Estatística - ME 414: Gabarito da Lista 1 [1]

1. a. Vitamina (A, B1, B2, B6, B12): Variável qualitativa nominal.

   b. Quantidade de calorias na batata frita: Variável quantitativa continua.

   c. Desfecho de uma doença (curado, não curado); Variável qualitativa nominal.

   d. Classificação de uma lesão (lesão fatal; severa; moderada; pequena): Variável qualitativa ordinal.

   e. Grupo sanguíneo (A, B, AB, O); Variável qualitativa nominal.

   f. Paridade (primeira gestação, segunda gestação, terceira ...): Variável qualitativa ordinal.

   g. Estado geral de um paciente (bom, regular, ruim); Variável qualitativa ordinal.

   h. Número de nascidos vivos em certo hospital em junho/99; Variável quantitativa discreta.

   i. Idade: Variável quantitativa continua pois está relacionada com o tempo que é uma variável quantitativa continua.

   j. Concentração de flúor na água: Variável quantitativa continua.

   k. Atividade esportiva preferida. Variável qualitativa ordinal.

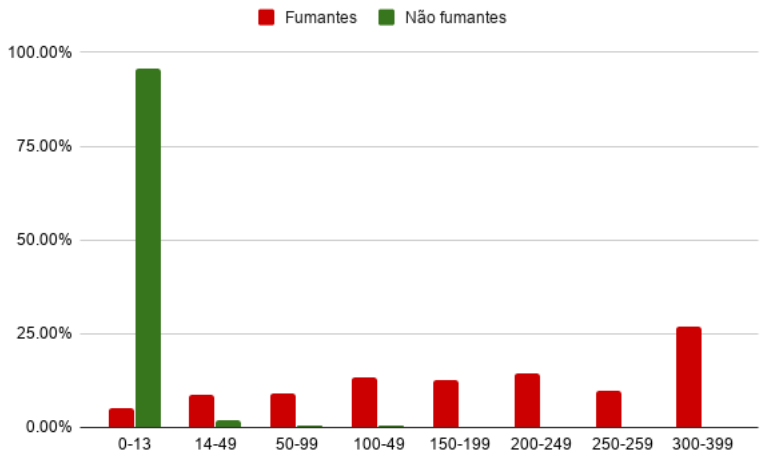
2. a. Não, pois todos os níveis de cotinina possuem diferentes frequências absolutas, isto é, por exemplo que o número de pessoas fumantes é maior ao número de não fumantes mas que em proporção ao total, poderiam ter o mesmo valor.

   b. Resposta

Nível de cotinina	Fumantes(p_i%)	Não fumantes(p_i%)
0-13	5.07%	95.79%
14-49	8.64%	2.09%
50-99	9.23%	0.67%
100-49	13.39%	0.44%
150-199	12.80%	0.20%
200-249	14.29%	0.23%
250-259	9.81%	0.26%
300-399	26.77%	0.32%
Total	100.00%	100.00%

   c. .

[1]PED: Lisbeth Corbacho Carazas, ra162526@ime.unicamp.br



d. Da figura e a tabela, note que a grande maioria dos não fumantes tem baixo nível de cotinina.

3. a. Ordenando os dados temos:

$$21.3 \quad 22.1 \quad 22.8 \quad 23.5 \quad 24.6 \quad 65.4 \quad 67.2 \quad 71.7 \quad 76.3 \quad 84.5.$$

Como são 10 dados, a mediana seria a semissoma dos valores intermédios, posição 5ta e 6ta isto é:

$$\frac{24.6 + 65.4}{2} = 45.$$

É claro que a mediana não representa o conjunto de dados, pois veja que os dados estão divididos em dois grupos.

b. .

b1.

$$21.3 \quad 22.1 \quad 22.8 \quad 23.5 \quad 24 \quad 24.6 \quad 65.4 \quad 67.2 \quad 71.7 \quad 76.3 \quad 84.5.$$

Veja que mediana está na 6ta posição pois $\frac{(11+1)}{2} = 6$, então

$$Md = 24.6.$$

A mediana representa somente os dados do lado esquerdo.

b2.

$$21.3 \quad 22.1 \quad 22.8 \quad 23.5 \quad 24.6 \quad 65.4 \quad 67.2 \quad 71.7 \quad 75 \quad 76.3 \quad 84.5.$$

Veja que mediana está na 6ta posição pois $\frac{(11+1)}{2} = 6$, então

$$Md = 65.4.$$

A Medina representa somente os dados do lado direito.

Do item b1. e b2. a Mediana seria uma medida não estável.

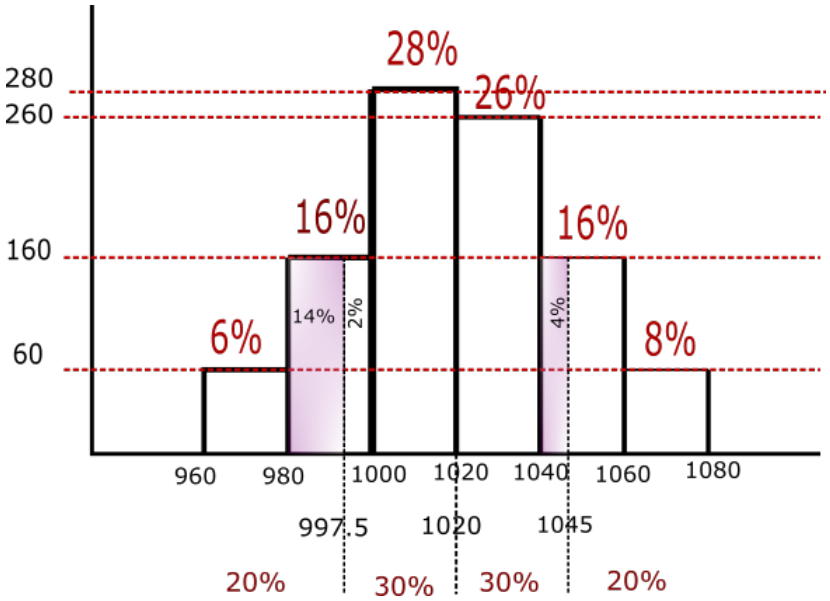
4. a. Sejam $m_i = \frac{L_i+U_i}{2}$ onde L_i é o limite inferior do intervalo e U_i é o limite superior do intervalo e $n = \sum_{i=1}^{6} n_i = 1000$

$$\begin{aligned} \overline{X} &= \sum_{i=1}^{6} \frac{n_i m_i}{n} \\ \overline{X} &= 1020.8 \end{aligned}$$

b.

$$\begin{aligned} \mathrm{Var}(X) &= \sum_{i=1}^{6} \frac{n_i (m_i - \overline{X})^2}{n-1} \\ \mathrm{Var}(X) &= \frac{691360}{999} = 692.061 \end{aligned}$$

c. Da seguinte figura podemos ver que o peso (em gramas) representativo está em torno de 1020 g, o que fornecido pela média calculada no item a..



d. Para um melhor entendimento desta parte do exercício ver o histograma do item c..

- Para os primeiros 20%

$$\begin{aligned} \frac{1000-980}{16\%} &= \frac{x-980}{14\%} && (1) \\ x &= 997.5 && (2) \end{aligned}$$

então o intervalo seria $(960, 997.5]$

- Os 30% seguintes: $x = 1029$, pois $28\% + 2\% = 30\%$ então o intervalo seria dado por: $(997.5, 1020]$

- Para os seguintes 30%

$$\begin{aligned} \frac{1060-1040}{16\%} &= \frac{x-1040}{4\%} && (3) \\ x &= 1045 && (4) \end{aligned}$$

então o intervalo seria $]1020, 1045]$

- Veja que só falta $12\% + 8\% = 20\%$, assim o último intervalo é dado por $(1045, 1085]$

e. Peso inferior $=\overline{X} - 2 \times Dp(\sqrt{962.0}) = 1020.80 - 2 \times 26.30 = 968.18$
Assim para o número de animais separados utilizamos o histograma do item a.

$$\frac{980 - 960}{6\%} = \frac{968.18 - 960}{x\%} \tag{5}$$
$$x\% = 2.454\% \tag{6}$$

Peso Superior $=\overline{X} + 1.52 \times Dp(\sqrt{962.0} = 1020.80 + 1.5 \times 26.30 = 1060.25$
Assim para o número de animais separados utilizamos o histograma do item a.

$$\frac{1080 - 1060}{8\%} = \frac{1060.25 - 1060}{x\%} \tag{7}$$
$$x\% = 0.1\% \tag{8}$$

Então são separados da seguinte forma :
(960,968.18] com 2.45%
(968.18,1060.25] com 89.64%
(1060.25,1080] com 7.9%.

