Disciplina: INFERÊNCIA ESTATÍSTICA
Curso: Graduação em Estatística
Código: EST0035
Semestre: 2025.1
Professor: Frederico Machado Almeida

LISTA DE EXERCÍCIOS #03

Observações:

- Questões para entregar: 3, 8, 10 *(b, d)*, 12 e 13.
- Demais questões são apenas para estudar.
- Prazo de entrega: **27/05/2025**

Q1. Seja $\bar{x} = 81,2$ a média observada de uma amostra aleatória simples de tamanho 20, proveniente de uma distribuição normal de média μ e variância $\sigma^2 = 80$. Encontre: (i) o intervalo de 90% de confiança para μ; (ii) o intervalo de 96% de confiança para θ, e compare a amplitude dos dois intervalos.

Resposta: Com base no enunciado, $\bar{x} = 81,2$ denota a estimativa pontual para μ obtida em uma amostra aleatória simples de tamanho $n = 20$. De igual forma, temos que μ é uma quantidade desconhecida, e $\sigma^2 = 80$. Nisso, segue imediatamente que,

(i) Para um nível de confiança de $\gamma = 1-\alpha = 0,90$ obtemos da tabela de distribuição normal padrão o quantil $z_{\alpha/2} = z_{1-\alpha/2} = 1,645$ (valores de $z_{\alpha/2} = 1,64$ ou 1,65 podem ser igualmente considerados). Nisso, com base na informação anterior, segue que, a quantidade pivotal para esse problema será dada pela estatística Z, isto é:

$$Q(\mathbf{X}, \mu) = \frac{(\bar{x} - \mu)}{\sqrt{\sigma^2/n}} = \mathbb{P}\left(-z_{\alpha/2} \leq \frac{(\bar{x} - \mu)}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right) \sim \mathcal{N}(0, 1).$$

Assim, para algum q_1 e q_2 da distribuição normal padrão, tal que $q_2 = -z_{\alpha/2} = -1,645$ e $q_1 = -q_2 = 1,645$ obtemos:

$$\mathbb{P}\left(-z_{\alpha/2} \leq Q(\mathbf{X}, \mu) \leq z_{\alpha/2}\right) = \mathbb{P}\left(-z_{\alpha/2} \leq \frac{(\bar{x} - \mu)}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right) = 1 - \alpha$$

$$\mathbb{P}\left(-z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \bar{x} - \mu \leq z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = \mathbb{P}\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha.$$

$$\mathbb{P}(81,2 - 1,645 \times 2 \leq \mu \leq 81,2 + 1,645 \times 2) = 0,90.$$

Portanto, o intervalo de 90% de confiança para μ é dado por: $IC_{0,9}(\mu) = [77,91; 84,49]$. Como interpretação, podemos afirmar que, construindo um número muito grande de intervalos de confiança nas mesmas condições, espera-se que 90% destes contenham o verdadeiro valor do parâmetro.

(ii) Agora, supondo um nível de confiança de $\gamma = 1 - \alpha = 0,96$ e assumindo todas as condições anteriores, temos que, $z_{\alpha/2} = z_{1-\alpha/2} = 2,054$. Desta forma, o intervalo de confiança para μ é obtido usando a mesma ideia de (i). Ou seja,

$$\mathbb{P}\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = \mathbb{P}\left(81,2 - 2,054 \times 2 \leq \mu \leq 81,2 + 2,054 \times 2\right) = 0,96$$

Ou seja, o intervalo de 96% de confiança para μ, é dado por: $IC_{0,96}(\mu) = [77,09; 85,31]$. Comparando as amplitudes dos dois intervalos, observamos que: de (i) $A_1 = L_s - L_i = 84,49 - 77,91 = 6,58$ e de (ii) $A_2 = L_s - L_i = 85,31 - 77,09 = 8,22$. Portanto, como já era esperado, o intervalo de confiança associado ao maior nível de confiança (intervalo obtido em (ii)) apresentou maior valor para amplitude.

Q2. Suponha que uma amostra aleatória de tamanho 17 tenha sido extraída da distribuição $X \sim \mathcal{N}(\mu, \sigma^2)$ tenha fornecido uma média de $\bar{x} = 4,7$ e variância $s^2 = 5,76$. Construa os intervalos de 95 e 99% de confiança para μ.

Resposta: Com base no enunciado, $\bar{x} = 4,7$ representa a estimativa pontual para μ, obtida a partir de uma amostra aleatória simples de tamanho $n = 17$. Como a variância populacional é desconhecida, utilizamos a variância amostral $s^2 = 5,76$. Nesse contexto, sabe-se que a quantidade pivotal para este problema segue uma distribuição t de Student com $n - 1 = 16$ graus de liberdade, dada por:

$$Q(\mathbf{X}, \mu) = \frac{\bar{x} - \mu}{S/\sqrt{n}} \sim t_{(16)}$$

(i) Intervalo de confiança de 95%

Para um nível de confiança $\gamma = 1 - \alpha = 0,95$, obtemos da tabela t de Student os quantis:

$$t_{\alpha/2;16} = t_{(0,025;16)} \approx 2,1199$$

Com base na distribuição da quantidade pivotal, temos:

$$\mathbb{P}\left(-t_{(0,025;16)} \leq \frac{\bar{x} - \mu}{S/\sqrt{n}} \leq t_{(0,025;16)}\right) = 0,95$$

Multiplicando todos os termos por $\dfrac{S}{\sqrt{n}}$ e reorganizando:

$$\mathbb{P}\left(\bar{X} - t_{(0,025;16)} \times \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{(0,025;16)} \times \frac{S}{\sqrt{n}}\right) = 0,95$$

Substituindo os valores numéricos:

- $\bar{x} = 4,7$
- $s = \sqrt{5,76} = 2,4$
- $\dfrac{s}{\sqrt{n}} = \dfrac{2,4}{\sqrt{17}} \approx 0,582$

$$\Rightarrow \mathbb{P}(4,7 - 2,1199 \times 0,582 \leq \mu \leq 4,7 + 2,1199 \times 0,582) = 0,95$$

$$\Rightarrow \mathbb{P}(3,47 \leq \mu \leq 5,93) = 0,95$$

(ii) Intervalo de confiança de 99%

Agora, para $\gamma = 0,99 \Rightarrow \alpha = 0,01$, temos:

$$t_{(0{,}005;16)} \approx 2,9208$$

Repetindo o mesmo processo:

$$\mathbb{P}\left(\bar{X} - t_{(0{,}005;16)} \times \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{(0{,}005;16)} \times \frac{S}{\sqrt{n}}\right) = 0,99$$

$$\Rightarrow \mathbb{P}(4,7 - 2,9208 \times 0,582 \leq \mu \leq 4,7 + 2,9208 \times 0,582) = 0,99$$

$$\Rightarrow \mathbb{P}(3,000 \leq \mu \leq 6,399) = 0,99$$

Conclusões finais:

- Intervalo de confiança de 95% para μ: $IC_{95\%}(\mu) = [3,47;\ 5,93]$
- Intervalo de confiança de 99% para μ: $IC_{99\%}(\mu) = [3,00;\ 6,40]$

Vê-se que a amplitude do $IC_{99\%}(\mu)$ é superior ao do $IC_{95\%}(\mu)$, por ser uma característica dos intervalos de confiança. Mantendo tudo constante, ao se aumentar o nível de confiança, também se aumenta a amplitude do intervalo.

