioxidantes no sangue. Chocolate contém flavonóides, antioxidantes que tem sido associado à diminuição de risco de doenças do coração e infarto.” A Mars Inc., a companhia de doces, e a Chocolate Manufactures Association financiaram grande parte da pesquisa. O que há de errado com esse estudo?*

Exercício 3. *O Senado do Estado do Havaí fez audiências públicas quando estava considerando uma lei que exigia que motociclistas usassem capacetes. Alguns motociclistas testemunharam que haviam sofrido acidentes nos quais o capacete não teria sido de qualquer utilidade. Qual grupo importante não pôde testemunhar?*

Exercício 4. *Você precisa realizar um estudo para determinar o tamanho médio de uma família em seu estado. Por conveniência, você coleta dados que consistem no número de irmãs e irmãos de sua universidade. Qual o grupo de famílias é omitido por essa abordagem? Os resultados serão representativos de todas as famílias do estado?*

Exercício 5. *Classifique como verdadeiro (V) ou falso (F) e justifique quando a afirmação for falsa.*

- () Uma amostra é um subconjunto da população.*
- () A estatística descritiva é a etapa na análise que obtém-se informações da amostra a partir de medidas resumo, gráficos e tabelas, dentre outros.*
- () A estatística inferencial trata de usar as informações da população toda para tirar uma conclusão sobre uma amostra correspondente.*
- () Uma população é um conjunto de alguns resultados, respostas, medidas ou contagens que são de interesse.*
- () Dados que estão no nível nominal são quantitativos.*
- () Dados que estão no nível ordinal são qualitativos.*

Exercício 6. *Se perguntarmos a eletricitistas se a troca de um transformador é ‘muito fácil’, ‘fácil’, ‘difícil’ ou ‘muito difícil’, obteremos dados ordinais ou nominais? E se codificarmos essas alternativas em 1, 2, 3 e 4, de acordo com o grau de dificuldade?*

Exercício 7. *Classifique as seguintes variáveis como quantitativa discreta ou contínua e qualitativa ordinal ou nominal.*

- (a) Região de residência (Norte, Nordeste, Centro-Oeste, Sudeste ou Sul).*
- (b) Renda familiar, em salários mínimos.*
- (c) Área em que reside (urbana ou rural).*
- (d) Classificação socioeconômica do chefe do domicílio (muito pobre, pobre, regular, rico, muito rico).*
- (e) Número de cheques utilizados numa conta corrente.*
- (f) Número de pessoas por domicílio.*
- (g) Unidades de um artigo em um inventário.*
- (h) Peso de cada produto fabricado.*
- (i) Tempo decorrido antes da primeira falha de um dispositivo.*
- (j) Número de indivíduos, em uma área geográfica, que recebem seguro desemprego.*
- (k) Montante de vendas, em dólares.*

Exercício 8. *Um economista seleciona aleatoriamente 10 assalariados de cada um dos 50 estados americanos. Para cada estado, ele encontra a média das rendas anuais e depois soma esses 50 valores e divide por 50. Esse resultado tem chance de ser uma boa estimativa da média de todos os assalariados nos Estados Unidos?*

Exercício 9. *Suponha que estávamos trabalhando com um conjunto de n observações relativas a uma determinada variável quantitativa, onde $n > 2$. Foram calculadas a média e a mediana desses dados. Em seguida, o valor da maior entre todas essas n observações sofreu um aumento. Como resultado disso, podemos afirmar que:*

- (a) Tanto a média quanto a mediana aumentam.*
- (b) Tanto a média quanto a mediana não se alteram.*

- (c) A média não se altera e a mediana aumenta.
- (d) A média aumenta e a mediana não se altera.
- (e) Nada se pode afirmar sobre a mudança na mediana.

Exercício 10. Comente as seguintes sentenças:

- (a) Sempre a metade dos dados está abaixo da média.
- (b) Sempre existe uma, e somente uma, moda em um conjunto de dados.
- (c) A mediana é a resposta que mais se repete em um conjunto de dados.
- (d) As medidas de posição e dispersão podem ser usadas tanto para variáveis quantitativas quanto para variáveis qualitativas.
- (e) A média amostral é um valor presente em um conjunto de dados.

Exercício 11. Na companhia A, a média dos salários é 10.000 unidades e o 3º quartil é 5000.

- (a) Se você se apresentasse como candidato a funcionário nessa firma e se o seu salário fosse escolhido ao acaso entre todos os possíveis salários, o que seria mais provável: ganhar mais ou menos que 5000 unidades?
- (b) Suponha que na companhia B a média dos salários seja 7000 unidades, a variância praticamente zero e o salário também escolhido ao acaso. Em qual companhia você se apresentaria para procurar emprego?