5. a. Seja X: Diâmetros do coração dos adultos maiores normais.
Para cálculo da média temos:

$$\overline{X} = \sum_{i=1}^{15} \frac{X_i}{15} = 131.0666667$$

Para calcular a mediana, ordenamos os dados em forma crescente:

$$X \ : \ 103\ 114\ 114\ 114\ 121\ 125\ 125\ 130\ 130\ 132\ 135\ 139\ 146\ 169\ 169$$

como são 15 dados é um número impar, então o cálculo é dado por: posição$= \frac{15+1}{2} = 8$, assim a mediana é dada pelo valor na oitava posição, isto é 130.

A moda é o valor mais frequente, neste casso é 114.

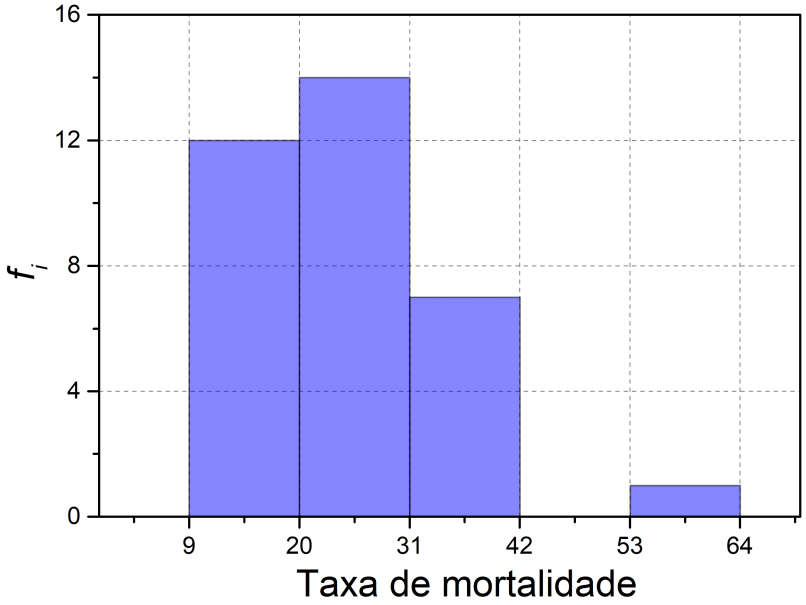
b.

$$\text{Var}(X) = \frac{\sum_{i=1}^{15}(X_i - \overline{X})^2}{14} = 358.495 \tag{9}$$
$$Dp(X) = 18.933 \tag{10}$$

6. a. .

Nro	taxa de morta.	f_i	N_i	$p_i\%$	$P_i\%$
1	(9,20]	12	12	35.29%	35.29 %
2	(20,31]	14	26	41.18%	76.47 %
3	(31,42]	7	33	20.59%	97.06 %
4	(42,53]	0	33	0.00%	97.06%
5	(53,64]	1	34	2.94%	100 %
Total		34		100%	

b. .



c. Seja $m_i = \frac{L_i+U_i}{2}$ onde L_i é o limite inferior do intervalo e U_i é o limite superior do intervalo.

$$n = \sum_{i=1}^{5} f_i = 34$$

$$\begin{aligned} \overline{X} &= \sum_{i=1}^{5} \frac{f_i m_i}{n} \\ \overline{X} &= 24.85 \end{aligned}$$

Para o cálculo da mediana utilizamos as proporções, para isso, observamos nas frequências acumuladas porcentuais tal que contenham 50 %, assim a $Md \in (20, 31]$.

$$\begin{aligned} \frac{Md - 20}{14.71\%} &= \frac{31 - 20}{41.18\%} \\ Md &= 23.92 \end{aligned}$$

Para o cálculo do 1° quartil , ou seja 25° percentil (q_1), observamos nas frequências acumuladas porcentuais tal que contenham 25 %, assim a

$q_1 \in (9, 20]$.

$$\begin{aligned} \frac{q(0.25) - 9}{25\%} &= \frac{20 - 9}{35.29\%} \\ q(0.25) &= 16.79 \end{aligned}$$

Para o cálculo do 3° quartil , ou seja 75° percentil (q_3), observamos nas frequências acumuladas porcentuais que contenham 75%, assim $q_3 \in (20, 31]$.

$$\begin{aligned} \frac{q(0.75) - 20}{(75 - 35.29)\%} &= \frac{31 - 20}{41.18\%} \\ q(0.75) &= 30.60. \end{aligned}$$

A variança e o desvio padrão é dado por:

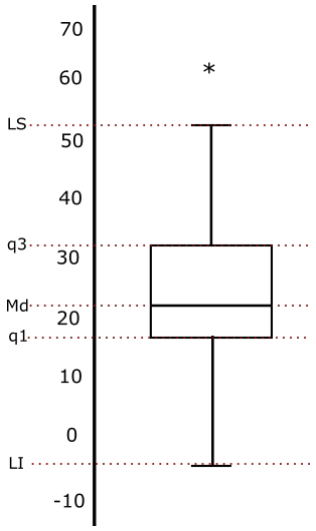
$$\begin{aligned} \mathrm{Var}(X) &= \sum_{i=1}^{5} \frac{f_i(m_i - \overline{X})^2}{n-1} \\ \mathrm{Var}(X) &= \frac{3373.765}{33} = 102.235 \\ Dp(X) &= 10.111 \end{aligned}$$

d. – Esquema de 5 números.

		34	
md		23.92	
q	16.79		30.60
md	9.9		62.2

– Para o Box Plots

$$\begin{aligned} d_q &= q_3 - q_1 = 30.69 - 16.79 = 13.9 \\ LI &= q_1 + 1.5(dq) = 16.79 - 1.5(13.9) = -4.06 \\ LS &= q_3 + 1.5(dq) = 30.60 + 1.5(13.9) = 51.45 \end{aligned}$$



– Ramos de folhas

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

7. a.

$$\overline{Y} = \frac{\sum_{i=1}^{10} Y_i}{10} = 15 \tag{11}$$

$$\text{Var}(Y) = \frac{\sum_{i=1}^{10} (Y_i - \overline{Y})^2}{9} = 16.66 \tag{12}$$

$$Dp(Y) = 4.081 \tag{13}$$

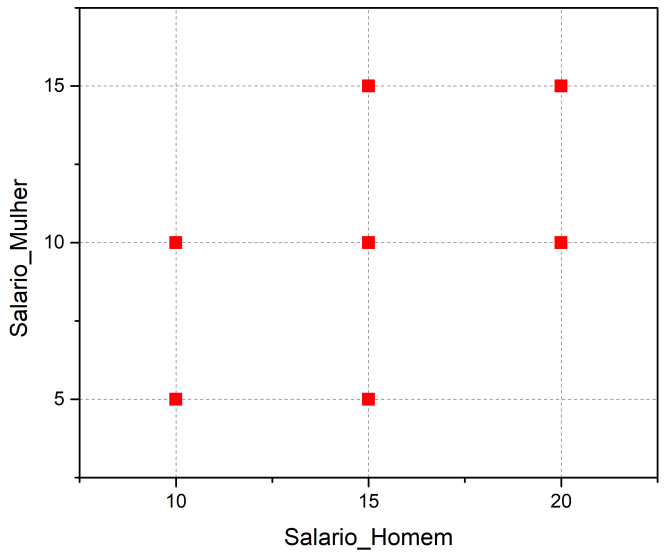
b.

$$\overline{X} = \frac{\sum_{i=1}^{10} X_i}{10} = 10$$

$$\text{Var}(X) = \frac{\sum_{i=1}^{10} (X_i - \overline{X})^2}{9} = 11.11$$

$$Dp(X) = 3.33$$

c. .



d.

$$Corr(X,Y) = \frac{1}{n} \sum_{i=1}^{10} \left(\frac{Y_i - \overline{Y}}{Dp(Y)} \right) \left(\frac{X_i - \overline{X}}{Dp(X)} \right) = 0.4088$$

Como $0.480 > 0$, então, existe uma correlação média entre as variáveis X e Y.

e. S= Salario total do casal.

