Universidade Federal da Paraíba

Bacharelado em Estatística

Pesquisa Aplicada à Estatística – Lista Completa de Exercícios

Duas Populações e Testes Especiais

Aluno: ____________________ **Matrícula:** ____________________

Questão 1

(Retirado de: "Lista Atividade 2 - Intervalo de Confiança para duas populações", Exercício número 1)

Uma fábrica deseja comparar a durabilidade (em horas) de dois tipos de lâmpadas. Sabe-se, por especificações técnicas, que os desvios padrão populacionais são $\sigma_A = 50$ h (tipo A) e $\sigma_B = 60$ h (tipo B). Foram testadas amostras aleatórias de 40 lâmpadas do tipo A e 35 do tipo B, obtendo-se médias amostrais $\bar{x}_A = 1200$ h e $\bar{x}_B = 1180$ h.

Resolução:
Extração de Dados:
$\sigma_A = 50$, $\sigma_B = 60$, $n_A = 40$, $n_B = 35$, $\bar{x}_A = 1200$, $\bar{x}_B = 1180$, $1 - \alpha = 0,95$.

(a) Construa um intervalo de confiança de 95% para a diferença entre as durabilidades médias populacionais $\mu_A - \mu_B$.

1. Valor Crítico: Como os desvios padrão populacionais são conhecidos, usamos a distribuição Normal Padrão (Z).

$$Z_{\alpha/2} = \boxed{1,96}$$

2. Diferença Pontual:

$$\bar{x}_A - \bar{x}_B = 1200 - 1180 = \boxed{20}$$

3. Erro Padrão (EP):

$$EP = \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}} = \sqrt{\frac{50^2}{40} + \frac{60^2}{35}} = \sqrt{62,5 + 102,8571} \approx \boxed{12,859}$$

4. Margem de Erro (E):

$$E = Z_{\alpha/2} \times EP = 1,96 \times 12,859 \approx \boxed{25,20}$$

5. Construção e Interpretação:

$$IC(\mu_A - \mu_B; 95\%) = [20 - 25,20\,;\, 20 + 25,20] = \boxed{[-5,20\,;\, 45,20]}$$

(b) Interprete o resultado obtido. É possível afirmar, com 95% de confiança, que existe diferença significativa entre os dois tipos de lâmpadas?
Interpretação: Estima-se, com 95% de confiança, que a verdadeira diferença entre as durabilidades médias esteja no intervalo de -5,20 horas a 45,20 horas. Como o intervalo construído **contém o valor 0**, não existe evidência estatística de diferença significativa entre as lâmpadas A e B.

(c) Se o nível de confiança fosse aumentado para 99%, o que aconteceria com a amplitude do intervalo? (Justifique sem recalcular.)
Aumentando o nível de confiança para 99%, o multiplicador crítico da distribuição normal ($Z_{\alpha/2}$) aumentará (de 1,96 para 2,576). Na fórmula Margem de Erro $E = Z_{\alpha/2} \times EP$, isso provoca um aumento direto de E, resultando em um **aumento da amplitude** do intervalo de confiança. Maior grau de certeza exige intervalos mais largos.

Questão 2

(Retirado de: "Lista Atividade 2 - Intervalo de Confiança para duas populações", Exercício número 2)

Um pesquisador quer avaliar o efeito de um novo método de ensino em matemática. Dois grupos independentes de alunos foram formados: 12 alunos no método tradicional (grupo 1) e 10 alunos no novo método (grupo 2). As notas em um teste padronizado apresentaram: $\bar{x}_1 = 72$, $s_1 = 8$, $\bar{x}_2 = 78$, $s_2 = 7$. Supondo que as variâncias populacionais sejam iguais e que as notas sigam distribuição normal:

Resolução:
Extração de Dados:
$n_1 = 12$, $\bar{x}_1 = 72$, $s_1 = 8$, $n_2 = 10$, $\bar{x}_2 = 78$, $s_2 = 7$, $1 - \alpha = 0,95$, com $\sigma_1^2 = \sigma_2^2$.