Q3. Assuma que $X_1, X_2, \cdots, X_n$ denotam cópias *iid's* de uma variável aleatória X com distribuição Poisson(θ). Assuma que uma amostra aleatória de tamanho 200 tenha fornecido uma estimativa de 3,4 para θ. Pede-se para construir um intervalo de confiança assintótico de 96% para θ.

Resposta: Para resolução dessa questão utilizaremos a Definição 6, que diz que se $\hat{\theta}_n = \hat{\theta}(X_1, ..., X_n)$ é um estimador pontual não-viesado de θ, e que $I_n(\theta) = nI_1(\theta) < \infty$, então, pelo TCL,

$$\hat{\theta}_n \sim N(\theta, I_n^{-1}(\theta)),\ \text{quando}\, n \to \infty.$$

Desse modo, a primeira etapa para encontrar a Informação de Fisher é a definição da verossimilhança. Para $X_i \sim$ Poisson(θ), a função de probabilidade é:

$$P(X_i = x_i) = \frac{e^{-\theta}\theta^{x_i}}{x_i!}$$

A função de verossimilhança para n observações independentes é:

$$L(\theta) = \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!} = e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} \frac{1}{x_i!}$$

A log-verossimilhança é:

$$\ell(\theta) = \log L(\theta) = -n\theta + \left(\sum_{i=1}^{n} x_i\right)\log\theta - \sum_{i=1}^{n}\log(x_i!)$$

Derivando em relação a θ e igualando a zero:

$$\frac{d}{d\theta}\ell(\theta) = -n + \frac{1}{\theta}\sum_{i=1}^{n} x_i = 0 \quad \Rightarrow \quad \hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{X}$$

Assim, o EMV de θ é $\hat{\theta} = \bar{X}$.

Sendo assim, pela Definição 6, a variância assintótica do estimador $\hat{\theta}_n$ é:

$$\mathrm{Var}(\hat{\theta}_n) = I_n^{-1}(\theta)$$

Para a distribuição de Poisson, a informação de Fisher para 1 observação é:

$$I_1(\theta) = \mathbb{E}\left[-\frac{d^2}{d\theta^2}\log f(X;\theta)\right] = \mathbb{E}\left[-\left(-\frac{X}{\theta^2}\right)\right] = \frac{\mathbb{E}[X]}{\theta^2} = \frac{\theta}{\theta^2} = \frac{1}{\theta}$$

Logo, para n observações:

$$I_n(\theta) = n \cdot I_1(\theta) = \frac{n}{\theta} \quad \Rightarrow \quad I_n^{-1}(\theta) = \frac{\theta}{n}$$

Como $\hat{\theta} \overset{a}{\sim} \mathcal{N}\left(\theta, \frac{\theta}{n}\right)$, substituímos θ por $\hat{\theta}$ na variância para obter um intervalo de confiança assintótico. Desse modo, considerando que é pedido um intervalo de confiança de 96%, então $\alpha = 0,04$:

$$IC_{0{,}96}(\theta) = \hat{\theta} \pm z_{1-\frac{0{,}04}{2}} \cdot \sqrt{\frac{\hat{\theta}}{n}} = \hat{\theta} \pm z_{0{,}98} \cdot \sqrt{\frac{\hat{\theta}}{n}}$$

,

sendo $z_{0{,}98} \approx 2,0537$. Substituindo os valores:

$$IC_{0{,}96}(\theta) = 3,4 \pm z_{0{,}98} \cdot \sqrt{\frac{3,4}{200}} \approx 3,4 \pm 2,0537 \cdot 0,1304 \approx 3,4 \pm 0,2676 \Rightarrow (3,13;\ 3,67)$$

Portanto, o intervalo de confiança assintótico de 96% para θ é:

$$IC_{0{,}96}(\theta) = (3,132;\ 3,668)$$

Q4. Com respeito às últimas eleições municipais foi efetuada uma sondagem sobre as intenções de voto dos eleitores de uma certa cidade Brasileira, recolhendo-se 500 respostas. Nestas, 200 estabeleciam a intenção de votar no atual Prefeito, 150 pronunciavam-se favoráveis ao candidato do outro partido e as restantes tem preferência por outros candidatos ou não expressam a sua preferencia. O estudo destina-se a avaliar as hipóteses de exito do atual Prefeito. Pergunta-se:

(a) Qual a população em causa? Justifique a escolha especificando quais os parâmetros da distribuição.

Resposta: A população em questão corresponde a todos os eleitores da cidade considerada na pesquisa eleitoral. A amostra, composta por $n = 500$ eleitores, tem por objetivo inferir características da população total, no caso o êxito do atual prefeito, sendo medida a proporção de eleitores que votarão nele.

O parâmetro de interesse é a proporção populacional p dos eleitores que votarão no atual prefeito. Supondo amostragem aleatória, a variável aleatória X que conta o número de eleitores que votam no atual prefeito segue a distribuição:

$$X \sim \text{Binomial}(n = 500, p).$$

(b) Indique, justificando, qual o melhor estimador para a proporção de eleitores da cidade em questão, que não votam no atual Prefeito. Com base na amostra recolhida indique uma estimativa para essa proporção.

Resposta: O melhor estimador pontual para a proporção populacional p é a proporção amostral:

$$\hat{p} = \frac{x}{n}$$

Onde x é o número de eleitores da amostra que não votarão no atual prefeito.

Pelos dados:

- 200 votam no atual prefeito;
- 150 votam no candidato do outro partido;
- 150 têm outra preferência ou não responderam.

Logo, o total que não vota no atual prefeito é $x = 150 + 150 = 300$.

Portanto:

$$\hat{p} = \frac{300}{500} = 0,60$$

Assim, a estimativa pontual para a proporção de eleitores que não votarão no atual prefeito é de 60%.

(c) Construa um intervalo de 90% para a proporção de eleitores que preferem votar no candidato do outro partido. Interprete o resultado.

