[1]

1 Estatística Descritiva.

1 Identifique cada uma das variáveis seguintes como quantitativa, qualitativa e como contínua, discreta, nominal, ordinal.

a) A concentração de impurezas em uma amostra de leite, em mg por litro.

b) A procedência de cada candidato ao vestibular da Unicamp em certo ano.

c) O tempo de reação de um indivíduo após submetido a certo estímulo.

d) A resposta de um indivíduo à questão: "É natural que pessoas de uma determinada raça queiram viver longe de pessoas de outras raças."

   i concordo plenamente
   ii concordo
   iii indeciso
   iv discordo
   v discordo plenamente

e) O número de moradores em cada residência de uma cidade.

f) A temperatura de certa região, em determinada época do ano.

g) A produção por hectare de determinado tipo de grão.

2 Em um estudo sobre contusões causadas durante a prática de esportes, 25 escolas de um estado brasileiro foram selecionadas, ao acaso, e entrevistadas. Foram coletados os dados abaixo, sobre o número de contusões classificadas como graves em atletas do sexo masculino para duas modalidades de esporte.

[1] Lista de exercícios seleção feita pela profa. Verónica González-López, com a contribuição do prof. Mario Gneri, Márcio Lanfredi Viola e Diego Bernardini - IMECC Unicamp .

Basquete:

1	2	4	4	7
3	3	2	4	5
2	4	3	5	3
2	4	3	6	5
5	6	4	6	5

Futebol:

1	7	7	6	1
2	6	1	7	2
1	3	2	7	5
6	1	7	4	1
5	7	6	3	2

a) Construa uma distribuição de frequências para as 50 observações.

b) Construa uma distribuição de frequências para cada modalidade.

c) Represente graficamente cada uma das distribuições.

d) Comente os resultados.

3 Os dados abaixo referem-se a dureza de 30 peças de alumínio

53.0	70.2	84.3	69.5	77.8	87.5
53.4	82.5	67.3	54.1	70.5	71.4
95.4	51.1	74.4	55.7	63.5	85.8
53.5	64.3	82.7	78.5	55.7	69.1
72.3	59.5	55.3	73.0	52.4	50.7

a) Faça uma tabela de distribuição de frequências.

b) Faça uma representação gráfica para a distribuição de frequências.

c) Calcule média, mediana e desvío padrão.

d) Apresente um histograma dos dados.

e) Faça um ramo-e-folhas, um esquema de cinco números, um box plot.

f) Comente os resultados.

4 Considere a altura (em polegadas) de 20 indivíduos

Indivíduo	1	2	3	4	5
Altura	67.75	72.25	66.25	72.25	71.25
Indivíduo	6	7	8	9	10
Altura	74.75	69.75	72.5	74	73.5
Indivíduo	11	12	13	14	15
Altura	74.5	76	69.5	71.25	69.5
Indivíduo	16	17	18	19	20
Altura	66	71	71	67.75	73.5

Considere os seguintes intervalos para as realizações da variável Altura

Intervalo	1	2	3	4	5
	[66,68)	[68,70)	[70,72)	[72,74)	[74,76]

a) Faça uma tabela de distribuição de frequências.

b) Faça uma representação gráfica para a distribuição de frequências.

c) Calcule média, variância, desvío padrão e desvío médio.

d) Apresente um histograma dos dados.

e) Faça um ramo-e-folhas, um esquema de cinco números, um box plot.

f) Comente os resultados.

5 As medidas de peso (em libras) e de cintura dos mesmos indivíduos do problema anterior são registradas

Indivíduo	1	2	3	4	5
Peso	154.25	173.25	154	184.75	184.25
Cintura	94.5	98.7	99.2	101.2	101.9
Indivíduo	6	7	8	9	10
Peso	210.25	181	176	191	198.25
Cintura	107.8	100.3	97.1	99.9	104.1
Indivíduo	11	12	13	14	15
Peso	186.25	216	180.5	205.25	187.75
Cintura	98.2	107.7	103.9	108.6	100.1
Indivíduo	16	17	18	19	20
Peso	162.75	195.75	209.25	183.75	211.75
Cintura	99.2	105.2	107	102.4	109

a) **Estudo marginal da variável Cintura**: considere os seguintes intervalos para as realizações da variável Cintura

1	2	3	4
[94,96)	[96,98)	[98,100)	[100,102)
5	6	7	8
[102,104)	[104,106)	[106,108)	[108,110)

i) Faça uma tabela de distribuição de frequências.

ii) Calcule média, variância, desvío padrão e desvío médio.

iii) Apresente um histograma dos dados.

iv) Faça um esquema de cinco números e um box plot.

b) **Estudo Conjunto**:

i) Calcule a Correlação existente entre os seguintes pares de variáveis:

Peso e Altura

Peso e Cintura

Altura e Cintura

ii) Se seu interesse for estudar a variável Peso, qual das outras duas variáveis (Altura e Cintura) poderia "explicar" melhor a variável Peso?. Justifique.

iii) Faça um diagrama de dispersão de Cintura vs Peso.

6 a) Esboce um histograma onde média, mediana e moda coincidam;

b) Esboce um histograma onde média e mediana coincidam, mas não a moda;

c) Esboce os histogramas de duas variáveis X e Y com as mesmas médias mas com variâncias diferentes.

7 Os pesos em kg de um conjunto de 10 pessoas, já ordenados de menor a maior, são:

$$21.3; 22.1; 22.8; 23.5; 24.6; 65.4; 67.2; 71.7; 76.3; 84.5$$

a) Calcule a mediana e questione a sua representatividade neste contexto;

b) Verifique a instabilidade da mediana neste caso supondo a entrada ao grupo de mais uma pessoa nas duas situações seguintes:

   i) a pessoa pesa 24 kg;

   ii) a pessoa pesa 75 kg.

(Dica: tem 2 grupos; exercício p/aula e não p/prova).

8 É feito um teste de velocidade para um grupo de 15 pessoas. Os tempos em segundos gastos em completar uma pista de 400 m são os seguintes:

$$28.7; 49.2; 49.8; 50.0; 50.1; 50.6; 71.9; 72.1;$$

$$74.1; 74.3; 74.8; 75.1; 190.8; 192.5; 196.1$$

Analise estes dados.
(Dica: vários clusters; este é exercício p/ ser trabalhado em aula, não para prova)

2 Probabilidade

1 Uma moeda e um dado são lançados. Dê um espaço amostral do experimento e depois represente-o como produto cartesiano dos dois espaços amostrais, correspondentes aos experimentos considerados individualmente.

2 Defina um espaço amostral para cada um dos seguintes experimentos aleatórios.

i) Lançamento de dois dados e uma moeda, anota-se a configuração obtida.

ii) Numa linha de produção conta-se o número de peças defeituosas num intervalo de uma hora.

iii) Investigam-se famílias com 4 crianças, anotando-se a configuração segundo o sexo.

iv) Numa entrevista telefônica com 250 assinantes, pergunta-se se o proprietário tem ou não máquina de secar roupa.

v) Um fichário de 10 nomes contém 3 nomes de mulheres. Seleciona-se ficha após ficha, até o último nome de mulher ser selecionado, e anota-se o número de fichas selecionadas.

vi) Um relógio mecânico pode parar a qualquer momento por falha técnica. Mede-se o ângulo em graus que o ponteiro dos segundos forma com o eixo imaginário orientado do centro ao número 12.

vii) De cada família entrevistada numa pesquisa, anotam-se a classe social a que pertence $\{A, B, C, D\}$ e o estado civil do chefe da família.

3 Sejam A, B e C eventos associados a um experimento aleatório. Expresse em notação de conjuntos e faça os diagramas de Venn dos seguintes eventos:

i) somente A ocorre;

ii) A e B ocorrem, mas C não;

iii) todos os três ocorrem;

iv) pelo menos um deles ocorre;

v) pelo menos dois deles ocorrem;

vi) exatamente um deles ocorre;

vii) exatamente dois deles ocorrem;

viii) nenhum deles ocorre;

ix) não mais que dois deles ocorrem;

x) no máximo 3 deles ocorrem.

4 Calcule as probabilidades dos eventos do exercício 3 supondo que

$$P(A) = 0,35; P(B) = 0,40; P(C) = 0,15; P(A \cap B) = 0,10; B \cap C = A \cap C = \emptyset.$$

5 Sejam A, B e C eventos associados a um experimento aleatório. Demonstre e interprete:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$

6 Sejam A e B eventos associados a um experimento aleatório. Demonstre que:

$$P(A \cup B) \leq P(A) + P(B).$$

7 Sejam A_1 e A_2 eventos associados a um experimento aleatório. Demonstre que:

a)
$$se P(A_1) = P(A_2) = 0 \Rightarrow P(A_1 \cup A_2) = 0;$$

b)
$$se P(A_1) = P(A_2) = 1 \Rightarrow P(A_1 \cap A_2) = 1.$$

8 Uma moeda equilibrada é lançada 3 vezes. Descreva o espaço amostral e use a definição clássica para calcular as probabilidades dos seguintes eventos:

i) duas caras ocorrem;

ii) o resultado do segundo lançamento é cara;

iii) o resultado do primeiro lançamento é igual ao do terceiro;

iv) o número de caras é igual ao de coroas.