Exercício 12. Um levantamento arguiu 2.013 adultos: “Você está satisfeito com a situação da economia dos Estados Unidos hoje?” (The Wall Street Journal, 2 de dezembro de 1997). As categorias das resposta eram insatisfeito, satisfeito e indeciso.

- (a) Qual foi o tamanho da amostra nesta pesquisa?
- (b) Os dados coletados eram quantitativos ou qualitativos?
- (c) Para um resumo dos dados para esta questão, faria mais sentido usar as médias ou as porcentagens?
- (d) Dos entrevistados, aproximadamente 28% disseram que estavam insatisfeitos com a situação da economia dos Estados Unidos. Quantos indivíduos forneceram esta resposta?

Exercício 13. O que acontece com a média, com a variância e com o desvio padrão de um conjunto de dados quando:

- (a) Cada observação é multiplicada por 2.
- (b) Soma-se 10 a cada observação.
- (c) Subtrai-se a média geral de cada observação.
- (d) De cada observação, subtrai-se a média geral e divide-se pelo desvio padrão.

Exercício 14. Considere uma amostra de tamanho $n = 15$ composta por crianças de 1 a 12 anos, sendo que a variável observada foi a altura (em cm) das crianças. A média da altura dessa amostra é 111,5 cm. Suponha que a menor altura da amostra é modificado de 85 cm para 70 cm após refazer as medições das crianças.

- (a) Qual é o valor da média após a modificação? Apresente os cálculos.
- (b) A modificação irá alterar o valor da mediana? Justifique.

Exercício 15. A espessura (mm) de 9 parafusos foi mensurada e os dados obtidos foram:

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

Calcule as medidas descritivas: mínimo, máximo, quartis, mediana, média, amplitude, desvio médio, variância, desvio padrão e coeficiente de variação.

Exercício 16. Uma indústria de componentes eletrônicos está interessada em determinar a vida útil de certo tipo de bateria. Uma amostra, em horas, é obtida e os tempos seguem abaixo:

123, 116, 122, 110, 175, 126, 125, 111, 118, 117.

- (a) Qual é o tamanho da amostra?
- (b) Encontre a média e a mediana amostrais.
- (c) Qual característica nessa amostra é responsável pela considerável diferença entre as duas?

Exercício 17. A Pensilvânia é o quinto maior produtor de árvores de Natal dos Estados Unidos, com uma colheita de 1,5 milhão de árvores em 1994. Os preços das árvores variam de US\$ 3,50 a US\$ 5,50 por unidade (The Philadelphia Inquirer, 3 dezembro de 1994). Suponha que os seguintes dados representam o preço por unidade para 16 árvores vendidas na área da Filadélfia.

3,90 4,20 3,90 5,10 4,20 4,50 4,10 5,10 4,30 4,20
4,00 5,20 4,50 4,20 4,50 5,10

Calcule as medidas de posição (média, mediana, moda e quartis) e dispersão (amplitudes total e interquartilica, variância, desvio-padrão e coeficiente de variação).

Exercício 18. Os seguintes tempos foram registrados pelos corredores de 400m e 1600m de um time de trilha de uma universidade (os tempos estão em minutos).

Tempos de 400m: 1,04 0,92 0,90 0,98 0,99
Tempos de 1600m: 4,60 4,52 4,70 4,35 4,50

Depois de ver essa amostra de corrida, um dos treinadores comentou que os corredores de 400m apresentam

tempos mais constantes. Use o desvio padrão e o coeficiente de variação para sintetizar a variabilidade nos dados. O uso do coeficiente de variação indica que a declaração do treinador é válida?

Exercício 19. Um estudo dos efeitos do tabagismo nos padrões de sono é conduzido. A medida observada é o tempo, em minutos, que se leva para dormir. Os dados obtidos são:

Fumantes: 69,3 56,0 22,1 47,6 53,2 48,1 52,7 34,4 60,2 43,8 23,2 13,8
 Não-fumantes: 28,6 25,1 26,4 34,9 29,8 28,4 38,5 30,2 30,6 31,8 41,6 21,1 36,0 37,9 13,9

- Encontre a média amostral em cada grupo.
- Encontre o desvio-padrão amostral em cada grupo.
- Encontre o coeficiente de variação amostral em cada grupo.
- Comente o tempo de impacto que o fumo aparenta ter no tempo que se leva para dormir.