Resposta: Neste item, o interesse é na proporção de eleitores que preferem o candidato do outro partido, ou seja:

$$\hat{p}_{\text{outro}} = \frac{150}{500} = 0,30$$

Pelo TLC, dado que o tamanho amostral é suficientemente grande ($n > 30$), o intervalo de confiança de 90% para a proporção p é dado por:

$$IC_{0,90}(p) = \hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Com:

- $\hat{p} = 0,30$
- $n = 500$
- $z_{\alpha/2} \approx 1,645$ (nível de 90%)

Calculando o erro padrão:

$$EP = \sqrt{\frac{0,30 \cdot (1-0,30)}{500}} = \sqrt{\frac{0,21}{500}} \approx 0,02049$$

Calculando o intervalo:

$$IC_{0,90}(p) = 0,30 \pm 1,645 \cdot 0,02049 \approx 0,30 \pm 0,0337$$

$$IC_{0,90}(p) \approx (0,266,\ 0,334).$$

Portanto, com 90% de confiança, estima-se que a proporção de todos os eleitores da cidade que preferem votar no candidato do outro partido está entre 26,63% e 33,37%. Isso significa que, se muitas amostras semelhantes fossem coletadas, aproximadamente 90% dos intervalos construídos conteriam a proporção real da população. Ou seja, como o intervalo de confiança da proporção de indivíduos que não votariam no candidato atual não contém 0,5; significa que é provável que o candidato atual tenha êxito na sua candidatura.

Q5. Uma farmacêutica pretende avaliar a eficácia de duas marcas de indutores do sono, A e B que detêm no mercado. Os ensaios clínicos efetuados em 10 doentes permitiram obter os tempos (em minutos) que os mesmos doentes levaram a adormecer com cada um dos indutores.

Paciente	1	2	3	4	5	6	7	8	9	10
Indutor A	5	6	5	5	4	3	5	5	6	6
Indutor B	2	4	5	4	4	4	3	4	4	6

Construa um intervalo de confiança a 95% para a diferença entre os tempos médios que os doentes levam a adormecer para os dois indutores do sono, supondo que, os dados foram extraídos de duas distribuições normal com médias μ_1 e μ_2 (desconhecidas):

(a) Com $\sigma_1^2 \neq \sigma_2^2$, sendo ambas desconhecidas.

Resposta: Com base no enunciado, temos que, $n_A = n_B = 10$, as duas amostras apresentadas forneceram as seguintes estimativas para a média, e variância: $\bar{x}_A = \sum_{i=1}^{10} x_i/10 = 5$ e $\bar{x}_B = \sum_{i=1}^{10} x_i/10 = 4$. De igual forma, $S_A^2 = \frac{1}{n_A - 1} \sum_{i=1}^{n_A} (x_i - \bar{x}_A)^2 =$

$0,889$ e $S_B^2 = \frac{1}{n_B-1}\sum_{i=1}^{n_B}(x_i - \bar{x}_B)^2 = 1,111$. Portanto, supondo variâncias desconhecidas e diferentes (com $\sigma_1^2 = \sigma_A^2$ e $\sigma_2^2 = \sigma_B^2$), segue que, a quantidade pivotal associada ao estimador $\hat{\Delta}_n = \bar{X}_A - \bar{X}_B$ de $\mu_A - \mu_B$ é dada por:

$$Q(\mathbf{X_A}, \mathbf{X_B}, \Delta) = \frac{\left(\hat{\Delta}_n - \Delta\right)}{\sqrt{Var\left(\hat{\Delta}_n\right)}} = \frac{\left(\bar{X}_A - \bar{X}_B\right) - (\mu_A - \mu_B)}{\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}}} \sim t_v,$$

com $v = \frac{\left(\frac{S_A^2}{n_A}+\frac{S_B^2}{n_B}\right)^2}{\frac{\left(\frac{S_A^2}{n_A}\right)^2}{n_A-1}+\frac{\left(\frac{S_B^2}{n_B}\right)^2}{n_B-1}}$, substituindo os valores na expressão anterior obtemos que, $v = 17,78 \approx 18$ graus de liberdade. Desta forma, para um nível de confiança de 95% obtemos que, $t_{(v=18,\alpha/2=0,025)} = 2,10$. Portanto, os limites do intervalo de confiança para a diferença de médias, Δ são dados por:

$$\mathbb{P}\left(-t_{(v,\alpha/2)} \leq Q(\mathbf{X_A}, \mathbf{X_B}, \Delta)\right) \leq t_{(v,\alpha/2)} = \mathbb{P}\left(-t_{(v,\alpha/2)} \leq \frac{\left(\hat{\Delta}_n - \Delta\right)}{\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}}} \leq t_{(v,\alpha/2)}\right) = 1 - \alpha$$

$$\mathbb{P}\left(\hat{\Delta}_n - t_{(v,\alpha/2)}\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}} \leq \Delta \leq \hat{\Delta}_n + t_{(v,\alpha/2)}\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}}\right) = 1 - \alpha$$

$$\mathbb{P}\left((5-4) - 2,10 \times \sqrt{\frac{0,889}{10} + \frac{1,111}{10}} \leq \Delta \leq (5-4) + 2,10\sqrt{\frac{0,889}{10} + \frac{1,111}{10}}\right).$$

Ou seja, o intervalo de 95% de confiança para Δ é dado por $IC_{0,95}(\Delta) = [0,06; 1,95]$.

(b) Com $\sigma_1^2 = \sigma_2^2 = \sigma^2$, sendo σ^2 desconhecidas.

Resposta: Bom, supondo agora que $\sigma_A^2 = \sigma_B^2 = \sigma^2$, com σ^2 (desconhecida), segue da abordagem dada em sala de aulas que,

$$Var\left(\hat{\Delta}_n\right) = \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B} = \frac{\sigma^2}{n_A} + \frac{\sigma^2}{n_B} = \sigma^2\left(\frac{1}{n_A} + \frac{1}{n_B}\right).$$

Portanto, um estimador para σ^2 é, $\hat{\sigma}^2 = S_p^2 = \frac{(n_A-1)S_A^2+(n_B-1)S_B^2}{(n_A-1)+(n_B-1)} = \frac{9\times 0,889+9\times 1,111}{9+9} \approx 1$. Desta forma, a quantidade pivotal será dada por:

$$Q(\mathbf{X_A}, \mathbf{X_B}, \Delta) = \frac{\left(\hat{\Delta}_n - \Delta\right)}{\sqrt{Var\left(\hat{\Delta}_n\right)}} = \frac{\left(\bar{X}_A - \bar{X}_B\right) - (\mu_A - \mu_B)}{S_p\sqrt{\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}} \sim t_{(n_A+n_B-2)},$$

onde, da tabela de distribuição t-Student obtemos o seguinte valor crítico, $t_{(18;0,025)} = 2,10$. Portanto, o intervalo de confiança para Δ será dado por:

$$\begin{aligned} IC_{0,95}(\Delta) &= \left[\hat{\delta}_n - 2,10 \times 1 \times \sqrt{1/10 + 1/10}; \hat{\delta}_n + 2,10 \times 1 \times \sqrt{1/10 + 1/10}\right] \\ &= \left[1 - 2,10 \times 1 \times \sqrt{1/10 + 1/10}; 1 + 2,10 \times 1 \times \sqrt{1/10 + 1/10}\right] \\ &= [0,061; 1,939]. \end{aligned}$$

Onde $\hat{\delta}_n = \bar{x}_A - \bar{x}_B = 5 - 4 = 1$ denota a estimativa pontual para a diferença de médias.