(a) Calcule a variância combinada S_p^2.
Como as variâncias populacionais são assumidas iguais, calculamos a variância combinada ponderando as variâncias amostrais pelos respectivos graus de liberdade:

$$S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$
$$S_p^2 = \frac{(12 - 1)8^2 + (10 - 1)7^2}{12 + 10 - 2}$$
$$S_p^2 = \frac{11 \times 64 + 9 \times 49}{20}$$
$$S_p^2 = \frac{704 + 441}{20} = \frac{1145}{20} = \boxed{57,25}$$

(b) Construa um intervalo de confiança de 95% para a diferença $\mu_2 - \mu_1$.
1. Valor Crítico: Como as variâncias populacionais são desconhecidas mas iguais, usamos a distribuição t de Student com $gl = n_1 + n_2 - 2 = 20$. Para $1 - \alpha = 0,95$, temos

$\alpha/2 = 0,025$:

$$t_{\alpha/2} = \boxed{2,086}$$

2. Diferença Pontual: Prestando atenção na ordem requerida $\mu_2 - \mu_1$:

$$\bar{x}_2 - \bar{x}_1 = 78 - 72 = \boxed{6}$$

3. Erro Padrão (EP):

$$EP = \sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$
$$EP = \sqrt{57,25\left(\frac{1}{12} + \frac{1}{10}\right)} = \sqrt{57,25\left(\frac{11}{60}\right)} = \sqrt{10,4958} \approx \boxed{3,240}$$

4. Margem de Erro (E):

$$E = t_{\alpha/2} \times EP = 2,086 \times 3,240 \approx \boxed{6,758}$$

5. Construção e Interpretação:

$$IC(\mu_2 - \mu_1; 95\%) = [6 - 6,758\,;\,6 + 6,758] = \boxed{[-0,758\,;\,12,758]}$$

(c) Com base no intervalo, é possível concluir que o novo método é mais eficaz? Justifique.
Interpretação: Como o intervalo de confiança $[-0,758\,;\,12,758]$ **contém o valor 0**, não há evidências estatísticas para afirmar, com 95% de confiança, que $\mu_2 > \mu_1$. Portanto, não se pode concluir estatisticamente que o novo método é mais eficaz do que o método tradicional.

Questão 3

(Retirado de: "Lista Atividade 2 - Intervalo de Confiança para duas populações", Exercício número 3)

Um estudo comparou o tempo de reação (em ms) de dois grupos de motoristas: um grupo que dormiu 6 horas (grupo 1, $n_1 = 15$) e outro que dormiu 8 horas (grupo 2, $n_2 = 12$). Os dados resumidos são: $\bar{x}_1 = 320$, $s_1 = 25$, $\bar{x}_2 = 300$, $s_2 = 18$. Admita que as populações são normais.

Resolução:
Extração de Dados:
$n_1 = 15$, $\bar{x}_1 = 320$, $s_1 = 25$
$n_2 = 12$, $\bar{x}_2 = 300$, $s_2 = 18$
$1 - \alpha = 0,95$, variâncias populacionais desconhecidas e assumidas diferentes devido ao item (b).

(a) Construa um intervalo de confiança de 95% para $\mu_1 - \mu_2$ e Interprete o resultado.
1. Valor Crítico: Utilizamos a aproximação de Welch para corrigir os graus de liberdade (ν), pois tratamos do problema de Behrens-Fisher (variâncias diferentes):

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}} = \frac{\left(\frac{625}{15} + \frac{324}{12}\right)^2}{\frac{(625/15)^2}{14} + \frac{(324/12)^2}{11}}$$
$$\nu = \frac{(41,667 + 27)^2}{\frac{1736,111}{14} + \frac{729}{11}} = \frac{4715,111}{124,008 + 66,273} = \frac{4715,111}{190,281} \approx 24,78$$

Arredondando ν conservadoramente para $\nu = 24$. Usando a distribuição t de Student para $gl = 24$ e $\alpha/2 = 0,025$:

$$t_{\alpha/2} = \boxed{2,064}$$

2. Diferença Pontual:

$$\bar{x}_1 - \bar{x}_2 = 320 - 300 = \boxed{20}$$

3. Erro Padrão (EP):

$$EP = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{625}{15} + \frac{324}{12}} = \sqrt{41,667 + 27} = \sqrt{68,667} \approx \boxed{8,287}$$