9 Um dado equilibrado é lançado duas vezes. Descreva o espaço amostral e use a definição clássica para calcular as probabilidades dos seguintes eventos:

i) a soma dos pontos é par;

ii) a soma é ímpar;

iii) primeiro lançamento menor que o segundo;

iv) soma igual a 7;

v) soma diferente de dois;

vi) soma ≤ 4 ou soma>2;

vii) primeiro lançamento menor que o segundo lançamento e soma par;

viii) soma ímpar e igual resultado em ambos lançamentos.

10 Consideremos o conjunto $\{a, b, c, d\}$
a)Calcular o número de amostras ordenadas com reposição.
b)Calcular o número de amostras ordenadas sem reposição.

11 O prefixo telefônico de uma universidade é 452.
a)Quantos números telefônicos de sete dígitos podem-se formar?
b)Quantos números telefônicos de sete dígitos diferentes podem-se formar? Qual a probabilidade de, obtido um número ao acaso, este apresentar os sete dígitos diferentes.?

12 Temos num plano 10 pontos não alinhados (não tem três pontos na mesma linha). Quantos triângulos, com vértices em ditos pontos ficam determinados?

3 Probabilidade Condicional, Independência e Teorema de Bayes

1 Uma pessoa é submetida a uma cirurgia e morre a causa de uma reação alérgica à anestesia. Os familiares e o cirurgião entram numa disputa quanto à escolha da anestesia. O cirurgião afirma que tinha duas opções: anestesia tipo 1 (AT1) e anestesia tipo 2 (AT2). Ele afirma ter escolhido a AT1 pois as estatísticas mostram que apenas 1 pessoa em 10000 apresenta reação alérgica à AT1, salientando que não existem testes prévios válidos. O cirurgião acrescenta que a AT2 é de eliminação lenta, o que dificultaria a recuperação pós-operatória.

Os familiares alegam que o médico fez a escolha errada, pois o paciente era hemofílico e que as estatísticas mostram que 20% dos hemofílicos reagem mal à AT1 e que portanto o cirurgião deveria ter utilizado a AT2.

Se ambas as informações atribuídas às estatísticas forem verdadeiras, quem tem razão? Quem utiliza argumento falacioso? Em que consiste a falácia? Utilize a notação probabilística usual na sua argumentação.

2 400 pessoas são classificadas segundo sexo e estado civil, obtendo-se a seguinte tabela:

	Solteiro(S)	Casado(C)	Desquitado(D)	Outros(O)
Feminino(F)	150	40	10	20
Maculino(M)	50	60	40	30

a) Calcule P(S/F) , P(C/F) , P(D/F) e P(O/F) . Verifique que:

$$P(S/F) + P(C/F) + P(D/F) + P(O/F) = 1;$$

b) repita substituindo F por M ;

c) Calcule P(F/S) e P(M/S). Verifique que: $P(F/S) + P(M/S) = 1$;

d) repita substituindo S por C, D e O ;

e) apresente formalmente as distribuições de estado civil, estado civil/F e estado civil/M ;

f) apresente formalmente as distribuições de sexo, sexo/S , sexo/C , sexo/D e sexo/O ;

g) repita todo o exercício substituindo a tabela acima por uma tabela equivalente onde constem apenas probabilidades em vez de freqüências absolutas.

3 Prove que:

a) se P(A).P(B) $\neq$ 0 , então: P(A/B).P(B) = P(B/A).P(A) ;

b) se P(A).P(B) $\neq$ 0 , então: P(B/A) = P(A/B).P(B)/P(A) .

4 Uma urna contém duas bolas brancas (B) e três vermelhas (V). Duas bolas são extraídas ao acaso, uma após a outra, sendo registrada a seqüência das cores. Calcule $P(2^aB/1^aV)$, $P(2^aV/1^aV)$ e as probabilidades de cada uma das 4 seqüências possíveis de cores nas seguintes situações:

a) as bolas são extraídas com reposição;

b) as bolas são extraídas sem reposição.

5 Uma fábrica tem 3 máquinas M_1, M_2 e M_3 que produzem a mesma peça, sendo que as mesmas respondem por 20%, 50% e 30% da produção total, respectivamente. Também é conhecida a proporção de peças defeituosas produzidas por cada uma delas: 15% na M_1, 2% na M_2 e 20% na M_3.

a) Calcule a percentagem global de peças defeituosas;

b) Se uma peça for defeituosa, qual é a probabilidade de que tenha sido produzida pela M_1, M_2 ou M_3?

c) Calcule

$$\sum_{j=1}^{3} P(defeituosa/M_j) \quad e \quad \sum_{j=1}^{3} P(M_j/defeituosa).$$

Uma soma vale 1 e a outra não. Explique o motivo;

d) Quanto vale

$$\sum_{j=1}^{3} P(M_j/boa)?$$

Argumente sem fazer a conta.

6 Verifique cuidadosamente as demonstrações dos 3 resultados seguintes e observe onde é utilizada cada uma das hipóteses.

Proposição A
H) Ω espaço amostral; $B, A_1, A_2, ..., A_n$ eventos tais que $\{A_1, A_2, ..., A_n\}$ é partição de Ω.

T)

$$P(B) = \sum_{j=1}^{n} P(B \cap A_j).$$

Demonstração: a tese decorre dos seguintes fatos:

$$B = \bigcup_{j=1}^{n} (B \cap A_j)$$

e se $i \neq j$, vale que

$$(B \cap A_i) \cap (B \cap A_j) = \emptyset.$$

Proposição B
H) Ω espaço amostral; $B, A_1, A_2, ..., A_n$ eventos tais que $\{A_1, A_2, ..., A_n\}$ é partição de Ω e além disso, $P(A_j) > 0$ se $1 \leq j \leq n$.

T) Fórmula da Probabilidade Total:

$$P(B) = \sum_{j=1}^{n} P(B/A_j)P(A_j).$$

Demonstração: substitua na tese do resultado anterior $P(B \cap A_j)$ por $P(B/A_j)P(A_j)$.

Proposição C
H) Ω espaço amostral; $B, A_1, A_2, ..., A_n$ eventos tais que $\{A_1, A_2, ..., A_n\}$ é partição de Ω e além disso, $P(A_j) > 0$ se $1 \leq j \leq n$, e $P(B) > 0$.

T) Para todo k tal que $1 \leq k \leq n$, vale que: (Fórmula de Bayes)

$$P(A_k/B) = \frac{P(B/A_k)P(A_k)}{\sum_{j=1}^{n} P(B/A_j)P(A_j)}.$$

Demonstração: observe que $P(A_k/B)P(B) = P(B/A_k)P(A_k)$, e depois isole $P(A_k/B)$ e use a fórmula da probabilidade total.

7 30% dos usuários de uma biblioteca universitária são alunos da graduação, 38% são alunos da pós e 32% professores. A consulta a livros estrangeiros é de 25%, 50% e 80% nas três categorias de usuários, respectivamente.

a) Qual é a probabilidade de que um usuário qualquer utilize um livro em português?

b) Se um usuário retirou um livro em português, calcule a probabilidade de que seja aluno da graduação, da pós ou que seja professor.

8 Sejam A e B eventos de Ω tais que $P(B) > 0$. Nestas condições, mostre que são equivalentes:

a) A e B são independentes;

b) P(A/B) = P(A).

9 Sejam A e B eventos de Ω. Mostre que as seguintes afirmações são todas equivalentes:

a) A e B são independentes;

b) A e B^c são independentes;

c) A^c e B^c são independentes;

d) A^c e B são independentes.

10 Mostre que:

a) Se $P(A) = 0$ e B é um evento qualquer, então A e B são independentes ;

b) Se $P(A) = 1$ e B é um evento qualquer, então A e B são independentes ;

c) Os eventos D e D^c são independentes se e somente se $P(D) = 0$ ou $P(D) = 1$;

d) Ache uma condição para que um evento E seja independente dele mesmo.

11 Uma moeda é jogada 3 vezes.

a) Ache uma fórmula para a probabilidade da seqüência (cara, coroa, cara);

b) Repita $a)$ no caso em que $P(\{cara\}) = p$, a mesma em todas as jogadas assumindo que os resultados de jogadas diferentes são independentes.

12 Uma urna contém duas bolas brancas (B) e três vermelhas (V). Duas bolas são extraídas ao acaso, uma após a outra, sendo registrada a sequência das cores. Considere cada uma das perguntas nas duas situações seguintes:

a) as bolas são extraídas com reposição;

b) as bolas são extraídas sem reposição.

Perguntas:

1) Calcule $P(1^aB/2^aV)$ e $P(1^aV/2^aV)$;

2) Os eventos $\{2^aV\}$ e $\{1^aB\}$ são independentes?