Casal	1	2	3	4	5	6	7	8	9	10
S=X+Y	15	20	20	20	25	25	30	30	30	35
Z=0.92Y+0.94X	13.9	18.6	18.6	18.5	23.2	23.2	27.9	27.8	27.8	32.5

Desta tabela temos:

$$\overline{S} = \frac{\sum_1^{10} S_i}{10} = 25$$

ou também

$$\overline{S} = \overline{X} + \overline{Y} = 10 + 15 = 25.$$

Para a variança de S, da tabela tem-se:

$$Var(S) = \sum_{i=1}^{10} \frac{(S_i - \overline{S})^2}{9} = 38.89$$

f. Z= Salario com desconto

$$\begin{aligned} &Z = (Y - 0.084Y) + (X - 0.06X) = 0.92Y + 0.94X \\ &\Rightarrow \overline{Z} = 0.92\overline{Y} + 0.94\overline{X} \\ &\Rightarrow \overline{Z} = 23.2 \end{aligned}$$

Para a variança de S, da tabela tem-se:

$$Var(Z) = \sum_{i=1}^{10} (\frac{Z_i - \overline{Z})^2}{9} = 33.53$$

Estatística - ME 414 - Gabarito da Lista 2

1. a.

$$\begin{aligned} P(A|B) &= \frac{P(A\cap B)}{P(B)} = P(A) \text{ pois } \; P(B) > 0 \\ &\Rightarrow \; P(A\cap B) = P(A)P(B) \end{aligned}$$

b.

$$\begin{aligned} P(A|B) &= \frac{P(A\cap B)}{P(B)} \\ P(A|B) &= \frac{P(A)P(B)}{P(B)} = P(A) \end{aligned}$$

Isto é, que o evento A não depende do evento B.

2.

$$\begin{aligned} P(A\cup B) &= P(A) + P(B) - P(A\cap B) && (14) \\ P(A\cup B) &+ P(A\cap B) = P(A) + P(B) && (15) \end{aligned}$$

Como $P(A\cap B) \geq 0$ então $P(A\cup B) \leq P(A) + P(B)$

3. O espaço amostral de dois dados lançados estaria dado por

$$\Omega = \{(1,1),(1,2),....,(6,6)\}$$

Ω possuí 36 valores.

4. a. Os resultados possíveis e as respectivas probabilidades estão dadas por:

$$\Omega = \{PP, VV, PV, VP\}$$

$$\begin{aligned} P(\{PP\}) &= \frac{5}{15}\times\frac{4}{14} = \frac{10}{105} \\ P(\{VV\}) &= \frac{10}{15}\times\frac{9}{14} = \frac{45}{105} \\ P(\{VP\}) &= \frac{10}{15}\times\frac{5}{14} = \frac{25}{105} = P(\{PV\}) \end{aligned}$$

b.

$$\begin{aligned} &i. && P(\{PP\}) = \frac{5}{15}\times\frac{4}{14} = \frac{10}{105} \\ &ii. && P(\{VP\},\{PP\}) = \frac{25}{105} + \frac{10}{105} = \frac{35}{105} \\ &iii. && P(\{VV\},\{VP\}) = \frac{45}{105} + \frac{25}{105} = \frac{70}{105}. \end{aligned}$$

5. • Evento A: viver 70 ou mais anos, então P(A)=0,6.

- Evento B: viver 80 ou mais anos, então P(B)=0,2.

  Probabilidade condicional $P(B|A)$

$$\begin{aligned} P(B|A) &= P(B \cap A)/P(A) = P(B)/P(A) \\ P(B|A) &= 0,2/0,6 \\ P(B|A) &= 1/3 \end{aligned}$$

6. Essa é uma questão sobre Probabilidade Condicional, Teorema de Bayes. Temos que analisar os dados da questão:
   Eventos dados pela questão:
   H: freguês é homem
   A: freguês prefere salada
   M: freguês é mulher
   B: freguês prefere carne

   Agora podemos inferir as respostas:

   20% dos homens preferem salada então 80% dos homens preferem Carne:

$$P(A|H) = 20\% \Rightarrow P(B|H) = 80\%$$

   30% da mulheres comem carne então 70% das mulheres preferem salada:

$$P(B|M) = 30\% \Rightarrow P(A|M) = 70\%$$

   75% dos clientes são homens então 25% dos clientes são mulheres:

$$P(H) = 75\% \Rightarrow P(M) = 25\%.$$

   Sejam:

   P(A) : Probabilidade de um cliente comer salada
   P(B) : Probabilidade de um cliente comer carne

   Para calcular $P(M|A)$ será utilizado o Teorema de Bayes que diz:

$$P(M|A) = \frac{P(A|M)P(M)}{P(A)}.$$

   Precisaremos então do valor de P(A) que pode ser calculado da seguinte forma:

$$\begin{aligned} P(A) &= P(A|M).P(M) + P(A|H).P(H) \\ P(A) &= 70\%(25\%) + 20\%(75\%) = 32.5\%. \end{aligned}$$

   Substituindo os valores conhecidos temos:

$$\begin{aligned} P(M|A) &= \frac{P(A|M)P(M)}{P(A)} \\ P(M|A) &= \frac{0.7 \times 0.25}{0.325} \\ P(M|A) &= 0.54 \end{aligned}$$

7. C : O indivíduo responde corretamente.
   SR: O indivíduo sabe a resposta.
   RA: O indivíduo responde ao acaso.

$$\begin{aligned} P(SR) &= p \\ P(C|SR) &= 1 \\ P(C|RA) &= \frac{1}{n} \end{aligned}$$

   a. Seja $P(C)$ a probabilidade que o teste seja respondido corretamente, dada por:

$$\begin{aligned} P(C) &= P(C|SR)P(SR) + P(C|RA)P(RA) \\ P(C) &= 1 \times p + \frac{1}{n}(1-p) \end{aligned}$$

   já que as únicas dois opções para responder o teste são: sabe a resposta ou responde ao acaso, então $P(SR) + P(RA) = 1$ assim $P(RA) = 1 - p$. Que não responda ao acaso, significa que, indivíduo sabe a resposta. A probabilidade do indivíduo saber a resposta dado que respondeu corretamente é dado por $P(SR|C)$ utilizando o Teorema de Bayes, como segui:

$$\begin{aligned} P(SR|C) &= \frac{P(C|SR)P(SR)}{P(C)} \\ P(SR|C) &= \frac{p}{p + \frac{1-p}{n}} \\ P(SR|C) &= \frac{np}{1 + (n-1)p} \end{aligned}$$

   b. $P(\overline{C}) = 1 - P(C)$, então:

$$P(\overline{C}) = 1 - p - \frac{1}{n}(1-p) \tag{16}$$

   Para a resposta substituir os valores para n=5 e p=0.20 na equação (3).