Exercício 20. Para se estudar o desempenho de 4 corretoras de ações, selecionaram-se de cada uma delas amostras de ações negociadas. Para cada ação negociada, computou-se a porcentagem de lucro durante um período fixado de tempo. Os dados estão a seguir:

Corretora A: 21 48 42 37 38 25 29 32 33 33
 Corretora B: 12 42 32 28 26 26 16 18 35
 Corretora C: 58 32 46 45 50 52 56 56 37
 Corretora D: 42 44 22 32 24 37 37 32 31 28

Existe diferença entre as corretoras com relação à porcentagem de lucro durante o período estudado? Justifique.

Exercício 21. Um experimento foi delineado para determinar o efeito da taxa de fluxo de C2F6 sobre a uniformidade na queima de chapas de silicone usadas na integração de um circuito de manufatura. Três taxas de fluxo são utilizadas no experimento e os resultados da uniformidade (em %) de seis chapas avaliadas são:

C2F6	Uniformidade					
125	3,6	3,8	3,8	4,2	4,5	4,6
160	3,6	4,2	4,2	4,6	4,9	5,0
200	5,1	5,4	5,6	5,7	6,2	6,5

Você acredita que a uniformidade na queima de chapas de silicone é a mesma para as diferentes taxas de fluxo de C2F6? Justifique.

Exercício 22. Uma distribuição de frequência relativa parcial é dada na tabela a seguir.

- Qual é a frequência relativa da classe D?
- O tamanho total da amostra é 200. Qual é a frequência absoluta de cada classe?

Classe	Frequência relativa
A	0,22
B	0,18
C	0,40
D	?

Exercício 23. A tabela abaixo representa a distribuição do grau de instrução dos funcionários de uma empresa.

Grau de Instrução	%
Fundamental	40
Médio Incompleto	10
Médio Completo	25
Superior	17
Pós-graduação	8
Total	100

- Sabendo que a empresa tem 200 funcionários, quantos têm pós-graduação?
- Construa um gráfico de barras (ou o gráfico de colunas) e o gráfico de setores da distribuição do grau de instrução.
- No gráfico de setores, quantos graus tem o setor que representa cada categoria de grau de instrução?

Exercício 24. Uma auditoria de 20 declarações de imposto de renda revelou os seguintes erros de cálculos:

0, 2, 0, 0, 1, 3, 0, 0, 0, 1, 0, 1, 0, 0, 2, 1, 0, 0, 1, 0

- Identifique e classifique a variável em questão.
- Construa uma tabela mostrando as frequências absolutas e relativas de declarações com 0, 1, 2 e 3 erros de cálculos.
- Esboce esta informação através de gráficos estatísticos adequados.

Exercício 25. Os dados a seguir representam a duração da vida útil, em anos, de 30 bombas de combustível:

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

- Estabeleça a distribuição de frequência relativas, acumulada e acumulada relativa.
- Construa um histograma para a vida, em anos, das bombas de combustível. Interprete os resultados.
- Calcule a média e o desvio padrão amostrais utilizando os dados brutos e pela aproximação pela tabela e compare os resultados.

Exercício 26. Calcule a média, a mediana e a variância amostrais aproximadas para o conjunto de dados com a seguinte distribuição de frequência e construa seu histograma.

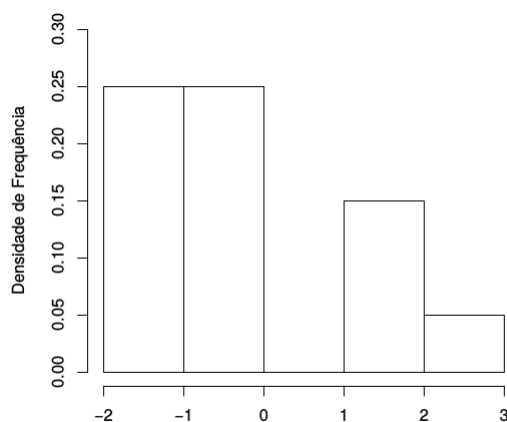
Intervalo de classe	Frequência absoluta
-10 0	3
0 10	8
10 20	12
20 30	16
30 40	9
40 50	4
50 60	2

Exercício 27. Considere a seguinte tabela:

Número médio de acidentes por mil homens/hora	Número de empresas
1,4 1,7	3
1,7 2,0	12
2,0 2,3	14
2,3 2,6	9
2,6 2,9	7
2,9 3,2	5
Total	2

- Determinar as frequências absolutas acumuladas, as frequências relativas e as frequências relativas acumuladas.
- Determinar as porcentagens.
- Determinar as porcentagens acumuladas.
- Qual a proporção de empresas que tiveram uma média de no mínimo 2,3 acidentes por mil homens/hora?
- Qual a porcentagem das empresas com um número médio de acidentes inferior a 2,0 acidentes por mil homens/hora?
- Esboce o histograma para esta tabela.