13 30% dos empregados de uma empresa são mulheres e o restante homens; 9% das pessoas são mulheres e fumantes, 59% das pessoas são homens e não fumantes. Calcule:

a) $P(\{mulher \;\; e \;\; fumante\})$;

b) $P(\{homem \;\; e \;\; fumante\})$;

c) probabilidade de um homem ser fumante;

d) probabilidade de um homem ser não fumante;

e) probabilidade de um fumante ser homem.

14 30% dos empregados de uma empresa são mulheres e o restante homens; 3/10 das mulheres são fumantes, 11/70 dos homens são fumantes. Calcule:

a) $P(\{mulher \;\; e \;\; fumante\})$;

b) $P(\{homem \;\; e \;\; fumante\})$;

c) probabilidade de um homem ser fumante;

d) probabilidade de um homem ser não fumante;

e) probabilidade de um fumante ser homem.

15 30% dos empregados de uma empresa são mulheres e o restante homens; 9% das pessoas são mulheres e fumantes, 11/70 dos homens são fumantes. Perguntas:

a) $P(\{mulher \quad e \quad fumante\})$;

b) $P(\{homem \quad e \quad fumante\})$;

c) probabilidade de um homem ser fumante;

d) probabilidade de um homem ser não fumante;

e) probabilidade de um fumante ser homem.

16 Suponha que a probabilidade de viver 70 ou mais anos é 0.6 e que a probabilidade de viver 80 ou mais anos é 0.2. Se uma pessoa faz 70 anos, qual é a probabilidade de que comemore o aniversario número 80?.

17 Considere uma urna com 3 bolas brancas e 7 bolas vermelhas. Duas bolas são retiradas da urna uma depois da outra sem repor a primeira delas na urna antes da retirada da segunda.
Assuma a seguinte notação: B_1V_2 representando que foi retirada uma bola branca na primeira retirada e uma bola vermelha na segunda. Calcule as seguintes probabilidades

$$P(B_1B_2),\ P(V_1B_2),\ P(B_1V_2),\ P(V_1V_2).$$

Considere que se faz mais uma extração de bolas da urna, recolocando na urna a segunda bola extraída anteriormente e calcule $P(B_1V_2B_3)$, onde $B_1V_2B_3$ representa que foi extraída branca na primeira, vermelha na sengunda e branca na terceira. Compute ainda,

$$P(B_1B_2B_3),\ P(B_1B_2V_3),\ P(V_1B_2B_3).$$

18 Suponha que se testam os chips para um circuito integrado e que a probabilidade de que sejam declarados com falhas quando efectivamente as tem é 0.95, sendo que a probabilidade de que sejam declarados em bom estado se efectivamente estão em bom estado é 0.97. Se 0.5% dos chips apresentam falhas, qual é a probabilidade de que um chip que foi declarado com falhas seja bom?

19 Um restaurante popular apresenta apenas dois tipos de refeições: salada completa ou um prato à base de carne. 20% dos fregueses do sexo masculino preferem salada, 30% das mulheres escolhem carne, 75% dos fregueses são homens. Considere os seguintes eventos:
H: freguês é homen, M: freguês é mulher,
A: freguês prefere salada, B: freguês prefere carne.
Calcular:

$$P(A|H),\ P(B|M),\ P(M|A).$$

20 Na tabela seguinte, os números que aparecem são probabilidades relacionadas com a ocorrência de $A, B, A \cap B$, etc. Assim
$P(A) = 0.10$, enquanto que $P(A \cap B) = 0.04$. Verifique se A e B são independentes.

	B	B^c	
A	0.04	0.06	0.10
A^c	0.08	0.82	0.90
	0.12	0.88	1.00

21 Reconsidere o problema 6, e usando as idéias do exercício 7 verifique se a escolha do prato depende do sexo do freguês.

4 Variáveis aleatórias discretas

1 A seguir aparece um estudo sobre o número de filhos dos 20 funcionários de uma empresa. Há 1 funcionário sem filhos, 5 com 1 filho, 9 com 2, 4 com 3 e 1 com 5. Calcule a média e variância destes dados e compare com a definição formal de média e variância de uma variável aleatória discreta.

2 Um alvo é feito com uma tábua quadrada pintada de branco, com exceção de um círculo no seu centro que é pintado de preto. As regras de uma prova são definidas da seguinte forma: o atirador que acertar no centro preto ganha 18 pontos, se acertar na parte branca da tábua ganha 8 pontos e se não acertar na tábua perde 2 pontos.

a) Um atirador atira no alvo: defina formalmente o espaço dos resultados deste experimento e a variável aleatória número de pontos;

b) O desempenho do atirador pode ser assim resumido:
$P\{acertar\ \ no\ \ centro\} = 0,2$ e
$P\{acertar\ \ na\ \ parte\ \ branca\} = 0,7$;
calcule média e variância do número de pontos para o atirador.

3 Seja X uma v.a. tal que: $P(\{X = -1\}) = 0.2$; $P(\{X = 0\}) = 0,1$ e $P(\{X = 6\}) = 0,7$.

a) Ache as funções de probabilidade das v.a.s: $Y = 3X + 2$, $Z = (-2)X + 1$, $U = X^2$ e $V = X^3$;

b) Verifique que $E(3X + 2) = 3E(X) + 2$ e que $E((-2)X + 1) = (-2)E(X) + 1$;

c) Verifique que $Var(3X + 2) = 9Var(X)$ e que $Var((-2)X + 1) = 4Var(X)$;

d) Generalize as observações dos items b) e c) para X qualquer v.a. discreta que assume um número finito de valores e para qualquer função do tipo $W = a + bX$;

e) Seja S uma v.a. tal que: $P(\{S = -1\}) = P(\{S = 7\}) = 0.3$ e $P(\{S = 0\}) = P(\{S = 1\}) = 0.2$; ache a função de probabilidade de S^2.

4 Seja X uma variável discreta assumindo um número finito de valores e $h : R \to R$ uma função qualquer. Utilize a experiência dos itens a) e e) do exercício anterior para dizer como construiria a função de probabilidade de $Y = h(X)$.

5 X é uma variável aleatória cuja função de distribuição acumulada F é dada por:

$$F(x) = 0 \;\; se \;\; x < -3;$$

$$F(x) = 0,2 \;\; se \;\; -3 \leq x < 4;$$

$$F(x) = 0,9 \;\; se \;\; 4 \leq x < 8 \;\; e$$

$$F(x) = 1 \;\; se \;\; x \geq 8.$$

a) Use F para calcular as probabilidades dos seguintes conjuntos: $\{3 < X \leq 7\}$; $\{3 \leq X \leq 7\}$; $\{3 \leq X < 7\}$; $\{3 < X < 7\}$; $\{-3 < X \leq 5\}$; $\{-3 \leq X \leq 5\}$; $\{-3 \leq X < 5\}$; $\{3 < X < 5\}$; $\{X \leq 6\}$; $\{X < 6\}$; $\{X \leq 4\}$; $\{X < 4\}$; $\{X > 3\}$; $\{X \geq 3\}$; $\{X > 4\}$; $\{X \geq 4\}$; $\{X > 10\}$; $\{X \geq 10\}$; $\{X > 5\}$; $\{X \geq -5\}$; $\{X < -11\}$, $\{X \geq 20\}$;

b) Ache a função de probabilidade de X e com ela responda às perguntas do item a).

c) Calcule E(X) e Variância(X).

6 Seja X uma variável aleatória discreta com $P(X = 0) = 0.25$, $P(X = 1) = 0.125$, $P(X = 2) = 0.125$, $P(X = 3) = 0.5$. Apresentar o gráfico a função de probabilidade e a função de distribuição acumulada. Calcular o valor esperado, a moda e a mediana de X. Calcular a variância de X. Determine as seguintes probabilidades:

$$P(0 < X < 1),\; P(X \leq 2),\; P(X > 3),\; P(X > 2.5).$$

7 Dada a função de distribuição acumulada

$$F(x) = \begin{cases} 0 & \text{se} \;\; x < 1 \\ 0.1 & \text{se} \;\; 1 \leq x < 2 \\ 0.3 & \text{se} \;\; 2 \leq x < 3 \\ 0.7 & \text{se} \;\; 3 \leq x < 4 \\ 0.8 & \text{se} \;\; 4 \leq x < 5 \\ 1 & \text{se} \;\; 5 \leq x \end{cases} .$$

Calcule a função de probabilidade da variável cuja $f.d.a.$ é $F(\cdot)$. Calcule ainda o valor esperado, a moda, a mediana e a variância de X. Determine as seguintes probabilidades:

$$P(1 \leq X < 2),\ P(X = 4),\ P(X > 3),\ P(X \leq 4).$$

8 Com dados do último censo, a assistente social de um centro de saúde constatou que para as famílias da região, 20% não têm filhos, 30% têm um filho, 35% têm dois e as restantes se dividem igualmente entre três, quatro ou cinco filhos. Determine a função de distribuição acumulada da variável N: número de filhos, e responda: se uma família é escolhida aleatóriamente nessa região qual a probabilidade de que o número de filhos nessa família seja maior o igual a 2?. Calcule o valor esperado e a variância da variável N.