8. Sejam os eventos:

   D: A peça produzida ser defeituosa.
   A: A peça defeituosa ter sido produzida pela fábrica A;$P(A) = 0.5$.
   B: A peça defeituosa ter sido produzida pela fábrica B, $P(B) = 0.25$.
   C: A peça defeituosa ter sido produzida pela fábrica C, $P(C) = 0.25$.

   As probabilidades de que um circuito integrado produzido por estas fábricas não funcione são:
   Fabrica A : $P(D|A) = 0.01$

Fabrica B : $P(D|B) = 0.04$
Fabrica C : $P(D|C) = 0.03$

Queremos $P(D)$, utilizando a regra da probabilidade total temos:

$$\begin{aligned}
P(D) &= P(A)P(D|A) + P(B)P(D|B) + P(C)P(D|C) \\
P(D) &= 0.5 \times 0.01 + 0.25 \times 0.04 + 0.25 \times 0.03 \\
P(D) &= 0.005 + 0.01 + 0.075 \\
P(D) &= 0.0225
\end{aligned}$$

A probabilidade de um circuito da produção conjunta, das três fábricas, não funcionar é 0,0225.

9. $X \sim P(\lambda = 2)$

   a.

$$\begin{aligned}
&P(X > 3) = 1 - P(X \leq 3) = 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)] \\
&P(X > 3) = 1 - \left(\frac{2^0 e^{-2}}{0!} + \frac{2^1 e^{-2}}{1!} + \frac{2^2 e^{-2}}{2!} + \frac{2^3 e^{-2}}{3!}\right) \\
&P(X > 3) = 0.1428
\end{aligned}$$

   b. E(X)=λ=2

10. X: Número de carros abandonados semanalmente.
    E(X)=$\lambda = 3$

    a. $P(X = 0) = \frac{6^0(e^{-3})}{0!} = e^{-3} = 0.04978$

    b.

$$\begin{aligned}
&P(X \geq 2) = 1 - P(X < 2) = 1 - [P(X = 0) + P(X = 1)] \\
&P(X \geq 2) = 1 - \left(e^{-3} + \frac{3e^{-3}}{1!}\right) \\
&P(X \geq 2) = 1 - 4e^{-3} = 0.8008.
\end{aligned} \tag{17}$$

11. .

X	1	2	25
P(X)	1/3	1/2	1/6

$$\begin{aligned}
E(X) &= 1 \times \frac{1}{3} + 2 \times \frac{1}{2} + 25 \times \frac{1}{6} \\
E(X) &= 5
\end{aligned}$$

A variança é dada por:

$$\begin{aligned}
V(X) &= (1-5)^2 \times \frac{1}{3} + (2-5)^2 \times \frac{1}{2} + (25-5)^2 \times \frac{1}{6} \\
V(X) &= 76.5
\end{aligned}$$

12. $X \sim \mathcal{B}(n,p)$ então $X \sim \mathcal{B}(5, 1/3)$.

Pela fórmula da variança, temos:

$$V(X) = E(X^2) - E(X)^2$$

então

$$V(X) + E(X)^2 = E(X^2). \tag{18}$$

Pela distribuição Binomial, temos que:

$E(X) = np = 5(\frac{1}{3})$

$V(X) = np(1-p) = 5\left(\frac{1}{3}\right)\left(\frac{2}{3}\right).$

Substituindo os valores de $E(X)$ e $V(X)$ em (5), temos que

$$E(X^2) = \frac{35}{9}$$

13. O espaço amostral de dois dados lançados estaria dado por

$$\Omega = \{(1,1),(1,2),....,(6,6)\}$$

Ω possuí 36 valores, e vemos que a probabilidade obter a soma dos resultados de dois dados lançados, são dadas pelas seguinte tabela:

X	2	3	4	5	6	7	8	9	10	11	12
P(X)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Então o valor esperado é dado por:

$$E(X) = 2 \times \frac{1}{36} + ... + 12 \times \frac{1}{36}.$$

14. .

X: Soma dos números anotados.

Nossas possíveis opções são:

$$\{1,2\};\{1,3\};\{2,1\};\{2,3\};\{3,1\};\{3,2\}$$

E a soma é dada por X

X	3	4	5
P(X)	2/6	2/6	2/6

Então,

$$E(X) = 3(2/6) + 4(2/6) + 5(2/6) = 4$$

ME 414 - Gabarito da Lista 3 [2]

Questão 1.

a. Sabemos que $\int_{-\infty}^{\infty} f(x)dx = 1$

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)dx &= \int_{-\infty}^{0} 0dx + \int_{0}^{1} kx^2dx + \int_{1}^{\infty} 0dx \\
&= \int_{0}^{1} kx^2dx = k\int_{0}^{1} x^2dx = k\frac{x^3}{3}\Big|_0^1 \\
&= k(\frac{1}{3}) = 1 \\
&\Rightarrow k = 3
\end{aligned}$$

b.

$$\begin{aligned}
P(1/4 < X < 1/2) &= \int_{1/4}^{1/2} 3x^2 = 3\left(\frac{x^3}{3}\right)\Big|_{1/4}^{1/2} \\
&= \left(\frac{1}{2^3} - \frac{1}{4^3}\right) = 0.1093.
\end{aligned}$$

c.

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} xf(x)dx \\
E(X) &= \int_{0}^{1} x(3x^2)dx = 3\int_{0}^{1} x^3dx \\
E(X) &= 3\left(\frac{x^4}{4}\right)\Big|_0^1 = 3/4 = 0.75.
\end{aligned}$$

Para a variança utilizamos

$$V(X) = E(X^2) - E(X)^2.$$

Vamos a calcular $E(X^2)$

$$\begin{aligned}
E(X^2) &= \int_{0}^{1} x^2f(x)dx = \int_{0}^{1} x^2(3x^2)dx \\
E(X^2) &= 3\int_{0}^{1} x^4dx \\
E(X^2) &= 3\,\frac{x^5}{5}\Big|_0^1 \\
E(X^2) &= 3/5.
\end{aligned}$$

Assim

$$V(X) = E(X^2) - E(X)^2 = 3/5 - 0.75^2 = 0.0375$$

[2]PED: Lisbeth Corbacho Carazas ra162526@ime.unicamp.br

Questão 2.

Seja $X \sim N(5, 16)$ assim $\mu = 5$ e $\sigma = 4$

a.

$$\begin{aligned} P(X \leq 13) &= P\left(Z \leq \frac{13-\mu}{\sigma}\right) \\ &= P\left(Z \leq \frac{8}{4}\right) = P(Z \leq 2) \\ &= 0.977 \end{aligned}$$

b.

$$\begin{aligned} P(X > 1) &= P\left(Z > \frac{1-\mu}{\sigma}\right) = P\left(Z > \frac{-4}{4}\right) \\ &= P(Z > -1) = P(Z \leq 1) = 0.8413 \end{aligned}$$

c.

$$\begin{aligned} P(X \leq a) &= P\left(Z \leq \frac{a-\mu}{\sigma}\right) = P\left(Z \leq \frac{a-5}{4}\right) = 0.04 \\ &\Rightarrow \phi(\frac{a-5}{4}) = 0.04 \\ &\Rightarrow \frac{a-5}{4} = -1.75 \\ &\Rightarrow a = 4(-1.75) + 5 = -2 \end{aligned}$$

Questão 3.

a. Mostraremos que $\int_{-\infty}^{\infty} f(x)dx = 1$

$$\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{\infty} \frac{xe^{-x/2}}{4} dx.$$

Para demostrar isso utilizaremos integração por partes.
$u = x \Rightarrow du = dx$ e $dv = e^{-x/2}dx \Rightarrow v = -2e^{-x/2}$

$$\begin{aligned} \int_{0}^{\infty} f(x)dx &= uv - \int_{0}^{\infty} vdu \\ &= \frac{1}{4}\left\{x(-2e^{-x/2}) - \int_{0}^{\infty} -2e^{-x/2}dx\right\} \\ &= \frac{1}{4}\left\{x(-2e^{-x/2}) + 2\int_{0}^{\infty} e^{-x/2}dx\right\} \\ &= \frac{1}{4}\left\{x(-2e^{-x/2}) + 2(-2e^{-x/2})\right\}\Big|_{0}^{\infty} \\ &= 1 \end{aligned}$$

note que para substituir os valores é preciso saber $e^{-\infty/2} = \frac{1}{e^{\infty/2}} \to 0$.

b. 6 meses equivale a meio ano (1/2)

$$\begin{aligned} p(X \leq 1/2) &= \int_0^{1/2} f(x)dx = \int_0^{1/2} \frac{xe^{-x/2}}{4} dx \\ &= \frac{1}{4}\left(-2xe^{-x/2} - 4e^{-x/2}\right)\Big|_0^{1/2} \\ &= \frac{1}{4}\left(-5e^{-1/4} + 4\right) = 2.65\% \end{aligned}$$

Questão 4.

