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.

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.

2 Probabilidade

2.1 Espaço Amostral

1 Uma moeda e um dado são lançados. Dé um espaço amostral do experimento e depois represente-o como produto cartesiano dos dois espaços amostrais, correspondente 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 familías 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. Selecionase 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.

2.2 Seleção de amostras e Cálculo de probabilidades

1 Numa sala com 15 pessoas, quantas amostras (de 15 pessoas) com a condição de que todas elas tenham nacido em días diferentes podem ser formadas?. Assuma que o ano tem 365 días.

2 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.

3 Uma moeda é lançada três vezes, registra-se a terna de possíveis resultados dos lançamentos.
a)Quantas amostras podem ser obtidas?
b)Quantas amostras de exactamente duas caras podem ser obtidas?. Qual a probabilidade de obter nos três lançamentos exatamente duas caras?.
c)Quantas amostras com pelo menos uma cara podem ser obtidas?. Qual a probabilidade de obter nos três lançamentos pelo menos uma cara?.

4 Dois dados são lançados registrando-se os resultados da face superior.
a)Dé um espaço amostral do experimento e depois represente-o como produto cartesiano dos dois espaços amostrais, correspondente aos experimentos considerados individualmente.
b)Quantas amostras podem ser obtidas?
c)Quantas amostras que apresentam resultados diferentes nos lançamentos entre um e outro dado podem ser obtidas?. Qual a probabilidade de obter resultados diferentes nos dois lançamentos?.

5 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.?

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

3 Probabilidade condicional. Indepêndencia. Teorema de Bayes

1 Assuma a seguinte distribuição de 300 estudantes, segundo sexo e área de estudo

Sexo	Biologia	Exactas	Humanas
Masculino	52	40	58
Feminino	38	32	80

um estudante é selecionado ao acaso
a)Qual é a probabilidade de que seja de sexo feminino e das humanas?.
b)Qual é a probabilidade de que seja de sexo masculino e não seja das biológicas?.
c)Dado que foi selecionado um estudante da área das humanas, qual é a probabilidade de que seja do sexo feminino?.
d)Dado que foi escolhido um estudante de sexo masculino, qual é a probabilidade de que não seja das biológicas?.

2 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?.

3 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,

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

4 Uma empresa monta rádios cujas partes são lançadas por três de suas fabricas, denominadas A_1, A_2 e A_3 respectivamente. Elas produzem $15\%, 35\%$ e 50% do total. As probabilidades de que as fabricas apresentem partes defeituosas são $0.01, 0.05$ e 0.02 respectivamente.
a)Uma peça é selecionada ao acaso desde a produção verificando-se que a peça é defeituosa, qual é a probabilidade de que a peça tenha sido lançada pela fabrica $A_i, i = 1, 2, 3$?.
b)Qual é a probabilidade de que uma peça selecionada ao acaso seja defeituosa?

5 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 o 0.5% dos chips apresentam falhas, qual é a probabilidade de que um chip que foi declarado com falhas seja bom?

6 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:

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

7 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

8 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 Seja X uma variável aleatória discreta com $P(X = 0) = 0.25$, $P(X = 1) = 0.125$, $P(X = 2) = 0.125$, $P(X = 3) = 0.5$. Graficar 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:

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

2 Dada a função de distribuição acumulada

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

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:

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

3 Com dados do último censo, a asistente social de um censo 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.

4.1 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 industria.

c)O mes do ano com maior número de enchentes na cidade de São Paulo.

2 Sendo X uma variável 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 fabrica 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 veiculo 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 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)$P(Y = 1|Y < 3)$

6 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 parametro $\lambda = 1$ por m^2. Uma chapa e 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.

7 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)$P(H > 0)$

8 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 defeituosas.
c)No mínimo 1 boa.

9 Suponha que a probabilidade de que um item produzido por uma máquina seja defeituoso e de 0.2. Se 10 itens produzidos por esta máquina são selecionados ao acaso, qual é a probabilidade de que no mais do que um defeituoso seja encontrado?. Use a Binomial e a Poisson e compare os resultados.

10 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).

5 Distribuição Normal

1 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.

2 Suponha que o tempo de duração dos pneus fabricados por Bolinha possuim distribuição normal de média μ e variância σ^2, um cliente compra um pneu e conjuntamente adquire um seguro de garantia do pneu, ele somente precisa escolher entre dois possíveis planos de seguro. O primero plano de seguro exige que o fabricante do pneu reembolse 100 reais ao cliente se o tempo de vida do pneu for menor o igual a $\mu - \sigma$ o fabricante tem um credito de 100 reais se o pneu durar mais de $\mu + \sigma$. O segundo plano castiga ao fabricante com um reembolso de 100 a favor do cliente se a vida do pneu for menor o igual a $\mu - 3\sigma$ e premia ao fabricante com um credito de 100 reais se este durar mais de $\mu + 3\sigma$. Em média, que plano resulta mais conveniente para o cliente ?, que plano resulta mais conveniente para o fabricante dos pneus?.

5.1 Aproximação Normal da Binomial

1 Seja Y 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 ao 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 ao 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 serem defeituosos. Use a aproximação Normal.

3 Seja p a proporção de individuos 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 individuos selecionados ao acaso desde a população, responda

a)Qual é a probabilidade de que a proporção amostral

$p^* = \frac{\text{no. de portadores na amostra}}{19}$

seja maior o igual a 5/19 ?. Calcule a probabilidade exata.

b)Calcule a) utilizando a aproximação Normal dada pelo Teorema Central do Limite.

c)Compare o valor exato $P(5/19 \leq p^*)$ calculado em a) com os valores calculados por aproximação Normal $P_N(5/19 \leq p^*)$ (desde o ponto b)), $P_N(4.5/19 \leq p^*)$, $P_N(5.5/19 \leq p^*)$ estes ultimos calculados utilizando a aproximação Normal.

Determine qual deles representa melhor ao verdadeiro valor dado no item a).

d)Calcule a probabilidade exata $P(p^* \leq 2/19)$ e compare o resultado com as aproximaçõ es 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 ao verdadeiro valor $P(p^* \leq 2/19)$?.

6 Intervalos de Confiança

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

2 Antes de uma aleição, 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, cons-

trua um intervalo de confiança (95%) para a proporção p.

3 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 de p, com coeficiente de confiança 95% .

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