9 Um sinal consite em uma série de vibrações de magnitude X, tendo os valores 1,0,-1, cada um com probabilidade 1/3. Um ruido consiste em uma série de vibrações de magnitude Y, tendo os valores 2,0,-2 com probabilidades 1/6,2/3,1/6, respectivamente. Se ruídos e sinais são combinados, a soma consiste em vibrações de magnitude $Z = X + Y$.
a)Construir e apresentar o gráfico a função de probabilidade para Z, calcular sua média e variância, admitindo a independência entre ruído e sinal.
b)Construir e apresentar o gráfico a função de distribuição acumulada para Z, F_Z, calcular $F_Z(1), F_Z(-1.5)$. Achar um valor z tal que $F_Z(z) = 11/18$, calcular o menor valor z tal que $F_Z(z) = 11/18$.
c)Um amplificador de vibrações permite a captação da magnitude $2Z$, determine a função de probabilidade, a acumulada, o valor esperado e a variância desta nova variável.

5 Distribuição Binomial

1 Joga-se uma moeda 3 vezes e observa-se a sequência de caras e coroas obtida.

a) Construa Ω={espaço dos resultados associado ao experimento};

b) Calcule a probabilidade de cada $\omega \in \Omega$ sob as seguintes hipóteses:

I) p = probabilidade de $\{cara\}$ é a mesma em todas as jogadas;

II) eventos associados a conjuntos disjuntos de jogadas são independentes.

c) defina explicitamente a variável aleatória X = 'número de caras obtido nas 3 jogadas', e observe que a probabilidade definida em Ω no item anterior induz uma probabilidade em R através de X.

2 Uma moeda cuja probabilidade de cara é $0,4$ é jogada 5 vezes, sendo satisfeitas as condições I) e II) enunciadas no Exercício 1. Compare os seguintes eventos e calcule suas probabilidades:

a) a seqüência $\{(CKCCK)\}$, onde C denota cara e K coroa;

b) {são obtidas exatamente 3 caras nas 5 jogadas}.

3 Sejam Ω o espaço de resultados de um experimento E e A um evento fixo de Ω. O experimento E é repetido n vezes e supomos verdadeiras as seguintes hipóteses:

I) p = probabilidade de A é a mesma em todas as repetições;

II) eventos associados a conjuntos disjuntos de repetições são independentes.

Seja X a variável aleatória 'número de ocorrências de A nas n repetições'. Calcule $Probabilidade(\{X = j\})$ para todo inteiro j tal que $0 \leq j \leq n$. Calcule E(X) e Var(X).

4 Uma prova consiste em 25 perguntas de tipo múltipla escolha. Cada questão tem 5 respostas, sendo que apenas uma delas é verdadeira. A nota X é igual ao número de respostas corretas. Uma pessoa lança um dado equilibrado e indica a resposta cujo número aparece na face de cima do dado (se sair {6} o lance é desconsiderado).

a) Qual é a distribuição da variável 'nota' nestas condições? Veja se no contexto do problema são válidos os supostos nos quais o modelo utilizado se baseia;

b) Calcule as probabilidades dos seguintes eventos utilizando o modelo escolhido em a): $\{2 \leq X < 4\}$, $\{3 \leq X < 6\}$, $\{4 \leq X < 8\}$, $\{1 < X < 4\}$, $\{14 \leq X < 22\}$, $\{X \geq 6\}$, $\{X \geq 10\}$, $\{X \leq 1\}$, $\{X < 1\}$, $\{X \leq 20\}$, $\{X < 20\}$, $\{X \geq 1\}$, $\{X > 4\}$, $\{X \geq 23\}$;

c) Quais das hipóteses do modelo deixariam de ser satisfeitas se a pessoa respondesse seriamente em vez de usar o esquema acima descrito?

5 Uma companhia de seguros vendeu apólices a 20 pessoas da mesma idade e condições de saúde. De acordo com as tábuas atuariais, a probabilidade de que uma pessoa nas condições dos assegurados sobreviva 10 anos à data dos contratos é de 0,9. Calcule as probabilidades dos seguintes eventos:
a) todas as pessoas sobrevivem;
b) nenhuma sobrevive;
c) sobrevivem ao menos 5 pessoas;
d) sobrevivem ao menos 15 pessoas;
e) morrem exatamente 3 pessoas;
f) morrem no máximo 2 pessoas;
g) morrem no mínimo 5 pessoas.
Calcule o número médio de sobreviventes e número médio de mortos e também as variâncias do número de mortos e do número de sobreviventes aos 10 anos do contrato.

6 Uma urna contém 50 bolas, sendo 20 brancas e 30 vermelhas. São extraídas 10 bolas, uma após outra, com reposição. Calcule as probabilidades dos seguintes eventos:
a) o número de bolas vermelhas extraídas é igual a 4;
b) o número de bolas brancas extraídas é igual a 1;
c) pelo menos duas bolas vermelhas são extraídas;
d) no máximo 3 bolas vermelhas são extraídas.
Qual é o número médio de bolas brancas (vermelhas) extraído? Quais são as variâncias do número de bolas brancas e vermelhas extraído? Analise a aderência às hipóteses do modelo utilizado para responder as perguntas acima caso as extrações sejam sem reposição.

7 Um comerciante deseja comprar um lote de 200 mesas a uma fábrica. O lote oferecido tem 10 mesas defeituosas (mas o comerciante desconhece este fato). O comerciante adota a seguinte regra de decisão: ele observará uma amostra de 20 mesas escolhida por sorteio e aceitará o lote se ele tiver até 2 mesas defeituosas. Qual é a probabilidade do comerciante aceitar o lote nas condições acima detalhadas?
Observação: nas situações reais a amostragem é feita sem reposição, mas para encarar este problema precisaremos estudar a distribuição hipergeométrica. Por enquanto ficamos devendo.

8 Sabe-se que quando a distribuição é perfeitamente simétrica a média e a mediana coincidem e são ambas iguais ao centro de simetria. É o que acontece com a binomial(n,1/2) para qualquer n natural. Mais ainda, neste caso podemos observar que também a moda coincide com o centro de simetria. Dê um exemplo de uma distribuição discreta simétrica onde a moda seja diferente da média e da mediana (Dica: considere uma distribuição discreta simétrica que assuma 3 valores).
Dê um exemplo de uma distribuição discreta em que média, mediana e moda sejam todas diferentes.

9 Considere os gráficos da distribuição binomial(n,1/10) para n igual a 5, 10, 20, 30, 60 e 100. Observe que para n=5 a distribuição é totalmente assimétrica e que a medida que n cresce a assimetria diminui, sendo que para n=100 ela é quase inexistente. Justifique a afirmação da simetria da binomial(100,1/10), mesmo levando em conta a longuíssima cauda a direita (isto é, explique porque é possível desconsiderar tal cauda).

10 Sejam X e Y variáveis aleatórias com distribuições $binomial(n, p)$ e $binomial(n, 1-p)$, respectivamente. Prove que:
para todo inteiro j, $0 \leq j \leq n$, vale que

$$P(\{Y = j\}) = P(\{X = n - j\}).$$

Verifique este fato nos gráficos, comparando a binomial(20,p) com a binomial(20,1-p), para p igual a 0,2; 0,3 e 0,4.
Prove também que

$$E(X) + E(Y) = n \ \ e \ \ que \ \ Var(X) = Var(Y).$$

6 Modelos discretos

1 Discuta a validade do modelo Uniforme Discreto nos seguintes casos:
a)A escolha de um aluno que vai representar a classe junto a direção da escola.
b)O dia da semana em que ocorrem mais acidentes de trabalho numa indústria.
c)O mês do ano com maior número de enchentes na cidade de São Paulo.

2 Sendo X uma variável aleatória seguindo o modelo Uniforme Discreto, com valores no conjunto $\{1, 2, \cdots, 10\}$, pergunta-se:
a)$P(7 \leq X)$
b)$P(3 < X \leq 7)$
c)$P(X < 2\, ou\, 8 \leq X)$
d)$P(5 \leq X\, ou\, 8 < X)$
e)$P(X > 3\, e\, X < 6)$
f)$P(X \leq 9 | 6 \leq X)$

3 Discuta a validade do modelo Binomial nos seguintes casos:
a)Dos alunos de uma grande universidade, sorteamos 5 e contamos quantos se declaram usuários de drogas.
b)Escolhemos 20 lâmpadas ao acaso na prateleira de um supermercado, sendo 10 de uma fábrica e 10 de outra. Contamos o número total de defeituosas.
c)Quinze automóveis 0 km de uma mesma marca e tipo são submetidos a um teste anti-poluição e contamos o número deles que passaram no teste.
d)Um motorista é submetido a um teste em que deve estacionar seu veículo num pequeno espaço (isto é popularmente chamado de fazer baliza). Em 10 tentativas, contamos o número de vezes em que o motorista estacionou corretamente.