(c) Com $\sigma_1^2 = \sigma_2^2 = \sigma^2$, sendo σ^2 conhecidas.

Resposta: Supondo agora que $\sigma_A^2 = \sigma_B^2 = \sigma^2$, com σ^2 (conhecida), ou seja, $\sigma^2 = \sigma_0^2$. Desse modo, segue da abordagem dada em sala de aulas que,

$$Var\left(\hat{\Delta}_n\right) = \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B} = \frac{\sigma_0^2}{n_A} + \frac{\sigma_0^2}{n_B} = \sigma_0^2\left(\frac{1}{n_A} + \frac{1}{n_B}\right).$$

Portanto, um estimador para σ^2 é,

$$Q\left(\mathbf{X_A}, \mathbf{X_B}, \Delta\right) = \frac{\left(\hat{\Delta}_n - \Delta\right)}{\sqrt{Var\left(\hat{\Delta}_n\right)}} = \frac{\left(\bar{X}_A - \bar{X}_B\right) - (\mu_A - \mu_B)}{\sigma_0\sqrt{\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}} \sim \mathcal{N}(0,1),$$

Portanto, o intervalo de confiança é dado por:

$$\begin{aligned} IC_{0,95}(\Delta) &= \left[\hat{\delta}_n - z_{1-\frac{0,05}{2}} \times \sigma_0 \times \sqrt{1/10 + 1/10}; \hat{\delta}_n + z_{1-\frac{0,05}{2}} \times \sigma_0 \times \sqrt{1/10 + 1/10}\right] \\ &= \left[1 - 1,96 \times \sigma_0 \times \sqrt{1/10 + 1/10}; 1 + 1,96 \times \sigma_0 \times \sqrt{1/10 + 1/10}\right] \\ &= \left[1 - 0,876\sigma_0; 1 + 0,876\sigma_0\right]. \end{aligned}$$

Onde $\hat{\delta}_n = \bar{x}_A - \bar{x}_B = 5 - 4 = 1$ denota a estimativa pontual para a diferença de médias.

(d) Com base nos resultados dos itens anteriores, explique qual dos indutores é o mais eficaz.

Resposta: Observe que, dos resultados anteriores, nossa inferência foi baseada na diferença dos tempos médios do indutor A com relação ao indutor B. Portanto, para escolher o indutor mais eficaz, podemos simplesmente olhar para os tempos médios amostrais, onde é possível concluir que o indutor A apresenta maior tempo médio de indução do sono. Já, o indutor B apresenta o menor tempo médio. Consequências dessas alegações podem ser observadas nos intervalos de confiança que foram construídos nos itens anteriores, pois, nos dois casos os intervalos não apresentam valores negativos para as faixas de confiança, suportando uma vez mais a alegação de que o indutor B apresenta maior eficácia. Lembrando quem $\hat{\Delta}_n = \bar{X}_A - \bar{X}_B$.

Q6. Seja $\bar{X}_n$ a média de uma amostra aleatória $X_1, X_2, \cdots, X_n$ proveniente de uma distribuição normal de média μ e variância $\sigma^2 = 9$. Encontre o tamanho de amostra n, tal que, $\mathbb{P}\left(\bar{X}_n - 1 < \mu < \bar{X}_n + 1\right) = 0,90$.

Resposta: Queremos encontrar:

$$\mathbb{P}\left(\bar{X}_n - 1 < \mu < \bar{X}_n + 1\right) = 0,90$$

Que pode ser reescrita como:

$$\mathbb{P}\left(-1 < \bar{X}_n - \mu < 1\right) = 0,90$$

Desse modo, como $X_i \sim \mathcal{N}(\mu, \sigma^2 = 9)$, temos que $\bar{X}_n \sim \mathcal{N}\left(\mu, \frac{9}{n}\right)$.
Seja Z a variável normal padrão:

$$Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\overline{X}_n - \mu}{3/\sqrt{n}}$$

Logo:

$$\mathbb{P}\left(-\frac{1}{3/\sqrt{n}} < Z < \frac{1}{3/\sqrt{n}}\right) = 0,90 \Rightarrow -\frac{\sqrt{n}}{3} < Z < \frac{\sqrt{n}}{3}$$

Utilizando a Tabela da Normal Padrão, sabemos que:

$$\mathbb{P}(|Z| < z) = 0,90 \Rightarrow z = z_{0{,}95} \approx 1,645$$

Igualamos:

$$\frac{\sqrt{n}}{3} = 1,645 \Rightarrow \sqrt{n} = 3 \cdot 1,645 = 4,935 \Rightarrow n \approx (4,935)^2 \approx 24,35$$

Portanto, como n deve ser inteiro:

$$n = 25.$$

Q7. Seja $\bar{x}$ a média observada de uma amostra aleatória de tamanho n proveniente de uma distribuição com média μ desconhecida, e variância σ^2 conhecida. Encontre o tamanho de amostra n tal que, $[\bar{x} - \sigma/4; \bar{x} + \sigma/4]$ seja um intervalo de 95% de confiança para μ.

Resposta: Com base no enunciado, temos que, $\bar{x}$ é a média observada de uma amostra aleatória de tamanho n. Supondo σ^2 conhecido, o intervalo de 95% de confiança para μ é: $IC_{0{,}95}(\mu) = [\bar{x} - \sigma/4; \bar{x} + \sigma/4]$, cuja sua amplitude é dada por: $A_1 = L_s - L_i = \bar{x} + \sigma/4 - \bar{x} + \sigma/4 = \sigma/2$. Entretanto, supondo que o intervalo de confiança apresenta a seguinte notação genérica, $IC_{1-\alpha}(\mu) = [\bar{x}_n - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}; \bar{x}_n + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}]$, sua amplitude é dada por: $A_2 = L_s - L_i = 2z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ (resultado dado em sala de aula).

Assim, igualando essas duas amplitudes, digamos $A_1 = A_2$ obtemos o seguinte resultado, $\sigma/2 = 2z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \iff \left(\sqrt{n}\right)^2 = (4 \times 1,96)^2 \approx 62$, que denota o tamanho de amostra minimo, para obter um intervalo de confiança de $[\bar{x} - \sigma/4; \bar{x} + \sigma/4]$. Ou seja, $n \approx 62$.

Q8. Seja Y uma variável aleatória, cuja sua fdp é dada por: $f(y|\theta) = \frac{2(\theta - y)}{\theta^2}$, com $y \in (0, \theta)$ e zero, caso contrário.

(a) Encontre a função de distribuição da variável aleatória Y.