X: Peso de um determinado produto tal que $X \sim N(\mu, 20^2 g^2)$

a.

$$\begin{aligned} P(X < 500) &= P(Z < \frac{500-\mu}{20}) = 0.1 \\ &\Rightarrow \phi\left(\frac{500-\mu}{20}\right) = 0.1 \\ &\Rightarrow \left(\frac{500-\mu}{20}\right) = -1.28 \\ &\Rightarrow \mu = 525.6 \end{aligned}$$

b. Sejam X_1, X_2, X_3, X_4 pesos dos 4 pacotes da amostra.

Estão pedindo a probabilidade do peso total menor que 2kg, isto é, $\sum_{i=1}^{4} X_i < 2000g$ e é o mesmo que, $\overline{X} < \frac{2000}{4}$, assim a média amostral

$$\overline{X} \sim N\left(525.6g, \frac{20^2}{4}g^2\right) = N\left(525.6g, \left(\frac{20}{2}g\right)^2\right)$$

então

$$\begin{aligned} P(\overline{X} < 500) &= P\left(Z < \frac{500-525.6}{10}\right) \\ &= P(Z < -2.56) = 0.0052 \end{aligned}$$

Questão 5.

a.

$$\overline{X} \sim N\left(525.6g, \left(\frac{20}{2}g\right)^2\right)$$

A probabilidade de ser feita uma parada desnecessária é dada por:

$$
\begin{aligned}
& P\left(\overline{X} < 495.6 \cup \overline{X} > 555.6\right) = P\left(\overline{X} < 495.6\right) + P\left(\overline{X} > 555.6\right) \\
= \; & P\left(Z < \frac{495.6 - 525.6}{10}\right) + P\left(Z > \frac{555.6 - 525.6}{10}\right) \\
= \; & P(Z < -3) + P(Z > 3) \\
= \; & 2P(Z > 3) = 2(1 - P(Z < 3)) \\
= \; & 2(1 - 0.9987) = 0.0026
\end{aligned}
$$

b. $\overline{X} \sim N\left(510g, \left(\frac{20}{2}g\right)^2\right)$. A probabilidade de continuar a produção fora dos padrões desejados seria:

$$
\begin{aligned}
P\left(495.6 < \overline{X} < 555.6\right) &= P\left(Z < 555.6\right) - P\left(\overline{X} < 495.6\right) \\
&= P\left(Z < \frac{555.6 - 510}{10}\right) - P\left(Z < \frac{495.6 - 510}{10}\right) \\
&= P(Z < 4.56) + P(Z < -1.44) \\
&= 1 - 0.0749 \\
&= 0.09251.
\end{aligned}
$$

Questão 6.

X: Quantidade de ácido xanturênico excretado na urina.

$$
X \sim N\left(\frac{4.8mg}{15ml}, \left(\frac{2mg}{15ml}\right)^2\right)
$$

a. Para fines do cálculo desconsideramos o denominador dos dados de entrada.

$$
\begin{aligned}
P(2.8 < X < 7) &= P\left(\frac{2.8 - 4.8}{2} < Z < \frac{7 - 4.8}{2}\right) \\
&= P(-1 < Z < 1.5) \\
&= \phi(1.1) - \phi(-1) = 0.7056
\end{aligned}
$$

b.

$$
\begin{aligned}
P(X \leq x) &= P\left(Z < \frac{x - 4.8}{2}\right) = 0.1 \\
&= \phi\left(\frac{x - 4.8}{2}\right) = 0.1 \\
&\Rightarrow \frac{x - 4.8}{2} = -1.28 \\
&\Rightarrow x = 2.24
\end{aligned}
$$

Deve ter uma quantidade de acido xanturênico de 2.24mg/15ml.

c. Y: Número de pessoas com quantidade de ácido xanturênico anormal.

Veja que a probabilidade da quantidade de acido xanturênico anormal é determinada a partir do item a.

$$p = 1 - P(2.8 < X < 7) = 1 - 0.7056 = 0.2944$$

Então para $n = 10$, $p = 0.225$ e $Y \sim \mathcal{B}(10, 0.225)$ temos:

$$\begin{aligned} P(Y \leq 2) &= P(Y = 0) + P(Y = 1) + P(Y = 2) \\ &= \binom{10}{0}(0.2944)^0(0.7055)^{10} + \binom{10}{1}(0.2944)(0.7056)^9 \\ &+ \binom{10}{2}(0.2944)^2(0.7056)^8 \\ &= 0.31 + 0.128 + 0.240 = 0.398 \end{aligned}$$

Questão 7.

Seja $X \sim U([0, 5])$, nosso interesse é $Y = X^2$, então

$$G(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F(\sqrt{y}) - F(-\sqrt{y})$$

derivando $G'(y)$ obtemos a densidade

$$\begin{aligned} G'(y) &= F(\sqrt{y})' - F(-\sqrt{y})' \\ &= F'(\sqrt{y})\frac{1}{2\sqrt{y}} - F'(-\sqrt{y})\frac{1}{2\sqrt{y}} \\ &= \frac{1}{2\sqrt{y}}(f(\sqrt{y}) - f(-\sqrt{y})) \end{aligned}$$

Se $f(x) = \frac{1}{5}$ para $0 < x < 5$, então

$$g(y) = \frac{1}{10\sqrt{y}}$$

Para os limites de integração tem-se $y = 0$ e $y = 25$ (para $Y = X^2$, assim:

$$\begin{aligned} E(Y) &= \int_0^{25} y\frac{1}{10\sqrt{y}}dy = \frac{1}{10}\int_0^{25}\sqrt{y}dy \\ &= \frac{1}{15}y^{3/2}\Big|_0^{25} = \frac{25}{3}. \end{aligned}$$

Existe uma forma mais fácil de fazer, é a seguinte:

$$E(Y) = E(X^2) = \int_0^5 x^2 f(x)dx = \int_0^5 \frac{x^2}{5} = \frac{x^3}{15}\Big|_0^5 = \frac{25}{3}.$$

Questão 8.

a. $P(X \leq 10) = 1 - e^{10} = 0.999$

b. $P(5 < X < 15) = P(X < 15) - P(X < 5) = e^{-5} - e^{-15} = 0.0067$

c. $P(X > t) = 0.01 \Leftrightarrow e^{-t} = 1/100 \Leftrightarrow t = 4.605$

Questão 9.

Seja $X \sim exp(3)$, nosso interesse é $Y = X^3$, então

$$E(Y) = E(X^3) = \int_0^\infty x^3 f(x)dx = \int_0^\infty 3x^3 e^{-3x} dx$$

Para resolver essa integral usaremos integração por partes, considerando $u = x^3$ e $dv = 3e^{-3x}dx \Rightarrow du = 3x^2dx$ e $v = -e^{-3x}$

$$E(Y) = -x^3 e^{-3x}\Big|_0^\infty + \int_0^\infty 3x^2 e^{-3x} dx = \int_0^\infty 3x^2 e^{-3x} dx.$$

Aplicando novamente integração por partes, com $u = x^2$ e $dv = 3e^{-3x}dx \Rightarrow du = 2xdx$ e $v = -e^{3x}$, então:

$$E(Y) = -x^2 e^{-3x}\Big|_0^\infty + \int_0^\infty 2xe^{-3x} dx = \int_0^\infty 2xe^{-3x} dx$$

Aplicando novamente integração por partes, com $u = x$ e $dv = 2e^{-3x}dx \Rightarrow du = 2xdx$ e $v = \frac{-2}{3}e^{3x}$, então:

$$E(Y) = -\frac{2x}{3}e^{-3x}\Big|_0^\infty + \int_0^\infty \frac{2}{3}e^{-3x} dx = -\frac{2}{9}e^{-3x}\Big|_0^\infty = \frac{2}{9}$$

Questão 10.