4. Margem de Erro (E):

$$E = t_{\alpha/2} \times EP = 2,064 \times 8,287 \approx \boxed{17,104}$$

5. Construção e Interpretação:

$$IC(\mu_1 - \mu_2; 95\%) = [20 - 17,104\,;\, 20 + 17,104] = \boxed{[2,896\,;\, 37,104]}$$

Interpretação: O intervalo de confiança **não contém o valor 0** (ambos os limites são positivos). Assim, estipula-se com 95% de confiança que a média temporal de reação do grupo 1 (dormiu 6h) é estatisticamente e significativamente maior do que a do grupo 2 (dormiu 8h).

(b) Pesquise na internet porque a aproximação utilizada para cálculo dos graus de liberdades é necessária.
A aproximação de Welch-Satterthwaite é necessária quando os desvios padrão populacionais são **desconhecidos e não podem ser considerados iguais** (falta de homocedasticidade). Sob estas condições, não se pode agrupar e calcular uma única variância combinada (S_p^2), resultando no problema de Behrens-Fisher, onde a diferença observada não segue uma t de Student perfeitamente padronizada. A fórmula aproxima graus de liberdade equivalentes baseados nas proporções e pesos de cada variância amostral, achatando artificialmente a distribuição (graus com finais não inteiros abaixados de forma conservadora). Seu uso garante que o intervalo final alcance, independentemente do erro,

a respectiva margem técnica de confiança especificada, protegendo as inferências contra conclusões falsas caso as dispersões sejam heterogêneas e severamente divergentes.

Questão 4

(Retirado de: "Lista Atividade 2 - Intervalo de Confiança para duas populações", Exercício número 4)

Um nutricionista aplicou uma dieta a 10 pacientes e mediu seus pesos (kg) antes e depois do programa. Os dados são:

Paciente	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Antes	85	92	78	88	95	82	90	79	84	91
Depois	82	88	75	84	91	78	86	76	81	87

Resolução:
Extração de Dados:
$d_i = \text{Depois} - \text{Antes}$, $n = 10$ pares amostrais, $1 - \alpha = 0,95$.

(a) Calcule as diferenças $d_i = \text{Depois} - \text{Antes}$ e obtenha a média $\bar{d}$ e o desvio padrão s_d dessas diferenças.
As diferenças par a par (d_i) para os 10 pacientes são:

$$\begin{aligned} d_1 &= 82 - 85 = -3 & \quad d_6 &= 78 - 82 = -4 \\ d_2 &= 88 - 92 = -4 & \quad d_7 &= 86 - 90 = -4 \\ d_3 &= 75 - 78 = -3 & \quad d_8 &= 76 - 79 = -3 \\ d_4 &= 84 - 88 = -4 & \quad d_9 &= 81 - 84 = -3 \\ d_5 &= 91 - 95 = -4 & \quad d_{10} &= 87 - 91 = -4 \end{aligned}$$

A soma das diferenças é $\sum d_i = -36$. Onde a média das diferenças $(\bar{d})$ é:

$$\bar{d} = \frac{\sum d_i}{n} = \frac{-36}{10} = \boxed{-3,6}$$

A soma do quadrado das diferenças é $\sum d_i^2 = 4 \times (-3)^2 + 6 \times (-4)^2 = 36 + 96 = 132$. A variância amostral s_d^2 e o desvio padrão s_d são:

$$s_d^2 = \frac{\sum d_i^2 - \frac{(\sum d_i)^2}{n}}{n-1} = \frac{132 - \frac{1296}{10}}{9} = \frac{132 - 129,6}{9} = \frac{2,4}{9} \approx 0,2667$$
$$s_d = \sqrt{s_d^2} = \sqrt{\frac{2,4}{9}} \approx \boxed{0,5164}$$

(b) Construa um intervalo de confiança de 95% para a verdadeira redução média de peso μ_d.