Resposta: temos que, Y é uma variável aleatória tal que, sua função densidade de probabilidade é dada acima. Assim, a função de distribuição de Y é dada por:

$$F_Y(y) = \int_0^y f(\theta|x)\,dx = \frac{2}{\theta^2}\int_0^y (\theta - x)\,dx = \frac{y(2\theta - y)}{\theta^2}, \text{ com } 0 < y < \theta.$$

(b) Mostre que $\frac{Y}{\theta}$ é uma quantidade pivotal.

Resposta: Seja $Q(Y,\theta) = \frac{Y}{\theta}$, logo sua função de distribuição acumulada é dada por:

$$\begin{aligned} F_Q(q) &= \mathbb{P}(Q \leq q) = \mathbb{P}\left(\frac{Y}{\theta} \leq q\right) = \mathbb{P}(Y \leq q\theta) = \int_0^{\theta q} f_Y(y|\theta)\, dy \\ &= F_Y(\theta q) - F_Y(0) = \frac{\theta q(2\theta - \theta q)}{\theta^2} = q(2-q), \text{ com } 0 < q < 2. \end{aligned}$$

Que não depende de θ. Então, Q é uma quantidade pivotal para θ.

(c) Use a quantidade pivotal do item (b) para obter os limites inferior e superior de θ, supondo $1-\alpha = 0,90$.

Resposta: Segue o fato de que Q é uma quantidade pivotal (resultado do item anterior), e para algum $0 < q_1 < q_2 < 2$ da distribuição de Q segue que,

$$\mathbb{P}(q_1 \leq Q(Y,\theta) \leq q_2) = \mathbb{P}\left(q_1 \leq \frac{Y}{\theta} \leq q_2\right) = \mathbb{P}\left(\frac{1}{q_2} \leq \frac{\theta}{Y} \leq \frac{1}{q_1}\right)$$
$$\mathbb{P}\left(\frac{Y}{q_2} \leq \theta \leq \frac{Y}{q_1}\right) = 1-\alpha.$$

Ou seja, o intervalo de confiança para θ é dado por: $IC_{0,90}(\theta) = \left[\frac{Y}{q_2}; \frac{Y}{q_1}\right]$. Assim, os quantis q_1 e q_2 podem ser obtidos resolvendo as equações:

$$\int_0^{q_1} f_Q(q)\, dq = F_Q(q_1) - F_Q(0) = \alpha/2$$
$$\int_{q_2}^{2} f_Q(q)\, dq = F_Q(2) - F_Q(q_2) = \alpha/2.$$

Portanto, por forma a simplificar as expressões, não vamos detalhar o cálculo dos quantis q_1 e q_2.

Q9. Seja $X_1, X_2, \cdots, X_n$ uma amostra aleatória proveniente da distribuição $\mathcal{N}(\mu,\sigma^2)$, com ambos os parâmetros sendo desconhecidos. O intervalo de confiança para σ^2 pode ser obtido usando a distribuição amostral de S^2. Isto é, $\mathbb{P}\left(\frac{(n-1)S^2}{\sigma^2} < b\right) = 1-\alpha/2$ e $\mathbb{P}\left(a < \frac{(n-1)S^2}{\sigma^2} < b\right) = 1-\alpha$.

(a) Obtenha o intervalo de 95% de confiança para σ^2.

Resposta: Pede-se para construir um intervalo de confiança para σ^2 com base na estatística:

$$Q(\mathbf{X},\sigma^2) = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

Escolhendo os quantis a e b da distribuição Qui-Quadrado com $n-1$ graus de liberdade, temos:

$$\mathbb{P}\left(a \leq \frac{(n-1)S^2}{\sigma^2} \leq b\right) = 1-\alpha$$

Isolando σ^2:

$$\mathbb{P}\left(\frac{(n-1)S^2}{b} \leq \sigma^2 \leq \frac{(n-1)S^2}{a}\right) = 1-\alpha$$

Portanto, o intervalo de confiança de nível $1-\alpha$ para σ^2 é:

$$IC_{1-\alpha}(\sigma^2) = \left[\frac{(n-1)S^2}{b}, \ \frac{(n-1)S^2}{a}\right]$$

com $a = \chi^2_{n-1,\ 1-\alpha/2}$ e $b = \chi^2_{n-1,\ \alpha/2}$.
Portanto,para $1-\alpha = 0,95$, temos:

$$IC_{0{,}95}(\sigma^2) = \left[\frac{(n-1)S^2}{\chi^2_{(n-1,\ 0{,}025)}}, \ \frac{(n-1)S^2}{\chi^2_{(n-1,\ 0{,}975)}}\right].$$

(b) Se $n = 9$ e $s^2 = 7,93$ obtenha o intervalo de 90% de confiança para σ^2.

Resposta: Com os valores $n = 9$, $s^2 = 7,93$, $\alpha = 0,10$, tem-se que os graus de liberdade serão: $df = 8$

Portanto, os quantis serão:

$$a = \chi^2_{8,\ 0{,}95} = 15,507, \quad b = \chi^2_{8,\ 0{,}05} = 2,733$$

Substituindo:

$$IC_{0{,}90}(\sigma^2) = \left[\frac{8 \times 7,93}{15,507}; \ \frac{8 \times 7,93}{2,733}\right] = \left[\frac{63,44}{15,507}; \ \frac{63,44}{2,733}\right].$$

Portanto:

$$IC_{0{,}90}(\sigma^2) \approx [4,09; \ 23,21].$$

(c) Supondo μ conhecido, explique o quão esse fato pode modificar o processo de construção de um intervalo de confiança para σ^2.

Resposta: Quando μ é desconhecida, usamos:

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

Porque se perde 1 grau de liberdade na estimação de μ. Caso μ sej conhecido, então a estatística se torna:

$$\frac{nS^2}{\sigma^2} \sim \chi^2_n$$

Pois não há a necessidade de se estimar μ, já que é conhecido.
Nesse caso, o intervalo de confiança torna-se:

$$IC_{1-\alpha}(\sigma^2) = \left[\frac{nS^2}{(\chi^2_{n,\ \alpha/2})}, \frac{nS^2}{\chi^2_{(n,\ 1-\alpha/2)}}\right]$$

O que resulta em um intervalo mais estreito, pois há menor incerteza na estimação de σ^2.

Q10. Seja $X_{11}, X_{12}, \cdots, X_{1n_1}$ e $X_{21}, X_{22}, \cdots, X_{2n_2}$ duas amostras independentes, provenientes das seguintes distribuições, $X_{1i} \overset{iid}{\sim} \mathcal{N}(\mu_1, \sigma_1^2)$ e $X_{2i} \overset{iid}{\sim} \mathcal{N}(\mu_2, \sigma_2^2)$.

(a) Discuta o processo de obtenção de um intervalo de 92% de confiança $\Delta = \mu_1 - \mu_2$, supondo que $\sigma_1^2 \neq \sigma_2^2$, e que ambos são conhecidos, e com $n_1 = 15$ e $n_2 = 12$.