Seja $T \sim exp(2)$

$$\begin{aligned} P(X = 0) &= P(0 \leq T < 1) = P(T < 1) - P(T \leq 0) \\ &= (1 - e^{-2}) - (1 - e^{0}) = 1 - e^{-2} = 0.865 \end{aligned}$$

$$\begin{aligned} P(X = 1) &= P(1 \leq T < 2) = P(T < 2) - P(T \leq 1) \\ &= (1 - e^{-4}) - (1 - e^{-2}) = e^{-2} - e^{-4} = 0.117 \end{aligned}$$

$$\begin{aligned} P(X = 2) &= P(T \geq 2) = 1 - P(T \leq 2) \\ &= 1 - (1 - e^{-4}) = e^{-4} = 0.018 \end{aligned}$$

Questão 11.

Seja $X \sim (\mu, \sigma^2)$

a.

$$\begin{aligned} P(X \leq \mu + 2\sigma) &= P(\frac{X-\mu}{\sigma} \leq 2) \\ &= P(Z \leq 2) = 0.9772 \end{aligned}$$

b. Utilize a propriedade

$$\begin{aligned} P(|X-\mu| \leq \sigma) &= P(-\sigma \leq X - \mu \leq \sigma) \\ &= P\left(-1 \leq \frac{X-\mu}{\sigma} \leq 1\right) \\ &= P(-1 \leq Z \leq 1) \\ &= P(Z \leq 1) - P(Z \leq -1) \\ &= 0.8413 - 0.1587 = 0.6826 \end{aligned}$$

c.

$$\begin{aligned} P(-a\sigma + \mu \leq X \leq a\sigma + \mu) &= P(-a\sigma \leq X - \mu \leq a\sigma) \\ &= P\left(-a \leq \frac{X-\mu}{\sigma} \leq a\right) \\ &= P(-a \leq Z \leq a) = 0.99 \\ &= P(Z \leq a) - P(Z \leq -a) \\ &= 1 - 2P(Z \leq -a) = 0.99 \\ &= 1 - 2(1 - P(Z \leq a)) \\ &= 2P(Z \leq a) - 1 = 0.99 \\ &\Rightarrow P(Z \leq a) = 0.995 \\ &\Rightarrow \phi(a) = 0.995 \\ &\Rightarrow a = 2.56 \end{aligned}$$

d.

$$\begin{aligned} P(X > a) &= 1 - P(X < a) = 0.90 \\ &\Rightarrow P(X < a) = 0.01 \\ &\Rightarrow P(\frac{X-\mu}{\sigma} < \frac{a-\mu}{\sigma}) = 0.01 \\ &\Rightarrow P(Z < \frac{a-1}{\sqrt{2}}) = 0.1 \\ &\Rightarrow \phi(\frac{a-1}{\sqrt{2}}) = 0.1 \\ &\Rightarrow \frac{a-1}{\sqrt{2}} = -1.285 \\ &\Rightarrow a = -0.8173 \end{aligned}$$

Estatística - ME 414: Gabarito da Lista 4 [3]

Questão 1.
X= número de pessoas acima de 40 anos que têm artrite Para uma amostra de $X = 240$ de $n = 4000$ tem-se:

a. $\hat{p} = \frac{X}{n} = \frac{240}{4000} = 0.06$

b. $IC(p; 0.95) = \left[\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$, $z_{\alpha/2}$ é o percentil $\alpha/2$ da distribuição Normal então $z_{\alpha/2} = z_{0.025} = 1.96$, assim:

$$\begin{aligned} IC(p; 0.95) &= \left[0.06 - 1.966\sqrt{\frac{0.06(1-0.06)}{4000}}, 0.06 + 1.96\sqrt{\frac{0.06(1-0.06)}{4000}}\right] \\ &= [0.052, 0.067] \end{aligned}$$

A um grau de confiança de 95% estimamos que a proporção de pessoas acima de 40 anos pertence ao intervalo [0.052,0.067]

Questão 2.
p=Proporção de pessoas que preferem a marca X de sabonete, $\hat{p} = 0.70$, n=625, $z_{\alpha/2} = z_{0.05} = 1.64$, então:

$$IC(p; 0.90) = \left[0.70 - 1.64\sqrt{\frac{0.70(1-0.70)}{625}}, 0.06 + 1.64\sqrt{\frac{0.70(1-0.70)}{625}}\right] = [0.67, 0.73] \tag{19}$$

a um grau de confiança de 90% é estimado que a proporção de pessoas preferem a marca X de sabonete pertence ao intervalo [0.67,0.73].

Questão 3.
p=Proporção de pessoas que são favoráveis a um determinado partido $\hat{p} = 0.6$ e $n = 100$.

- a. Seja ME a margem de erro $\alpha = 0.2$

$$ME = z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = z_{0.1}\sqrt{\frac{0.6(1-0.6)}{n}} = 0.01 \Rightarrow n = \frac{0.24(1.285)^2}{0.01^2} = 3962.94 \tag{20}$$

assim, o tamanho da amostra seria 3963 para que o erro cometido na estimativa seja não máximo 0.01 com compatibilidade 0.8

b. Se $n = 3963$ $\hat{p} = 0.55$ e $z_{\alpha/2} = z_{0.025} = 1.96$

$$\begin{aligned} IC(p; 0.95) &= \left[0.55 - 1.96\sqrt{\frac{0.55(1-0.55)}{3963}}, 0.55 + 1.96\sqrt{\frac{0.0.55(1-0.55)}{3963}}\right] \\ &= [0.534, 0.565] \end{aligned}$$

[3]PED: Lisbeth Corbacho Carazas, ra162526@ime.unicamp.br

com grau de confiança ao 95% é estimado que a proporção de pessoas são favoráveis a um determinado partido está no intervalo [0.534,0.565].

Questão 4.
p=Proporção de consumidores de um certo produto.
$n = 300$ $\hat{p} = \frac{1}{3}$.

a. $z_{\alpha/2} = z_{0.025} = 1.96$

$$IC(p; 0.95) = \left[\frac{1}{3} - 1.96\sqrt{\frac{\frac{1}{3}(1-\frac{1}{3})}{300}}, \frac{1}{3} + 1.96\sqrt{\frac{\frac{1}{3}(1-\frac{1}{3})}{300}}\right] = [0.280; 0.387] \tag{21}$$

b. ME margem de erro

$$\begin{aligned} ME &= z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = z_{0.025}\sqrt{\frac{\frac{1}{3}(1-\frac{1}{3})}{n}} = 1.96\sqrt{\frac{2/9}{n}} = 0.02 \\ \Rightarrow n &= \frac{2(1.96)^2}{9(0.02)^2} = 2134.22 \end{aligned}$$

Portanto, o tamanho da amostra seria 2135 para que o erro cometido na estimativa seja não máximo 0.02.

Questão 5.
Sejam $X_1, ..., X_n$ amostras aleatórias com distribuição $N(\mu, \sigma^2)$

$$E(T) = E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} E(a_i X_i) = \sum_{i=1}^{n} a_i E(X_i) = \sum_{i=1}^{n} a_i \mu_i = \mu \sum_{i=1}^{n} a_i = \mu(1) = \mu.$$

Portanto, T é um estimador não viciado a média μ

Questão 6.