4 Uma certa doença pode ser curada através de procedimento cirúrgico em 80% dos casos. Dentre os que tem essa doença, sorteamos 15 pacientes que serão submetidos a cirurgia. Fazendo alguma suposição adicional que julgar necessária, responda qual é a probabilidade de:
a)Todos serem curados?.

b)Pelo menos dois não serem curados?.
c)Ao menos 10 ficarem livres da doença?.

5 Bactérias de certa classe aparecem na água a razão de 0,8 por cm^3. Qual é a probabilidade de que em 5 cm^3 de água tenhamos:
a) no mínimo duas bactérias;
b) pelo menos 13 bactérias;
c) nenhuma;
d) no máximo sete.

6 Um digitador comete 0.5 erros por folha em média ao transcrever um texto. Qual é a probabilidade de que num texto de 15 páginas cometa 8 ou mais erros?

7 A taxa de suicídios num certo país é de 1 para cada 250.000 habitantes por semana.
Considere uma cidade de 500.000 habitantes:
a) calcule a probabilidade de ter 6 ou mais suicídios numa semana?
b) utilizaria o mesmo modelo se em vez de suicídios se tratasse de dengue? Justifique.

8 Os trabalhadores de certa fábrica sofrem em média dois acidentes por mês. Calcule as probabilidades dos seguintes eventos:
a) ocorrem 5 acidentes ou menos num período de um mês (2 meses, 3 meses);
b) ocorrem 8 ou mais acidentes num período de um mês (2 meses, 3meses);
c) $2 \leq$ número de acidentes < 5 no mês de abril e também em junho.

9 A variável aleatória Y tem densidade Poisson com parametro $\lambda = 2$. Obtenha:
a)$P(Y < 2)$
b)$P(2 \leq Y < 4)$
c)$P(Y > 0)$
d)$P(Y = 1|Y < 3)$

10 A aplicação de fundo anti-corrosivo em chapas de aço de 1 m^2 é feita mecanicamente e pode produzir defeitos (pequenas bolhas na pintura), de acordo com uma variável aleatória Poisson de parâmetro $\lambda = 1$ por m^2. Uma chapa é sorteada ao acaso para ser inspecionada, pergunta-se

a probabilidade de:
a)Encontrarmos pelo menos 1 defeito.
b)No máximo 2 defeitos serem encontrados.
c)Encontrar de 2 a 4 defeitos.
d)Não mais de um defeito ser encontrado.

11 Um banco de sangue necessita sangue do tipo 0-Rh negativo. Suponha que a probabilidade de uma pessoa ter este tipo de sangue seja 0.10. Doadores permanentes chegam ao hemocentro para fazer sua doação rotineira. Calcule as probabilidades de que o primeiro doador com sangue do tipo 0-Rh negativo seja:
a) o primeiro a chegar;
b) o segundo;
c) o quarto;
d) o sétimo.

12 No contexto do exercício anterior, calcule:
a) probabilidade de que o primeiro doador com sangue no grupo 0-Rh negativo apareça a partir do quarto doador;
b) probabilidade de que o primeiro doador com sangue no grupo 0-Rh negativo apareça no máximo em 5 tentativas.

13 Um supermercado vende uma caixa com 20 lâmpadas, das quais 4 são inúteis e as restantes boas. Um comprador decide testar 5 das lâmpadas (obviamente sem reposição) escolhidas ao acaso e comprar a caixa caso haja no máximo duas defeituosas entre as lâmpadas testadas. Qual é a probabilidade de comprar a caixa? Ache a distribuição do número de itens defeituosos.

14 Uma urna contém bolas vermelhas (V) e brancas (B). São extraídas 5 bolas da urna.
Calcule as probabilidades de extrair 2 vermelhas e 3 brancas nas seguintes condições:
a) a urna tem 4V e 6B, extração sem reposição (resposta: $10/21 \approx 0{,}47619$);
b) a urna tem 8V e 12B, extração sem reposição (resposta: $385/969 \approx 0{,}39732$);
c) a urna tem 16V e 24B, extração sem reposição (resposta: $10120/27417 \approx 0{,}3691$);

d) a urna tem 48V e 72B, extração sem reposição (resposta: 934360/2646917 $\approx$ 0,3530);
e) a urna tem 40% das bolas V e 60% B, extração com reposição (resposta: 0,3456).

15 A variável H segue o modelo Hipergeométrico com parâmetros $N = 10$, $n = 5$ e $r = 4$. Determine:
a)$P(H = 2)$
b)$P(H \leq 1)$
c)$P(H > 0)$

16 Por engano 3 peças defeituosas foram misturadas com boas formando um lote com 12 peças no total. Escolhendo ao acaso 4 dessas peças, determine a probabilidade de encontrar:
a)Pelo menos 2 defeituosas.
b)No máximo uma defeituosa.
c)No mínimo 1 boa.

17 Suponha que a probabilidade de que um item produzido por uma máquina seja defeituoso é de 0.2. Se 10 itens produzidos por esta máquina são selecionados ao acaso, qual é a probabilidade de que não mais do que um defeituoso seja encontrado?. Use a Binomial e a Poisson e compare os resultados.

18 Numa central telefônica, o número de chamadas chega segundo uma distribuiçao de Poisson, com media de 8 chamadas por minuto. Determinar qual é a probabilidade de que num minuto se tenha:
a)10 ou mais chamadas;
b)menos do que 9 chamadas;
c)entre 7 (inclusive) e 9 (exclusive).

19 Numa fábrica de pregos sabe-se que a proporção de itens defeituosos é igual a 0.1. A produção mensal é de 100.000 artigos/mês. Qual é a probabilidade de que uma amostra de tamanho 4 dos artigos produzidos num mês contenha:
a) nenhum defeituoso;
b) exatamente um defeituoso;
c) exatamente dois defeituosos;
d) não mais do que 2 defeituosos;
Calcule a esperança e variância do número de defeituosos na amostra.

20 Entre as 14:00 e 17:00 horas de um dia útil passam por um pedágio em média 150 carros por hora. Calcule:

a) probabilidade de que passem até 4 carros entre 15.30 e 15.32;

b) probabilidade de que passem até 4 carros entre 16.48 e 16.50;

c) probabilidade de que passem exatamente 3 carros entre 14.16 e 14.17;

d) probabilidade de que passem até 4 carros entre 16.00 e 16.01;

e) probabilidade de que passem 7 ou mais carros entre 13.15 e 13.18.

7 Modelos contínuos

1 Seja X a v.a. contínua cuja densidade de probabilidade é

$$f(x) = 2x, \quad 0 \leq x \leq 1$$

a)Calcule a distribuição acumulada $F(x)$ o valor esperado $E(X)$, a variância $Var(X)$ e o desvío padrão $\sigma(X)$.
b)Calcule $P(0 < X < 1/2), P(1/3 < X \leq 1)$.
c)Grafique $F(x)$ e determine o valor de x_0 tal que $F(x_0) = 0.95$. Calcule $P(x_0 < X \leq 1)$. Interprete.

2 Seja X a v.a. contínua cuja densidade de probabilidade é

$$f(x) = \begin{cases} kx^2 & \text{se} \quad 0 \leq x \leq 1 \\ 0 & \quad\text{caso contrario} \end{cases}.$$

a)Determinar o valor k.
b)Calcular $E(X), Var(X)$.
c)Determine a f.d.a. de X.

3 O comprimento do lado de um quadrado aleatório é uma v.a. uniforme em [0,5]. Calcular a área esperada do quadrado.

4 Seja X uma v.a. com distribuição uniforme em [-1,1] (Notação $X \sim U[-1,1]$). Isto significa que X tem função de densidade f e função de distribuição acumulada F dadas pelas fórmulas abaixo.

$$f(x) = 1/2 \;\; se \;\; -1 \leq x \leq 1; \;\; f(x) = 0 \;\; caso \;\; \text{contrário}.$$

$$F(x) = 0 \;\; se \;\; x < -1; \;\; F(x) = (x{+}1)/2 \;\; se \;\; -1 \leq x \leq 1; \;\; F(x) = 1 \;\; se \;\; x > 1.$$

Calcule as probabilidades dos seguintes conjuntos usando f e F:

$$\{X < -2\}, \{X \leq 0\}, \{X < 0\}, \{-1 < X < 0,8\}, \{-3 \leq X \leq 0,8\},$$
$$\{0,2 < X < 1,5\}, \{-4 < X < -3\} \cup \{0,2 \leq X \leq 2\},$$
$$\{0,2 \leq X < 1\}, \{X > -2\}, \{X > 0\}, \{X \geq 1,2\}, \{X \geq 4\},$$
$$\{-1/2 \leq X < 1/2\} \cup \{8 \leq X < 24\}, \{X = 0\}, \{X = 4\}.$$