Resposta: Observe que, $X_{11}, X_{12}, \cdots, X_{1n_1} \sim \mathcal{N}(\mu_1, \sigma_1^2)$ e $X_{21}, X_{22}, \cdots, X_{2n_2} \sim \mathcal{N}(\mu_2, \sigma_2^2)$. Assim, supondo $\sigma_1^2 \neq \sigma_2^2$, sendo ambas conhecidas, vamos denotar por $\bar{X}_1 = \sum_{i=1}^{n_1} X_{1i}/n_1$ e $\bar{X}_2 = \sum_{i=1}^{n_2} X_{2i}/n_2$ as médias amostrais. Desta forma, segue imediatamente que, $\hat{\Delta}_n = \bar{X}_1 - \bar{X}_2$ é o estimador pontual para $\Delta = \mu_1 - \mu_2$. Observe que,

$$Var\left(\hat{\Delta}_n\right) = Var\left(\bar{X}_1 - \bar{X}_2\right) = Var\left(\bar{X}_1\right) + Var\left(\bar{X}_2\right) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

Porque σ_1^2 e σ_2^2 são quantidades conhecidas, segue imediatamente que,

$$Q(\mathbf{X_1}, \mathbf{X_2}, \Delta) = \frac{\hat{\Delta}_n - \Delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim \mathcal{N}(0, 1).$$

Portanto, para um nível de confiança de 92%, segue da tabela de distribuição normal padrão que, $z_{\alpha/2} = 1,75$. Desta forma,

$$\mathbb{P}\left(-z_{\alpha/2} \leq Q(\mathbf{X_1}, \mathbf{X_2}, \Delta) \leq z_{\alpha/2}\right) = \mathbb{P}\left(-z_{\alpha/2} \leq \frac{\hat{\Delta}_n - \Delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \leq z_{\alpha/2}\right) = 1 - \alpha$$

$$\mathbb{P}\left(\hat{\Delta}_n - 1,75\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq \Delta \leq \hat{\Delta}_n + 1,75\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) = 0,92.$$

Ou seja, o intervalo de 92% de confiança para Δ é dado por, $IC_{0,92}(\Delta) = \left[\hat{\Delta}_n - 1,75\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}; \hat{\Delta}_n + 1,75\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$, supondo $\sigma_1^2 \neq \sigma_2^2$ sendo ambas conhecidas.

(b) Assuma agora que, $\sigma_1^2 = 3\sigma_2^2$, com σ_2^2 sendo desconhecidos. Com base nessa informação, obtenha a quantidade pivotal correspondente, e o respectivo intervalo de 95% de confiança para $\Delta = \mu_1 - \mu_2$.

Resposta: Bom, por simplicidade, vamos supor que $\sigma_1^2 = 3\sigma_2^2 = 3\sigma^2$, sendo σ^2 desconhecida, tal como σ_2^2. Desta forma, a variância do estimador $\hat{\Delta}_n$ será dada por:

$$\begin{aligned} Var\left(\hat{\Delta}_n\right) &= Var\left(\bar{X}_1\right) + Var\left(\bar{X}_2\right) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{3\sigma_2^2}{n_1} + \frac{\sigma_2^2}{n_2} \\ &= \frac{3\sigma^2}{n_1} + \frac{\sigma^2}{n_2} = \sigma^2\left(\frac{3}{n_1} + \frac{1}{n_2}\right). \end{aligned}$$

Segue de forma similar a questão 4(b) que, o melhor estimador para σ^2 é a variância ponderada, $S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{(n_1-1)+(n_2-1)}$, com S_1^2 e S_2^2 sendo as variâncias amostrais, cujo as expressões são idênticas com aquelas apresentadas na Questão 4. Desta forma, a quantidade pivotal é dada por,

$$Q\left(\mathbf{X_1}, \mathbf{X_2}, \Delta\right) = \frac{\hat{\Delta}_n - \Delta}{\sqrt{Var\left(\hat{\Delta}_n\right)}} = \frac{\hat{\Delta}_n - \Delta}{S_p\sqrt{\left(\frac{3}{n_1} + \frac{1}{n_2}\right)}} \sim t_{n_1+n_2-2}.$$

Com isso supondo um nível de confiança de 95%, o intervalo de confiança para Δ é dado por:

$$IC_{0,95}\left(\Delta\right) = \left[\hat{\Delta}_n - t_{(n_1+n_2-2;0,025)}S_p\sqrt{\left(\frac{3}{n_1} + \frac{1}{n_2}\right)}; \hat{\Delta}_n + t_{(n_1+n_2-2;0,025)}S_p\sqrt{\left(\frac{3}{n_1} + \frac{1}{n_2}\right)}\right].$$

Portanto, os quantis da distribuição t-Student com n_1+n_2-2 graus de liberdade serão conhecidos, depois que forem especificados os tamanhos de amostra n_1 e n_2, respectivamente.

(c) Assuma que uma amostra de tamanho $n_1 = n_2 = 10$ forneceu as seguintes estimativas, $\bar{x}_1 = 4,80$; $s_1^2 = 8,64$; $\bar{x}_2 = 5,60$ e $s_2^2 = 7,88$. Construa o intervalo de 98% de confiança para o parâmetro $\Delta = \mu_1 - \mu_2$.

Resposta: Bom, com base nos valores fornecidos, vamos nos focar apenas em obter o intervalo de confiança no item (b). Para tal, temos que,

$$S_p^2 = \frac{\left(n_1-1\right)S_1^2 + \left(n_2-1\right)S_2^2}{\left(n_1-1\right)+\left(n_2-1\right)} = \frac{\left(10-1\right)8,64 + \left(10-1\right)7,88}{\left(10-1\right)+\left(10-1\right)} \approx 8,26.$$

Desta forma, segue da tabela de distribuição t-Student que, $t_{(18,0,001)} = 2,55$, com os graus de liberdade dados por: $n_1 + n_2 - 2 = 9 + 9 = 18$.

$$\begin{aligned}
L_i &= \hat{\delta}_n - t_{(18;0,025)}S_p\sqrt{\left(\frac{3}{n_1} + \frac{1}{n_2}\right)} = (4,80-5,60) - 2,55 \times \sqrt{8,26\left(\frac{3}{10} + \frac{1}{10}\right)} = -5,44.\\
L_s &= \hat{\delta}_n + t_{(18;0,025)}S_p\sqrt{\left(\frac{3}{n_1} + \frac{1}{n_2}\right)} = (4,80-5,60) + 2,55 \times \sqrt{8,26\left(\frac{3}{10} + \frac{1}{10}\right)} = 3,84.
\end{aligned}$$

Assim, concluímos que $IC_{0,98}\left(\Delta\right) = [L_i; L_s] = [-5,44; 3,84]$.

(d) Supondo que $\sigma_1^2 = \sigma_2^2 = \sigma^2$, com σ^2 sendo uma quantidade desconhecida, construa o intervalo de 94% de confiança para σ^2.