Seja $n = 50$, $\bar{x} = 20$ e $\sigma = 2$

a.

$$\begin{aligned} IC(\mu, 1-\alpha) &= \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] \\ IC(\mu, 0.99) &= \left[20 - 2.575\frac{2}{\sqrt{50}}, 20 + 2.575\frac{2}{\sqrt{50}}\right] = [19.772, 21.228] \end{aligned}$$

A um grau de confiança igual a 99% é estimado que a média populacional do tempo gasto para montar o brinquedo está entre 19.772 e 21.228.

b. $\alpha = 0.01 \Rightarrow z_{0.005} = 2.575$ e $ME = \epsilon = 0.10$

$$\epsilon = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \Rightarrow n = \sigma^2\left(\frac{z_{\alpha/2}}{\epsilon}\right) = 4\left(\frac{2.575}{0.10}\right)^2 = 2652.25 \tag{22}$$

Portanto o tamanho amostral deverá ser $m = 2653$.

Questão 7.

Sejam $n = 25$, $\alpha = 0.05$ então $t_{24;0.025} = 2.064$ $\bar{x} = 97.64$ e $s = 17.821$.

$$\begin{aligned} IC(\mu, 1-\alpha) &= \left[\bar{x} - t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}, \bar{x} + t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}\right] \\ IC(\mu, 0.95) &= \left[97.64 - 2.064\frac{17.821}{\sqrt{25}}, 97.64 + 2.064\frac{17.821}{\sqrt{25}}\right] = [90.283, 104.997]. \end{aligned}$$

A um grau de confiança do 95% estima-se a média populacional do nível de glicêmico pertence ao intervalo [90.283,104.997].

Questão 8.

Sejam $n = 100$, $\bar{x} = 500$, $\sigma = 5$ e $\alpha = 0.05$ então $z_{0.025} = 1.96$

$$\begin{aligned} IC(\mu, 1-\alpha) &= \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] \\ IC(\mu, 0.95) &= \left[500 - 1.96\frac{5}{\sqrt{100}}, 500 + 1.96\frac{5}{\sqrt{100}}\right] = [499.02, 500.98]. \end{aligned}$$

A um grau de confiança do 95% estima-se a média populacional do tempo de vida de uma peça de equipamento pertence ao intervalo [499.02,500.98].

Questão 9.

Seja $\alpha = 0.10$ então $z_{0.05} = 1.645$ e C o cumprimento do intervalo de confiança

$$\begin{aligned} IC(\mu, 1-\alpha) &= \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] \\ \Rightarrow \quad & C = 2z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < 0.2\sigma \\ \Rightarrow \quad & n > \left(\frac{2z_{\alpha/2}\sigma}{0.20\sigma}\right)^2 = \left(\frac{2(1.645)}{0.20}\right) = 16.45^2 = 270.60 \\ \Rightarrow \quad & n \geq 271 \end{aligned}$$

Questão 10.

Os dois grupos são amostras aleatórias de duas populacionais independentes e normalmente distribuídas. Os tamanhos das amostras são $n_1 = n_2 = 18$, as medidas

amostrais $\bar{x}_1 = 6.622$ e $\bar{x}_2 = 5.750$ e as variâncias amostrais $s_1^2 = 1.325$ e $s_2^2 = 0.739$. Vamos supor que as variâncias populacionais são iguais e desconhecidas, portanto o estimador da variança populacional é:

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{17(1.325)+17(0.739)}{34} = 1.032. \tag{23}$$

Assim, o intervalo de confiança ao 95% para diferença das médias populacionais para $\alpha = 0.05$, $t_{34;0.025} = 2.032$ é dado por:

$$\begin{aligned}
IC(\mu_1-\mu_2, 1-\alpha) &= \left[(\bar{x_1}-\bar{x_2}) - t_{n_1+n_2-2,\alpha/2}\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}, (\bar{x_1}-\bar{x_2}) + t_{n_1+n_2-2,\alpha/2}\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right] \\
IC(\mu_1-\mu_2, 0.95) &= \left[(6.622-5.750) - 2.032\sqrt{1.032\left(\frac{2}{18}\right)}, (6.622-5.750) + 2.032\sqrt{1.032\left(\frac{2}{18}\right)}\right] \\
IC(\mu_1-\mu_2, 0.95) &= [0.184, 1.560].
\end{aligned}$$

A um grau de confiança de 95% estima-se que a diferença entre os resultados médios obtidos entre o grupo 1 e grupo 2 está no intervalo [0.184,1.560].

Questão 11.
Sejam $n_A = 100$ $\hat{p}_A = 0.75$, $n_B = 100$, $\hat{p}_B = 0.65$ e $\hat{p} = \frac{75+65}{200} = 0.70$

Será testado:

$$H_0 : p_A - p_B = 0 \quad vs \quad H_a : p_A - p_B > 0$$

$$z_{obs} = \frac{\hat{p}_A - \hat{p}_B}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}} = \frac{0.75-0.65}{\sqrt{0.70(1-0.70)\left(\frac{1}{100}+\frac{1}{100}\right)}} = 1.543 \tag{24}$$

valor de p: $P(Z \geq z_{obs}) = P(Z \geq 1.543) = 1 - P(Z < 1.543) = 1 - 0.939 = 0.061$. Se consideramos $\alpha \leq 0.05$ então com valor de p maior que α não rejeitamos a hipótese nula de que a verdadeira proporção de curados usando soro é igual a verdadeira proporção de curados não usando soro.

Questão 12.
Sejam $\hat{p}_1 = \frac{54}{150} = 0.36$, $\hat{p}_2 = \frac{33}{100} = 0.348$ e $\hat{p} = \frac{75+65}{200} = 0.70$

Será testado:

$$H_0 : p_1 - p_2 = 0 \quad vs \quad H_a : p_1 - p_2 > 0$$

$$z_{obs} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{0.36-0.33}{\sqrt{0.348(1-0.348)\left(\frac{1}{150}+\frac{1}{100}\right)}} = 0.488 \tag{25}$$

valor-de-p: $P(Z \geq z_{obs}) = P(Z \geq 0.488) = 1 - P(Z < 0.488) = 1 - 0.6872 = 0.3128$. Como o valor-de-p maior que 0.05 não rejeitamos a hipótese nula, isto é, que dois métodos tem a mesma proporção de sucesso para provocar chuva.

Questão 13.

Seja p: a proporção de associados que apoiam a política de privatização do governo.

Sejam $\alpha = 0.05$ e $n = 80$.

Será testado:

$$H_0 : p = 0 \quad vs \quad H_a : p > 0.60$$

Estatística do Teste: $Z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$.

Valor crítico $z_{crit} = z_{0.05} = 1.64$.

Região de rejeição: Rejeitamos H_0 se $Z \geq 1.64$ então:

$$C = \left\{ (X_1, X_2, ..., X_n) : \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \geq 1.64 \right\}.$$

Questão 14.

a. Sejam $n = 15$, $\bar{X} = 83.9$, $s = 18.2$ e $\alpha = 0.10$.
Será testado:

$$H_0 : \mu = 85 \quad vs \quad H_a : \mu < 85$$

Estatística do Teste: $t_{obs} = \frac{\bar{X}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{83.9-85}{\frac{18.2}{\sqrt{15}}} = -0.234$.

Valor crítico $t_{crit} = t_{14:0.10} = -1.345$ menor que o valor observado da estatística do teste ($t_{obs} = -0.234$) não rejeitamos a hipótese que a média é 85.

b. Sejam $n = 15$, $\bar{X} = 79.1$, $s = 11.8$ e $\alpha = 0.10$.
Será testado:

$$H_0 : \mu = 76 \quad vs \quad H_a : \mu \neq 76$$

Estatística do Teste: $t_{obs} = \frac{\bar{X}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{70.10-76}{\frac{11.8}{\sqrt{15}}} = 1.017$
Valor crítico $t_{crit} = t_{14:0.05} = 1.761$ menor que o valor observado da estatística do teste ($t_{obs} = 1.017$) não rejeitamos a hipótese que a média é 76.