Exercício 28. O histograma a seguir mostra os lucros líquidos mensais (em milhares de R\$) de uma empresa no último ano. Esta faltando um retângulo no histograma. Qual é a sua altura?



Exercício 29. A associação de ex-alunos de uma universidade patrocina excursões mensais para associados solteiros. Os registros mostram que nos últimos quatro anos essas excursões contaram com as seguintes quantidades de ex-alunos associados solteiros.

28 51 38 27 35 33 40 37 28 33 27 33 31 41
46 40 36 53 23 33 27 40 30 33 22 37 38 36 48
22 36 45 34 26 28 40 42 43 41 35 50 31 48
38 33 39 35

- Construa a distribuição de frequência para esses dados, incluindo frequência relativa, porcentagem e porcentagem acumulada. Em seguida, construa o histograma das frequências absolutas e interprete o gráfico.
- Construa o políno de frequência e o polígono de frequência acumulada.
- 95% das excursões amostradas apresentam um número de ex-alunos associados solteiros abaixo de qual valor?

Exercício 30. Para estudar o comportamento de duas variedades de cana-de-açúcar, realizou-se um experimento do qual foram obtidos os resultados indicados a seguir:

Variedade 1: 65 68 75 76 77 78 80 80 82 86 97
Variedade 2: 88 89 90 91 92 93 95 96 97 97 99

- Calcule os quartis, Q_1 , Mediana (Q_2) e Q_3 para cada uma das variedades.
- Faça o gráfico box plot para os dados das variedades 1 e 2.
- Compare os dois conjuntos de dados através do desenho do item (b), considerando suas principais características.

Exercício 31. A seguir estão as porcentagens de encolhimento na secagem de 40 espécimes de cerâmica plástica.

20,3 16,8 21,7 19,4 15,9 18,3 22,3 17,1 19,6 21,5 21,5
17,9 19,5 19,7 13,3 20,5 24,4 19,8 19,3 17,9 18,9 18,1
23,5 19,8 20,4 18,4 19,5 18,5 17,4 18,7 18,3 20,4 20,1
18,5 17,8 17,3 20,0 19,2 18,4 19,0

- Agrupe essas porcentagens numa distribuição de frequências.
- Construa o histograma, polígono de frequência e o polígono de frequência acumulada.
- 60% das 40 espécimes amostradas apresentam porcentagem de encolhimento na secagem abaixo de qual valor?

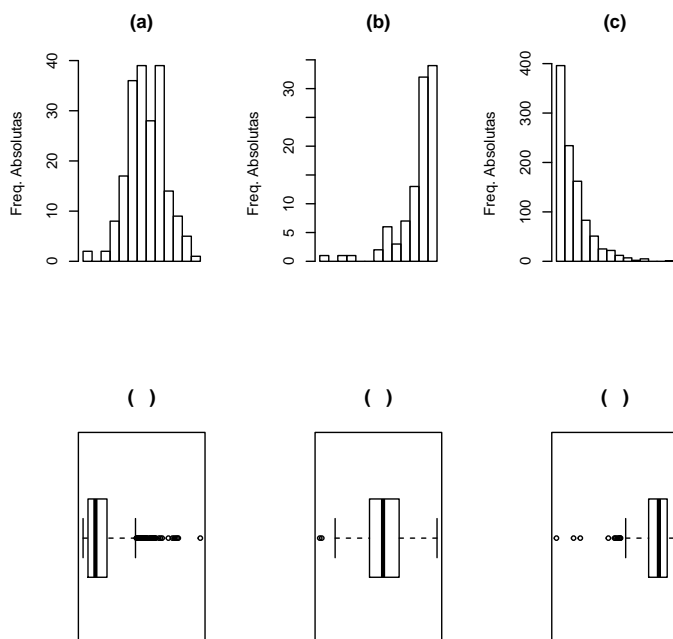
Exercício 32. A seguir são apresentados os números de situações de perigo potencial em pistas de pouso de aviões, listados em ordem para cada ano a partir de 1990 (com base em dados da Federal Aviation Administration).

281 242 219 186 200 240 275 292 325 321 421

Há uma tendência? Em caso afirmativo, qual?