Resposta: Supondo agora que $\sigma_1^2 = \sigma_2^2 = \sigma^2$, com σ^2 (desconhecida), segue da abordagem dada em sala de aulas que a estimativa para σ^2 é:

$$\hat{\sigma}^2 = S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2.}$$

Além disso, tem-se a relação de que:

$$\frac{(n_1+n_2-2)S_p^2}{\sigma^2} = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{\sigma^2} \sim \chi^2_{n_1+n_2-2}$$

Portanto, tem-se o intervalo:

$$
\begin{aligned}
1-\alpha &= \mathbb{P}\left(\chi^2_{(\alpha/2,\, n_1+n_2-2)} < \frac{(n_1+n_2-2)S_p^2}{\sigma^2} < \chi^2_{(1-\alpha/2,\, n_1+n_2-2)}\right)\\
&= \mathbb{P}\left(\frac{\chi^2_{(\alpha/2,\, n_1+n_2-2)}}{(n_1+n_2-2)S_p^2} < \frac{1}{\sigma^2} < \frac{\chi^2_{(1-\alpha/2,\, n_1+n_2-2)}}{(n_1+n_2-2)S_p^2}\right)\\
&= \mathbb{P}\left(\frac{(n_1+n_2-2)S_p^2}{\chi^2_{(1-\alpha/2,\, n_1+n_2-2)}} < \sigma^2 < \frac{(n_1+n_2-2)S_p^2}{\chi^2_{(\alpha/2,\, n_1+n_2-2)}}\right)\\
&= \mathbb{P}\left(\frac{18\times S_p^2}{\chi^2_{(1-0,06/2;\ 18)}} < \sigma^2 < \frac{18\times S_p^2}{\chi^2_{(0,06/2;\ 18)}}\right)\\
&= \mathbb{P}\left(\frac{18\times 8,26}{30,845} < \sigma^2 < \frac{18\times 8,26}{8,512}\right)\\
&= \mathbb{P}\left(4,82 < \sigma^2 <\ 17,47\right).
\end{aligned}
$$

Portanto, tem-se que o intervalo de confiança de 94% para σ é:

$$IC_{0,94}(\sigma) = (4,82;\ 17,47).$$

Q11. Seja $X_1, \cdots, X_{100}$ e $Y_1, \cdots, Y_{100}$ duas amostras aleatórias independentes, cada uma com distribuição Binomial de parâmetros θ_1 e θ_2, respectivamente. Assuma que nas duas populações tenham sido observados o número de sucessos, $S_x = 50$ e $S_y = 40$. Construa o intervalo de confiança aproximada de 90% de confiança para $\theta_1 - \theta_2$.

Resposta: As estimativas pontuais das proporções amostrais são dadas por:

$$\hat{\theta}_{1n} = \frac{S_X}{n_1} = \frac{50}{100} = 0,5, \qquad \hat{\theta}_{2n} = \frac{S_Y}{n_2} = \frac{40}{100} = 0,4$$

Desse modo, as variâncias se dão por:

$$\text{Var}(\hat{\theta}_{1n}) = \frac{\hat{\theta}_{1n}(1-\hat{\theta}_{1n})}{n} = \frac{0,5(1-0,5)}{100} = \frac{0,25}{100} = 0,0025$$

$$\text{Var}(\hat{\theta}_{2n}) = \frac{\hat{\theta}_{2n}(1-\hat{\theta}_{2n})}{n} = \frac{0,4(1-0,4)}{100} = \frac{0,24}{100} = 0,0024$$

Substituindo para calcular a variância de $\hat{\Delta}_n = \hat{\theta}_1 - \hat{\theta}_2$:

$$\text{Var}(\hat{\theta}_1 - \hat{\theta}_2) = \text{Var}(\hat{\Delta}_n) \approx 0,0025 + 0,0024 = 0,0049$$

Logo, o erro padrão é:

$$EP = \sqrt{0,0049} \approx 0,07$$

Para um nível de confiança de 90%, tem-se $\alpha = 0,10$, e o valor crítico correspondente é:

$$z_{1-\alpha/2} = z_{0,95} \approx 1,645$$

Aplicando a fórmula geral do intervalo assintótico (TCL) para a diferença de proporções:

$$\begin{aligned} IC_{0,9}(\Delta) &= \hat{\Delta}_n \pm z_{1-\alpha/2} \times EP \\ &= 0,1 \pm 1,645 \cdot 0,07 \\ &= 0,1 \pm 0,1152 \end{aligned}$$

Portanto, o intervalo de confiança de 90% para Δ é:

$$IC_{0,9}(\Delta) = (-0,015;\ 0,215).$$

Q12. Seja $\bar{X}_n$ e $\bar{Y}_n$ as médias de duas amostras aleatórias independentes de mesmo tamanho n, provenientes das distribuições $\mathcal{N}(\mu_1, \sigma^2)$ e $\mathcal{N}(\mu_2, \sigma^2)$, respectivamente. Ou seja, pressupõem-se que as duas populações têm a mesma variância σ^2 conhecida. Pede-se para encontrar o valor de n, tal que

$$\mathbb{P}\left(\bar{X}_n - \bar{Y}_n - \frac{\sigma}{5} < \mu_1 - \mu_2 < \bar{X}_n - \bar{Y}_n + \frac{\sigma}{5}\right) = 0,90.$$

Resposta:

Pelo enunciado temos que $\overline{X}_n \sim \mathcal{N}(\mu_1, \frac{\sigma^2}{n})$ e $\overline{Y}_n \sim \mathcal{N}(\mu_2, \frac{\sigma^2}{n})$, e como $\sigma_1^2 = \sigma_2^2 = \sigma^2$ conhecido, então:

$$\overline{X}_n - \overline{Y}_n \sim \mathcal{N}\left(\mu_1 - \mu_2, \sigma^2\left(\frac{1}{n} + \frac{1}{n}\right)\right) = \mathcal{N}\left(\mu_1 - \mu_2, \frac{2\sigma^2}{n}\right)$$

Desse modo, queremos:

$$\mathbb{P}\left(\bar{X}_n - \bar{Y}_n - \frac{\sigma}{5} < \mu_1 - \mu_2 < \bar{X}_n - \bar{Y}_n + \frac{\sigma}{5}\right) = 0,90.$$

Ou seja,

$$\mathbb{P}\left(-\frac{\sigma}{5} < (\bar{X}_n - \bar{Y}_n) - (\mu_1 - \mu_2) < \frac{\sigma}{5}\right) = 0,90.$$

Definindo a variável padronizada:

$$Z = \frac{(\overline{X}_n - \overline{Y}_n) - (\mu_1 - \mu_2)}{\sqrt{2\sigma^2/n}} \sim \mathcal{N}(0,1)$$

Logo, tem-se que:

$$\mathbb{P}\left(\frac{-\frac{\sigma}{5}}{\sqrt{2\sigma^2/n}} < Z < \frac{\frac{\sigma}{5}}{\sqrt{2\sigma^2/n}}\right) = 0,90.$$