5 Seja X uma v.a. cuja função de distribuição acumulada é dada por:

$$F(x) = 0 \quad se \quad x < -1; \quad F(x) = (x+1)/4 \quad se \quad -1 \leq x < 1;$$

$$F(x) = 2x - \frac{x^2}{2} - 1 \quad se \quad 1 \leq x < 2; \quad F(x) = 1 \quad se \quad x \geq 2.$$

a)Calcule a função de densidade de X. É possível e/ou importante calcular o valor da densidade em -1, 1 e 2? Veja os gráficos da densidade e da acumulada: por que motivo X é considerada uma v.a. contínua?
b) Calcule E(X), E(X^2), Var(X), Mediana(X) e Moda(X);
c) Calcule as probabilidades dos seguintes conjuntos:

$$\{X = -2\}, \{X = -1\}, \{X = 0\}, \{X = 1/2\}, \{X = 1\}, \{X = 2\},$$

$$\{X = 8\}, \{X < -2\}, \{X \leq 0\}, \{X < 1,5\}, \{X < 4\}, \{X \geq 0\},$$

$$\{X > 1,5\}, \{X > 4\}, \{0 < X < 1,5\}, \{-2 < X < 0\}, \{0 \leq X < 1,2\},$$

$$\{1,1 \leq X \leq 3\}, \{-2 \leq X < 9\}, \{-2 \leq X < 1,5\}, \{3 \leq X < 6\}.$$

6 0 tempo de vida em horas X de um transistor é uma v.a. com função de densidade:

$$f(x) = 0 \quad se \quad x < 0; \quad f(x) = 500^{-1}e^{-x/500} \quad se \quad x \geq 0;$$

Calcule a função de distribuição acumulada, a média, variância, mediana e moda de X.
Calcule $P(\{X > x\})$ para todo x real.

7 Assuma que o tempo de duração X de uma consulta médica tenha distribuição exponencial com média de 10 minutos. Calcule a probabilidade dos seguintes eventos:
a) uma consulta demora 20 minutos no máximo;
b) uma consulta demora mais de 20 minutos;
c) uma consulta demora mais que o tempo médio.
Calcule a probabilidade do evento $\{X > E(X)\}$ para $X \sim Exp(\alpha)$, para todo $\alpha > 0$.

8 Seja $X \sim Normal(\mu, \sigma^2)$.
a) Prove que $Z = (X - \mu)/\sigma$ tem distribuição Normal(0,1).

(Dica: chame de F à função de dist. acumulada de X e de ϕ à de Z, e verifique que

$$\phi(t) = F(\mu + \sigma t)$$

e portanto que a relação entre as densidades φ de X e φ_0 de Z é

$$\varphi_0(t) = \sigma\varphi(\mu + \sigma t).$$

b) Verifique também as relações inversas entre as acumuladas e densidades de X e Z:

$$F(x) = \phi((x - \mu)/\sigma) \quad e \quad \varphi(x) = \sigma^{-1}\varphi_0((x - \mu)/\sigma).$$

9 Suponha que a duração de uma componente eletrônica possui distribuição exponencial de parâmetro $\lambda = 1$, calcule:
a)A probabilidade de que a duração seja menor a 10.
b)A probabilidade de que a duração esteja entre 5 e 15.
c)O valor t tal que a probabilidade de que a duração seja maior a t assuma o valor 0.01.

10 A longitude do lado de um cubo aleatório é uma v.a. contínua $Exp(3)$. Calcule o volume esperado do cubo.

11 Seja T a v.a. contínua de distribuição exponencial de parâmetro 2 e seja X a v.a. discreta definida como

$$X = \begin{cases} 0 & \text{se} \quad 0 \leq T < 1 \\ 1 & \text{se} \quad 1 \leq T < 2 \\ 2 & \text{se} \quad 2 \leq T \end{cases}.$$

Determine a função de probabilidades de X.

12 Assumindo que X possui distribuição $N(\mu, \sigma^2)$, calcule:
a)$P(X \leq \mu + 2\sigma)$
b)$P(|X - \mu| \leq \sigma)$
c)o número a tal que $P(\mu - a\sigma \leq X \leq \mu + a\sigma) = 0.99$
d)o número a tal que $P(X > a) = 0.90$.
Por simplicidade assuma primeiramente que $\mu = 1$ e $\sigma = \sqrt{2}$. Logo, determine as quantidades requeridas para μ e σ geral.

13 Seja $X \sim Normal(\mu, \sigma^2)$. Calcule as probabilidades dos seguintes intervalos:

$$\{\mu - \sigma < X < \mu + \sigma\}, \{\mu - 2\sigma < X < \mu + 2\sigma\}, \{\mu - 3\sigma < X < \mu + 3\sigma\},$$
$$\{-\infty < X < \mu\}, \{\mu < X < \infty\}; \{\mu - \sigma < X < \mu\}, \{\mu < X < \mu + \sigma\},$$
$$\{\mu - \sigma < X < \mu + 2\sigma\}, \{\mu - 2\sigma < X < \mu + \sigma\}.$$

14 Considere o peso de um puma macho adulto como uma variável aleatória com distribuição $Normal(\mu, \sigma^2)$. Sabe-se que 33,0 % destes animais tem peso inferior a 82.8 kg e também que 0,4% tem peso superior a 98,25 kg. Calcule μ e σ.

8 Aproximação Binomial-Normal e Intervalos de Confiança para proporções

1 Seja Y uma variável aleatória com distribuição binomial de parâmetros $n = 10$ e $p = 0.4$.
a)Determine o valor exato e a aproximação Normal para $P(7 \leq Y)$. Calcule pela aproximação Normal o valor $P(6.5 \leq Y)$ e conclua qual dos dois valores dados pela aproximação representa melhor o verdadeiro valor de $P(7 \leq Y)$, onde Y é Binomial.
b)Determine o valor exato e a aproximação Normal para $P(Y < 5)$. Calcule pela aproximação Normal o valor $P(Y \leq 4.5)$ e conclua qual dos dois valores dados pela aproximação representa melhor o verdadeiro valor de $P(Y < 5)$, onde Y é Binomial.

2 De um lote de produtos manufaturados, extraímos 100 itens ao acaso. Se 10% dos itens são defeituosos, calcular a probabilidade de 12 itens dentre os selecionados serem defeituosos. Use a aproximação Normal.

3 Seja p a proporção de indivíduos com glaucoma na cidade de Campinas. Se o Ministério da Saúde informa que a proporção atual é igual a 0.1 e temos uma amostra de 19 indivíduos selecionados ao acaso desta população, responda:
a)Qual é a probabilidade de que a proporção amostral

$$p^* = \frac{n^o \text{ de portadores na amostra}}{19}$$

seja maior ou igual a 5/19 ? Calcule a probabilidade exata.
b)Resolva o item a) utilizando a aproximação Normal dada pelo Teorema Central do Limite.
c)Compare o valor exato de $P(5/19 \leq p^*)$ calculado em a) com os valores calculados por aproximação Normal: $P_N(5/19 \leq p^*)$ (do item b)), $P_N(4.5/19 \leq p^*)$, $P_N(5.5/19 \leq p^*)$, estes últimos calculados utilizando a aproximação Normal.
Determine qual deles representa melhor o verdadeiro valor dado no item a).
d)Calcule a probabilidade exata $P(p^* \leq 2/19)$ e compare o resultado com as aproximações pela Normal:

$P_N(p^* \leq 1.5/19), P_N(p^* \leq 2/19), P_N(p^* \leq 2.5/19)$. Qual destes valores representa melhor o verdadeiro valor $P(p^* \leq 2/19)$?

4 Uma amostra aleatória de 625 donas-de-casa revela que 70% delas preferem a marca X de detergente. Construir o intervalo de confiança de 90% (intervalo por estimativa pontual e intervalo conservador) para $p =$ proporção das donas-de-casa que preferem X.

5 Antes de uma eleição, um determinado partido está interessado em estimar a proporção p de eleitores favoráveis a seu candidato. Uma amostra piloto de tamanho 100 revelou que 60% dos eleitores eram favoráveis ao candidato em questão.

(a) Determine o tamanho da amostra necessário para que o erro cometido na estimação seja no máximo 0.01 com probabilidade de 80%.

(b) Se na amostra fina (com tamanho dado por (a)) observou-se que 55% dos eleitores eram favoráveis ao candidato em questão, construa um intervalo de confiança (95%) para a proporção p.

6 Suponha que estejamos interessados em estimar a porcentagem de consumidores de um certo produto. Se uma amostra de tamanho 300 forneceu 100 indivíduos que consomem o dado produto, determine:

(a) O intervalo de confiança para p, com coeficiente de confiança 95%.

(b) O tamanho da amostra para que o erro da estimativa não exceda 0.02 unidades com probabilidade de 95%.

7 Um pedido de auxílio, feito pelo correio, teve 412 respostas a 5000 cartas enviadas, e outro pedido, mais dispendioso, teve 312 respostas a 3000 cartas enviadas. Obtenha o intervalo de confiança de 90% para a diferença de proporções entre os dois pedidos.