Exercício 33. O corpo docente do departamento de Estatística de uma universidade consiste em três professores titulares, seis professores adjuntos, doze professores assistentes e oito professores substitutos. Disponha essa informação na forma gráfica. Identifique e classifique a variável em questão.

Exercício 34. Associe cada histograma ao boxplot que represente o mesmo conjunto de dados.



Exercício 35. Para estudar o efeito da condição do solo no crescimento de uma planta, amostras provenientes de 3 tipos de solo e seus crescimentos são classificados em 3 categorias, como mostrado na tabela a seguir.

Crescimento	Tipo de solo		
	Com pedregulho	Com areia	Argila
Fraco	16	8	14
Médio	31	16	21
Bom	18	36	25
Total	65	60	60

Para comparar a qualidade do crescimento para diferentes tipos de solo:

- Fixe as categorias da variável Crescimento e calcule o percentual correspondente a cada tipo de solo.
- Fixe as categorias da variável Tipo de solo e calcule o percentual correspondente a cada categoria da variável Crescimento.

A partir destes percentuais, o que se pode afirmar sobre o efeito do tipo de solo no crescimento da planta?

Exercício 36. No estudo de uma certa comunidade de 10000 indivíduos verificou-se que:

- A proporção de indivíduos solteiros é de 0,4.
 - A proporção de indivíduos que recebem até 10 salários mínimos é de 0,2.
 - A proporção de indivíduos que recebem até 20 salários mínimos é de 0,7.
 - A proporção de indivíduos casados entre os que recebem mais de 20 salários mínimos é de 0,7.
 - A proporção de indivíduos que recebem até 10 salários mínimos entre os solteiros é de 0,3.
- Construa a distribuição conjunta (tabela de contingência) das variáveis estado civil e faixa salarial e as respectivas distribuições marginais.
 - Você diria que existe relação entre as duas variáveis consideradas?

Exercício 37. Uma pesquisa de clima organizacional é uma pesquisa de opinião através da qual se pretende investigar o nível de satisfação e motivação dos empregados de uma determinada empresa. Suponha que foi feita uma pesquisa de clima organizacional na empresa X. Os empregados a serem entrevistados foram selecionados através de um processo de amostragem dentro de cada um dos três departamentos da empresa. Em cada um dos três departamentos, o número de entrevistas realizadas (No.) foi proporcional ao contingente total de empregados naquele departamento. As respostas foram: 'Insatisfeito' (I), 'Parcialmente Satisfeito' (PS) ou 'Satisfeito' (S). O relatório correspondente indica que, com relação às condições de trabalho oferecidas aos empregados, dentro de cada um dos três departamentos dessa empresa, as opiniões se dividem conforme indica a tabela a seguir.

Departamento	No.	I	PS	S
Comercial	75	84%	10,67%	5,33%
Pessoal	75	53,33%	40%	6,67%
Produção	200	42%	36%	22%

Quais das afirmações estão corretas e quais não estão? Por quê?

- O número de empregados insatisfeitos é maior no Departamento de Produção do que no Departamento Comercial.
- Considerando somente os empregados parcialmente satisfeitos, o percentual daqueles que estão no Departamento Pessoal é maior do que os percentuais correspondentes aos outros dois departamentos.

Exercício 38. Uma indústria utiliza um determinado reagente químico (X) para obter uma dada substância (Y). Foram realizados 5 experimentos e as quantidades do reagente e da substância em cada um deles são as seguintes:

	1	2	3	4	5
X: Reagente (mg)	6,0	5,5	5,0	7,5	12,5
Y: Substância (mg)	2,0	3,0	3,5	4,0	5,0

- Construa o gráfico de dispersão entre as variáveis X e Y.

- (b) *Determine o coeficiente de correlação de Pearson entre X e Y .*
- (c) *A quantidade de reagente utilizada está relacionada com a quantidade de substância produzida? De que forma? Justifique.*

Exercício 39. *Muitas vezes, a determinação da capacidade de produção instalada para certo tipo de indústria é um processo difícil e custoso. Como alternativa, pode-se estimar a capacidade de produção através da escolha de uma outra variável de medida mais fácil e que esteja linearmente relacionada com ela. Suponha que foram observados os valores para as variáveis: capacidade de produção instalada em toneladas (variável X), potência instalada em 1000kW (variável Y) e área construída por 100m (variável Z).*

	1	2	3	4	5	6	7	8	9	10
X	4	5	3	5	8	9	10	11	12	12
Y	1	2	1	3	3	5	5	7	6	7
Z	6	7	10	11	11	9	12	12	11	14

Com base num critério estatístico, qual das variáveis você escolheria para estimar a capacidade de produção instalada?