Manipulando de forma algébrica tem-se que:

$$\frac{\frac{\sigma}{5}}{\sqrt{2\sigma^2/n}} = \frac{1}{5} \cdot \frac{1}{\sqrt{2/n}} = \frac{1}{5} \cdot \sqrt{n/2} = \frac{\sqrt{n}}{5\sqrt{2}}$$

Portanto:

$$\mathbb{P}\left(\frac{-\sqrt{n}}{5\sqrt{2}} < Z < \frac{\sqrt{n}}{5\sqrt{2}}\right) = 0,90$$

Utilizando a Tabela da Normal Padrão, sabemos que $\mathbb{P}(|Z| < z) = 0,90 \Rightarrow z = z_{0,95} \approx 1,645$

Igualando:

$$\frac{\sqrt{n}}{5\sqrt{2}} = 1,645$$

Desse modo teremos que:

$$\sqrt{n} = 1,645 \times 5\sqrt{2} = 1,645 \times 5 \times 1,4142 \approx 11,632$$

Portanto:

$$n \approx (11,632)^2 \approx 135,29$$

Como n deve ser inteiro, então: $n = 136$.

Observação: a segunda forma seria encontrar a amplitude do $IC_{0,9}(\Delta)$ que seria tal que, $AIC - L_s - L_i = 2\sigma/5$ e de seguida calcular a margem de erro como $E = AIC/2 = \sigma/5$ (resultado dado em sala de aula). Por fim, é do conhecimento de todos que $E = z_{\alpha/2}\sigma_{\Delta_n}$, onde $\sigma_{\Delta_n} = \sigma\sqrt{2/n}$ (segue do resultado apresentando anteriormente). Igualando as duas expressões de E e resolvendo como função de n obtemos o valor fornecido anteriormente.

Q13. Seja $X_1, X_2, \ldots, X_n$ e $Y_1, Y_2, \ldots, Y_m$ duas amostras independentes, tais que, $X_i \overset{iid}{\sim} \mathcal{N}(\mu_1, \sigma_1^2)$ e $Y_i \overset{iid}{\sim} \mathcal{N}(\mu_2, \sigma_2^2)$, com os 4 parâmetros sendo desconhecidos. Pede-se para:

(a) Explicar o tipo de distribuição associada à variável aleatória R.

Resposta: Sendo $R = \frac{\sigma_1^2}{\sigma_2^2}$, tem-se os estimadores de variância amostral:

$$S_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2, \quad S_2^2 = \frac{1}{m-1}\sum_{j=1}^{m}(Y_j - \bar{Y})^2$$

Sabe-se que, sob a suposição de normalidade:

$$\frac{(n-1)S_1^2}{\sigma_1^2} \sim \chi^2_{n-1}, \quad \frac{(m-1)S_2^2}{\sigma_2^2} \sim \chi^2_{m-1}$$

Além disso, essas duas variáveis qui-quadrado são independentes. Assim, a estatística:

$$R = \frac{\frac{\chi^2_{n-1}}{(n-1)}}{\frac{\chi^2_{m-1}}{(m-1)}} = \frac{\frac{\frac{(n-1)S_1^2}{\sigma_1^2}}{(n-1)}}{\frac{\frac{(m-1)S_2^2}{\sigma_2^2}}{(m-1)}} = \frac{\left(\frac{S_1^2}{\sigma_1^2}\right)}{\left(\frac{S_2^2}{\sigma_2^2}\right)} = \frac{S_1^2}{S_2^2}\frac{\sigma_2^2}{\sigma_1^2} \sim F_{(n-1,m-1)}$$

Portanto, a estatística amostral $\frac{S_1^2}{S_2^2}$ está relacionada a $R = \frac{\sigma_1^2}{\sigma_2^2}$ pela distribuição F:

$$\frac{S_1^2}{S_2^2} \sim R \cdot F_{(n-1,m-1)}$$

(b) Obter, usando a tabela apropriada, o valor dos quantis q_1 e q_2 tais que, $\mathbb{P}(R < q_2) = 0,975$ e $\mathbb{P}(q_1 < R < q_2) = 0,95$

Resposta: Pede-se para encontrar os quantis q_1 e q_2 tais que:

$$\mathbb{P}(q_1 < R < q_2) = 0{,}95$$

Como tem-se que $\frac{S_1^2}{S_2^2}\frac{\sigma_2^2}{\sigma_1^2} \sim F_{(n-1,m-1)}$, então:

$$\mathbb{P}\left(f_{(1-\frac{\alpha}{2})(n-1)(m-1)} < \frac{S_1^2}{S_2^2}\frac{\sigma_2^2}{\sigma_1^2} < f_{(\frac{\alpha}{2},n-1,m-1)}\right) = 1-\alpha$$

$$\Rightarrow \mathbb{P}\left(\frac{S_2^2 f_{(1-\frac{\alpha}{2},n-1,m-1)}}{S_1^2} < \frac{\sigma_2^2}{\sigma_1^2} < \frac{S_2^2 f_{(\frac{\alpha}{2},n-1,m-1)}}{S_1^2}\right) = 1-\alpha$$

$$\Rightarrow \mathbb{P}\left(\frac{S_1^2}{S_2^2 f_{(\frac{\alpha}{2},n-1,m-1)}} < R < \frac{S_1^2}{S_2^2 f_{(1-\frac{\alpha}{2},n-1,m-1)}}\right) = 1-\alpha.$$

Logo, os quantis são dados por:

$$q_1 = \frac{S_1^2}{S_2^2 f_{(1-\frac{\alpha}{2},n-1,m-1)}}, \quad q_2 = \frac{S_1^2}{S_2^2 f_{(\frac{\alpha}{2},n-1,m-1)}}.$$

Observe que, a obtenção de q_1 e q_2 requer especificar os valores de n e m respectivamente.

(c) Construir o intervalo de $(1-\alpha)100\%$ de confiança para a razão de variâncias $R = \sigma_1^2/\sigma_2^2$, supondo que $\frac{(n-1)S_1^2}{\sigma_1^2} \sim \chi^2_{n-1}$ e $\frac{(m-1)S_2^2}{\sigma_2^2} \sim \chi^2_{m-1}$. Sendo S_1^2 e S_2^2 os estimadores pontuais para σ_1^2 e σ_2^2, respectivamente.

Resposta: A partir dos resultados anteriores, o intervalo de confiança para R com nível de confiança $1-\alpha$ é:

$$IC_{(1-\alpha)100\%\left(\frac{\sigma_1^2}{\sigma_2^2}\right)} = \left[\frac{S_1^2}{S_2^2} \times \frac{1}{f_{(\frac{\alpha}{2},n-1,m-1)}}; \quad \frac{S_1^2}{S_2^2} \times \frac{1}{f_{(1-\frac{\alpha}{2},n-1,m-1)}}\right].$$