9 Teste de hipóteses para proporções

1 Os dados correspondem a uma distribuição $Bin(n,p)$. Conduzir o seguinte teste
$H_0 : p = 0.75$ vs $H_1 : p < 0.75$. Assuma $n = 150$.
a)Se $p^* = 0.72$ for uma estimativa pontual de p. Determine a força da evidência contida nos dados (π-value).
b)Verifique se a entimativa p^* apresenta evidência suficiente para rejeitar H_0 ao nível $\alpha = 0.01$.

2 Suspeita-se da honestidade de um dado de 6 faces. Procurando suporte para tal afirmação considera-se o número de vezes que a face 2 é obtida numa sequência de n lançamentos independentes.
a)Determine a hipótese nula H_0 e a alternativa H_1.
b)Em $n = 20$ lançamentos independentes obtem-se 2 vezes a face 2. Calcule a força da evidência contida nos dados e responda: Para que níveis de significância α, a hipótese H_0 é rejeitada?. Interprete.
Calcule o π-value utilizando a aproximação normal e responda: Para que níveis de significância α, a hipótese H_0 é rejeitada?. Compare.
c)Em $n = 20$ lançamentos independentes obtem-se 6 vezes a face 2. Calcule a força da evidência contida nos dados e determine se os dados resultam significantes ao nível $\alpha = 0.10$.

3 Membros de uma associação profissional desejam provar que menos da metade dos eleitores apoiam as medidas tomadas pela equipe econômica do governo para enfrentar a crise financeira internacional. Seja p a proporção de eleitores que apoiam as medidas.
a)Determine a hipótese nula e a alternativa de um teste que permita avaliar a situação.
b)Se uma pesquisa com 500 eleitores selecionados ao acaso revela que 228 apoiam as medidas econômicas, podemos dizer que os dados são significantes ao nível $\alpha = 0.05$?.

4 Dois grupos, A e B, são formados por pessoas distintas que possuem a mesma enfermidade. É ministrado um soro ao grupo A mas não ao grupo B. Das 100 pessoas que formaram o grupo A, 75 se curaram e, das 100 pessoas que formaram o grupo B, 65 obtiveram a cura. Verifique se o soro é eficiente na cura da enfermidade.

5 Numa amostra aleatória de visitantes de um museu, 22 de 100 famílias provenientes da região Sul e 33 de 120 famílias provenientes de São Paulo compraram alguma coisa nas lojas do museu. Podemos considerar que a proporção de pessoas, provenientes da região Sul e de São Paulo, que compraram alguma coisa nas lojas do museu, são iguais?

6 Um método de borrifar nuvens (para provocar chuva) obteve sucesso em 54 dentre 150 tentativas, enquanto que o outro método obteve sucesso em 33 dentre 100 tentativas. Pode-se concluir que o primeiro método é superior ao segundo?

7 Sendo X o número de sucessos em $n = 10$ provas de Bernoulli, queremos testar $H_0 : p = 0,6$. Se o teste for unilateral e rejeitarmos H_0 para valores pequenos de X, detemine o p-valor se o valor observado de X for 3. Conclua sobre a rejeição ou não de H_o.

8 Membros de uma associação patronal desejam demonstrar que mais de 60% dos seus associados apoiam a política de privatização do governo. Determine a região crítica do teste de hipótese para essa situação, para um nível de significância $\alpha = 0.05$, supondo que os dados são colhidos de uma amostra com 80 associados selecionados ao acaso.

10 Estimação por Intervalo para populações Normais

1 Considere a seguinte amostra aleatória de tamanho 20 proveniente de uma distribuição Normal de média desconhecida μ e variância σ^2: 13.736, 14.579, 14.025, 13.542, 14.294, 13.815, 13.615, 13.633, 13.893, 14.105, 14.129, 15.029, 13.814, 14.516, 13.982, 14.174, 13.900, 14.319, 13.822, 13.728

a) Calcule o $100\gamma\%$ I.C. para μ sabendo que $\sigma^2 = 0.36$, para os níveis de confiança $\gamma = 0.9, 0.95, 0.99$. Calcule o comprimento de cada intervalo de confiança. Que evidência há na relação entre o comprimento do intervalo e o nível de confiança?

b) Calcule o $100\gamma\%$ I.C. para μ supondo σ^2 desconhecido, para os níveis de confiança $\gamma = 0.9, 0.95, 0.99$.

2 Considere a seguinte amostra aleatória de tamanho 15 proveniente de uma distribuição Normal de média μ e variância desconhecida σ^2: 5.055, 6.916, 5.812, 5.044, 4.914, 5.665, 4.772, 5.502, 3.841, 5.782, 4.579, 5.477, 7.158, 5.254, 5.276

a) Calcule o $100\gamma\%$ I.C. para σ sabendo que $\mu = 5$, para os níveis de confiança $\gamma = 0.9, 0.95, 0.99$. Calcule o comprimento de cada intervalo de confiança.

b) Calcule o $100\gamma\%$ I.C. para σ supondo μ desconhecido, para os níveis de confiança $\gamma = 0.9, 0.95, 0.99$. Calcule o comprimento de cada intervalo de confiança.

3 Considere a média amostral de uma a.a. de uma população com média μ e variância igual 10. Encontre o valor de n para que o intervalo aleatório$(\bar{X} - 0,5; \bar{X} + 0,5)$ tenha probabilidade aproximada de 0,954 de conter μ.

4 Sejam $X_1, ..., X_n$ a.a. tendo distribuição $N(\mu, \sigma^2)$. A margem de erro de um IC de nível $100(1-\alpha)\%$ para μ utilizando-se $\bar{X}$, se σ conhecido, é dado por $e = z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}$. Considerando $\sigma = 4$, qual é o tamanho amostral de modo a ter 90% de certeza de que o erro de estimação não exceda 0,8?

11 Teste de hipóteses para populações Normais

1 Denotemos por μ a verdadeira média de nível de radioatividade (picocuries por litro). O valor 5 pCi/L é considerado como linha divisoria entre água segura e não segura. Qual dos seguintes testes recomenda conduzir?
$H_0 : \mu = 5$ vs $H_1 : \mu > 5$
$H_0 : \mu = 5$ vs $H_1 : \mu < 5$
Explique seu raciocínio em termos dos erros tipo I e II.

2 Para cada situação apresentada a seguir, verifique se os dados apresentam evidência suficiente para rejeitar a hipótese nula sendo que s denota o desvio padrão amostral $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$.

a) População normal, $n = 15$, $\overline{X} = 83.9$, $s = 18.2$, $\alpha = 10\%$, para o teste $H_0 : \mu = 85$ vs $H_1 : \mu < 85$.

b) População normal, $n = 15$, $\overline{X} = 79.1$, $s = 11.8$, $\alpha = 10\%$, para o teste $H_0 : \mu = 76$ vs $H_1 : \mu \neq 76$.

3 Sabendo que a resistência à tensão, de uma peça de algodão possui distribuição normal.

a) A resistência é medida em 15 peças selecionadas ao acaso, observando-se uma média amostral igual a 39.3 e um desvio padrão amostral igual a 2.6. Verifique se os dados são significantes ao nível $\alpha = 10\%$, para o teste $H_0 : \mu = 40$ vs $H_1 : \mu \neq 40$.

b) Determine a região crítica dos teste enunciado em a) para $\alpha = 10\%$.

c) A resistência é medida em 54 peças selecionadas ao acaso, observando-se uma média amostral igual a 42.4 e um desvio padrão amostral igual a 3.1. Calcule a força da evidência contida nos dados e determine para quais níveis de significância H_0 é rejeitada.

d) Melhorias implementadas no tratamento da fibra de algodão permitem suspeitar que a resistência tem aumentado. Perante esta afirmação reformule o teste. Se essa resistência foi medida em 15 peças observando-se uma média amostral de 41.3 com um desvio padrão amostral igual a 2.6. Verifique se os dados são significantes ao nível $\alpha = 0.05$. Determine a região crítica do teste para $\alpha = 0.05$.

4 A demanda biológica de oxigênio (DBO) é um índice de poluição controlado nas indústrias de papel e celulose (para preservar o equilibrio ambiental toda indústria de papel deve consumir uma quantidade de oxigênio que não supere um "valor limite"). Em 43 medidas coletadas numa indústria, no peíodo: Setembro 1999-Fevereiro 1999, a média e o desvio padrão dos dados observados foram 3242 ppd e 757 ppd, respectivamente. Aquela empresa tinha estabelecido como valor limite 3000 ppd para o DBO médio. Julgaria que os dados amostrais suportam que a meta foi atingida ao nível $\alpha = 5\%$?

5 Uma empresa mineira acredita que a exploração de urânio é possível numa certa região, isto é, na região a concentração média de urânio é superior a 10. Admitindo-se que a distribuição desta concentração é normal e que as medições em 13 pontos selecionados ao acaso na região são 7.92, 10.29, 19.89, 17.73, 10.36, 13.50, 8.81, 6.18, 7.02, 11.71, 8.33, 9.32, 14.61

a) Verifique se há evidência suficiente contra a hipótese de abandono da área .

b) Qual seria a região crítica do teste ao nível de significância $\alpha = 2\%$?