Questão 15.

Sejam $n = 16$, $\bar{X} = 94.32$, $\sigma = 1.20$ e $\alpha = 0.10$.

a. Será testado:

$$H_0 : \mu = 95 \quad vs \quad H_a : \mu \neq 95$$

Estatística do Teste: $z_{obs} = \frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{94.32-95}{\frac{1.2}{\sqrt{16}}} = -2.267$.

Valor-de-p $P(|Z| \geq |z_{obs}|) = P(|Z| \geq 2.267) = 2P(Z \geq 2.267) = 2[1 - P(Z < 2.267)] = 0.0234$.

Com valor-de-p maior que 0.01 é concluído que existe evidência para não rejeitar a hipótese de que o ponto de desvanecimento médio de uma certe marca de vegetais hidrogenados é 95.

b. Seja $\alpha = 0.01$ então $z_{crit} = 2.575$ e $\mu' = 94$ assim $\bar{X} \sim N(\mu', \sigma^2/n))$.

$$\begin{aligned}
\beta &= P(\text{não rejeitar} H_0 | H_0 \text{é falsa}) = P(|Z| < z_{crip} | \mu' = 94) \\
&= P\left(\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| < 2.575 | \mu' = 94\right) \\
&= P\left(-2.575 < \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < 2.575 | \mu' = 94\right) \\
&= P\left(-2.575 < \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < 2.575 | \mu' = 94\right) \\
&= P\left(-2.575 + \frac{\mu_0-\mu'}{\sigma/\sqrt{n}} < \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < 2.575 + \frac{\mu_0-\mu'}{\sigma/\sqrt{n}} | \mu' = 94\right) \\
&= P\left(-2.575 + \frac{95-94}{1.20/\sqrt{16}} < \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < 2.575 + \frac{95-94}{1.20/\sqrt{16}}\right) \\
&= P(0.758 < Z < 5.908) = 0.2242
\end{aligned}$$

Questão 16.

Sejam $n = 16$, $\bar{X} = 5.25$, $\sigma = 0.30$.
Será testado:

$$H_0 : \mu = 5.5 \quad vs \quad H_a : \mu \neq 5.5$$

Estatística do Teste: $z_{obs} = \frac{\bar{X}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{5.25-5.5}{\frac{0.30}{\sqrt{16}}} = -3.33$.

Valor-de-p: $P(|Z| \geq |z_{obs}|) = P(|Z| \geq 3.333) = 2(1 - P(Z < 3.333)) = 0.0008$.

Com um valor-de-p muito menor que 0.05 rejeitamos a hipótese de que o ponto médio de SiO_2 em certo tipo de cimento aluminoso é de 5.5.

Questão 17.

Sejam $n_H = 97$, $\bar{X}_H = 10.40$, $\sigma_H = 4.83$, $n_M = 148$, $\bar{X}_M = 9.26$, $\sigma_M = 4.86$ e $\alpha = 0.05$.

Será testado:

$$H_0 : \mu_H - \mu_M = 0 \;\; vs \;\; H_a : \mu_M - \mu_H > 0.60$$

Assumindo que a classificação fornecida pela escala possui distribuição normal e que as variancas são conhecidas e diferentes.

Estatística do Teste: $z_{obs} = \frac{\bar{X}_H - \bar{X}_M - \triangle_0}{\sqrt{\frac{\sigma_H^2}{\sqrt{n_H}} + \frac{\sigma_M}{\sqrt{n_M}}}} = \frac{10.40-9.26}{\sqrt{\frac{4.83^2}{97} + \frac{4.86^2}{148}}} = 1.802.$

Valor-de-p: $P(Z > z_{obs}) = P(Z > 1.802) = 1 - P(Z < 1.802) = 0.036.$

Com valor-de-p menor que 0.05 existe evidência para rejeitar a hipótese de que não existe diferença entre a facilidade que os estudantes homens es as estudantes mulheres entediam-se.

Questão 18.

Supondo que a duração das superfícies de rodagem possuem distribuição Normal. Sejam $\alpha = 0.05$, $m = 40$, $\bar{X} = 36500$, $\sigma_1 = 2200$, $n = 40$, $\bar{Y} = 33400$ e $\sigma_2 = 1900$.

Será testado:

$$H_0 : \mu_1 - \mu_2 = 0 \;\; vs \;\; H_a : \mu_1 - \mu_2 \neq 0$$

Estatística do Teste: $z_{obs} = \frac{\bar{X} - \bar{Y} - \triangle_0}{\sqrt{\frac{\sigma_1}{\sqrt{m}} + \frac{\sigma_2}{\sqrt{n}}}} = \frac{36500-33400}{\sqrt{\frac{2200^2}{40} + \frac{1900^2}{40}}} = 6.745.$

Valor-de-p: $P(|Z| > z_{obs}) = P(|Z| > 6.745) = 2(1 - P(Z < 6.745)) \simeq 0.$

Com valor-de-p menor que 0.05 existe evidência para rejeitar a hipótese de que os pontos médios de duração das superfícies de rodagem são iguais.

Questão 19.

Supondo que os dado possuem uma distribuição Normal. Sejam $\alpha = 0.01$, $m = 40$, $\bar{X} = 18.12$, $\sigma_1 = 1.60$, $n = 32$, $\bar{Y} = 16.87$ e $\sigma_2 = 1.40$.

Será testado:

$$H_0 : \mu_1 - \mu_2 = 0 \;\; vs \;\; H_a : \mu_1 - \mu_2 > 0$$

Estatística do Teste: $z_{obs} = \frac{\bar{X} - \bar{Y} - \triangle_0}{\sqrt{\frac{\sigma_1}{\sqrt{m}} + \frac{\sigma_2}{\sqrt{n}}}} = \frac{18.12-16.87}{\sqrt{\frac{1.60^2}{40} + \frac{1.40^2}{32}}} = 3.532.$

Valor-de-p: $P(Z > z_{obs}) = P(Z > 3.532) = 1 - P(Z < 3.532) = 1 - 0.9998 = 0.0002.$

Com valor-de-p menor que 0.01 existe evidência para rejeitar a hipótese nula de que a resistência média de coesão à tensão do morteiro modificado de látex de polímeros é igual a do morteiro não modificado.

Questão 20.

Supondo que o= conteúdo de nicotina das duas marcas tem distribuição Normal que as duas variança populacionais são iguais. Sejam $\alpha = 0.05$, $n_A = 40$, $\bar{X}_A = 20$, $s_A^2 = 6$, $n_B = 5$, $\bar{X}_B = 21$ e $s_B^2 = 5$. Com as variança populacionais desconhecidas e iguais, será determinado o estimados da variança:

$$S_p^2 = \frac{S_A^2(n_A - 1) + S_B^2(n_B - 1)}{n_A + n_B - 2} = \frac{6(4-1) + 5(4-1)}{4+5-2} = 5.429.$$

Será testado:

$$H_0 : \mu_1 - \mu_2 = 0 \quad vs \quad H_a : \mu_1 - \mu_2 \neq 0$$

Estatística do Teste: $t_{obs} = \frac{\bar{X}_A - \bar{X}_B - \triangle_0}{\sqrt{S_p^2\left(\frac{1}{\sqrt{n_A}} + \frac{1}{\sqrt{n_B}}\right)}} = \frac{20-21}{\sqrt{5.429(\frac{1}{4}+\frac{1}{5})}} = -0.639.$

Valor-de-p: $P(|T| > t_{obs}) = P(|T| > 0.639) = 2(1 - P(Z < 0.639)) = 2(1 - 0.728) = 0.544$

Com valor-de-p menor que 0.05 existe evidência para não rejeitar a hipótese de que o conteúdo médio de nicotina das duas marcas de cigarros é o mesmo.