6 Em 18 condenações por "posse de drogas" num tribunal norte-americano as condenações atribuídas tiveram média de 38 meses e desvio padrão amostral de 4 meses. Considerando que as condenações são normalmente distribuídas, verifique se os dados suportam ao nível de significância $\alpha = 5\%$ a suspeita de que nesse tribunal as condenações por "posse de drogas" é em média maior do que 36 meses. Faça o mesmo teste considerando $\alpha = 1\%$. Interprete.

7 Um fabricante de aparelhos de TV afirma que são necessários no máximo 250 microamperes (μA) para atingir um certo grau de brilho num tipo de TV. De uma amostra de 20 aparelhos obtivemos uma média amostral de $\overline{X} = 257.3$ μA. Denotemos por m a verdadeira média de

μA necessário para atingir o grau de brilho desejado e suponhamos que m é a média de uma população normal com σ conhecido e igual a 15.

a) Calcule a força da evidência contida nos dados para o nível $\alpha = 0.05$ conduzindo o teste cuja hipótese nula especifica que m é no máximo 250 μA.

b) Calcule a região crítica do teste para o nível $\alpha = 0.05$.

c) Se $m = 260$, Qual é a probabilidade de cometer um erro tipo II?

d) Para qual valor de n (tamanho amostral) a probabilidade de cometer o erro tipo II resulta igual a 0.01.

8 O ponto de desvanecimento de cada uma de 16 amostras de uma certa marca de vegetais hidrogenados foi determinado, resultando numa média amostral $\overline{X} = 94.32$. Considerando que o ponto de desvanecimento possui distribuição normal de desvio conhecido $\sigma = 1.20$.

a) Verifique se a amostra apresenta evidência suficiente para rejeitar H_0 ao nível $\alpha = 0.01$, calculando o π-value onde $H_0 : \mu = 95$ vs $H_1 : \mu \neq 95$.

b) Se $\alpha = 0.01$ e $\mu' = 94$. Qual é a probabilidade de cometer erro tipo II?

c) Se $\alpha = 0.01$. Que valor de n (tamanho amostral) é necessário para obter uma probabilidade de cometer erro tipo II, com $\mu' = 94$ igual a 0.1?

9 O ponto médio desejado de SiO_2 em certo tipo de cimento aluminoso é de 5.5. Para provar se o verdadeiro ponto médio da porcentagem numa planta de produção em particular é 5.5, foram coletadas 16 amostras. Supondo que a porcentagem de SiO_2 numa amostra está normalmente distribuída com desvio conhecido e igual a $\sigma = 0.3$ e sabendo que na amostra selecionada obteve-se $\overline{X} = 5.25$, responda:

a) Os dados indicam de forma conclusiva que o verdadeiro ponto médio de porcentagem não é $\mu = 5.5$?.

b) Se $\alpha = 0.01$, qual o valor de n (tamanho amostral) é necessário para obter uma probabilidade de cometer erro tipo II, com $\mu' = 5.6$ igual a 0.01?

10 Um experimento para comparar a resistência de coesão à tensão do morteiro modificado de látex de polímeros, com a resistência do morteiro não modificado, supondo que os dados tem distribuição Normal; resultou em $\overline{X} = 18.12\ kfg/cm^2$ para o morteiro modificado e em $\overline{Y} = 16.87\ kfg/cm^2$ para o morteiro não modificado. Sejam μ_1 e μ_2 as verdadeiras resitências de coesão à tensão para os morteiros modificado e não modificado respectivamente. Verifique se os dados suportam a rejeição de H_0. Onde $H_0 : \mu_1 - \mu_2 = 0$ vs $H_1 : \mu_1 - \mu_2 > 0$ com nível de significância $\alpha = 0.01$, nas seguintes situações:

a) Se para o morteiro modificado foi utilizada uma amostra de tamanho $m = 40$ e para o não modificado foi utilizada uma amostra de tamanho $n = 32$. Os valores dos desvios são conhecidos σ_1 e σ_2 (associados respectivamente ao morteiro modificado e ao não modificado), $\sigma_1 = 1.6$ e $\sigma_2 = 1.4$. Proponha uma estatística para conduzir o teste e verifique se os dados indicam a rejeição de H_0.

b) Se para o morteiro modificado foi utilizada uma amostra de tamanho $m = 40$ e para o não modificado foi utilizada uma amostra de tamanho $n = 32$. Os valores dos desvios são conhecidos σ_1 e σ_2 (associados respectivamente ao morteiro modificado e ao não modificado), $\sigma_1 = \sigma_2 = 1.6$. Proponha uma estatística para conduzir o teste e verifique se os dados indicam a rejeição de H_0.

c) Sabendo que

$$m = 30, \sum_{i-1}^{m}(X_i - \overline{X})^2 = 40.1,\ n = 22, \sum_{j=1}^{n}(Y_j - \overline{Y})^2 = 53.22.$$

Assumindo que os desvios σ_1 e σ_2 são desconhecidos e iguais, proponha uma estatística para o teste e determine se os dados indicam a rejeição de H_0.

11 Os estudantes universitários homens entediam-se mais facilmente que as estudantes mulheres?. Esta pergunta foi examinada pelo artigo "Boredom in Young Adults Gender and Cultural Comparisons"(*J. of Cross Cultural Psych.* pp. 209-223). Os autores aplicaram uma escala denominada *Escala Proneness de tedio* a 97 estudantes homens e a 148 estudantes mulheres, todos eles de universidades norteamericanas. Assumindo que a classificação fornecida pela escala Proneness possui distribuição normal verifique se a seguinte informação apoia a hipótese da

investigação.
Faça o teste adequado utilizando um nível de significância $\alpha = 0.05$ e os dados da seguinte tabela

Gênero	Tamanho amostral	Média am.	Desvio verdadeiro (σ)
Homens	97	10.40	4.83
Mulheres	148	9.26	4.86

12 Denotemos por μ_1 e μ_2 aos verdadeiros pontos médios de durações de superfícies de rodagem para duas marcas competidoras de medida FR78-15 de pneus radiais. Faça o seguinte teste de hipótese assumindo que a duração das superfícies de rodagem possui distribuição normal $H_0 : \mu_1 - \mu_2 = 0$ vs $H_1 : \mu_1 - \mu_2 \neq 0$ com nível de significância $\alpha = 0.05$, usando a seguinte informação: $m = 40$, $\overline{X} = 36500$, $\sigma_1 = 2200$ (valor verdadeiro do desvio) e $n = 40$, $\overline{Y} = 33400$, $\sigma_2 = 1900$ (valor verdadeiro do desvio).

13 (Teste sobre a variância com a média desconhecida) Seja $X_1, ..., X_n$ uma amostra aleatória normal $N(\mu, \sigma^2)$, μ desconhecido. Considere $H_o : \sigma^2 = 100$ versus $H_1 : \sigma^2 \neq 100$. Sabendo que $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ (distribuição Qui-Quadrado com $n-1$ graus de liberdade) onde $S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$ e utilizando $\alpha = 0,05$, $n = 16$, obtenha a região crítica do teste. Sabendo-se que $S^2 = 169$, qual é a conclusão do teste? Lembre-se que a distribuição Qui-Quadrado não é simétrica?

14 Uma fábrica de embalagens para produtos químicos está estudando dois processos, A e B, para combater a corrosão de suas latas especiais. Considerando $X_1, ..., X_n$ e $Y_1, ..., Y_m$ a.a. dos processos A e B, respectivamente. As amostras X_i, $i = 1, ..., n$ e Y_i, $i = 1, ..., m$ são independentes e possuem distribuição, respectivamente, $N(\mu_1, \sigma_1^2)$ e $N(\mu_2, \sigma_2^2)$, σ_1^2 e σ_2^2 conhecidos. Queremos testar o efeito dos tratamentos A e B, isto é, queremos testar as hipóteses $H_o : \mu_1 = \mu_2$ versus $H_1 : \mu_1 \neq \mu_2$. A região crítica associada às hipóteses é $RC = \{\hat{X} - \hat{Y} | \hat{X} - \hat{Y} \leq c_1 ou \hat{X} - \hat{Y} \geq c_2\}$. Utilizando $\alpha = 0,05$ e os dados apresentados na seguinte tabela, conclua sobre os dois tratamentos.

Método	Amostra	Média	Desvio Padrão
A	15	48	10
B	12	52	15

15 (Teste sobre a variância com a média conhecida) Seja $X_1, ..., X_n$ uma amostra aleatória normal $N(\mu, \sigma^2)$, μ conhecido igual a 45. Considere $H_o : \sigma^2 = 7$ versus $H_1 : \sigma^2 < 7$. Sabendo que $\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \sim \chi_n^2$ (distribuição Qui-Quadrado com n graus de liberdade) e utilizando $\alpha = 0,05$, $n = 15$, obtenha a região crítica do teste. Sabendo-se que $\sum_{i=1}^{n}(X_i - \mu)^2 = 42$, qual é a conclusão do teste? Utilizando $\alpha = 0,01$, qual é a conclusão do teste?