1. Valor Crítico: Como analisamos os desvios amostrais para os pares ligados e o grau de precisão global é oculto, usamos a distribuição t de Student. Os graus de liberdade são $gl = n - 1 = 9$. Para $1 - \alpha = 0,95$, temos $\alpha/2 = 0,025$:

$$t_{\alpha/2} = \boxed{2,262}$$

2. Diferença Pontual: Previamente calculada na alínea anterior: $\bar{d} = -3,6$.

3. Erro Padrão (EP):

$$EP = \frac{s_d}{\sqrt{n}} = \frac{0,5164}{\sqrt{10}} = \frac{0,5164}{3,1622} \approx \boxed{0,1633}$$

4. Margem de Erro (E):

$$E = t_{\alpha/2} \times EP = 2,262 \times 0,1633 \approx \boxed{0,3694}$$

5. Construção e Interpretação:

$$IC(\mu_d; 95\%) = [\bar{d} - E\,;\, \bar{d} + E] = [-3,6 - 0,3694\,;\, -3,6 + 0,3694]$$
$$IC(\mu_d; 95\%) = \boxed{[-3,9694\,;\, -3,2306]}$$

(c) Qual suposição é necessária para a validade desse intervalo? Com base no intervalo, é possível afirmar que a dieta reduz o peso?
A suposição necessária para amostras ligadas pequenas ($n = 10 < 30$) é que a verdadeira população associada às diferenças d_i venha de uma **distribuição Normal**.
Sobre o intervalo de confiança construído $[-3,9694\,;\, -3,2306]$, ele **não engloba o valor 0** e localiza-se estritamente na faixa negativa, reforçando a natureza da métrica adotada (Depois − Antes < 0). **Sim**, pode-se afirmar estatisticamente com 95% de confiança que a dieta contribuiu para a redução efetiva dos pesos dos pacientes (entre 3,2 kg e 3,9 kg em média de emagrecimento).

Questão 5

(Retirado de: "Lista Atividade 2 - Intervalo de Confiança para duas populações", Exercício número 5)

Uma pesquisa de mercado quer comparar a proporção de consumidores que preferem a marca A em duas regiões. Na região Norte, de 250 entrevistados, 150 preferem a marca A; na região Sul, de 200 entrevistados, 100 preferem a marca A.

Resolução:
Extração de Dados:
Região Norte: $n_N = 250$, sucesso $x_N = 150 \implies \hat{p}_N = \frac{150}{250} = 0,60$.
Região Sul: $n_S = 200$, sucesso $x_S = 100 \implies \hat{p}_S = \frac{100}{200} = 0,50$.
Nível de confiança $1 - \alpha = 0,95$.

(a) Verifique se as condições para o uso da aproximação normal estão satisfeitas.

Para o uso da aproximação normal, é necessário que o número de sucessos e falhas observadas em ambas as amostras seja maior ou igual a 5. Validando:

$$n_N\hat{p}_N = 150 \geq 5 \quad \text{e} \quad n_N(1-\hat{p}_N) = 100 \geq 5$$
$$n_S\hat{p}_S = 100 \geq 5 \quad \text{e} \quad n_S(1-\hat{p}_S) = 100 \geq 5$$

Como todas as contagens são substancialmente maiores que 5, **as condições estão plenamente satisfeitas**.

(b) Construa um intervalo de confiança de 95% para a diferença entre as proporções populacionais $p_N - p_S$.

1. Valor Crítico: Utilizamos a distribuição Normal Padrão (Z) para amostras grandes em proporções. Para $1-\alpha = 0,95$, temos $\alpha/2 = 0,025$:

$$Z_{\alpha/2} = \boxed{1,96}$$

2. Diferença Pontual:

$$\hat{p}_N - \hat{p}_S = 0,60 - 0,50 = \boxed{0,10}$$

3. Erro Padrão (EP):

$$EP = \sqrt{\frac{\hat{p}_N(1-\hat{p}_N)}{n_N} + \frac{\hat{p}_S(1-\hat{p}_S)}{n_S}}$$
$$EP = \sqrt{\frac{0,60 \times 0,40}{250} + \frac{0,50 \times 0,50}{200}} = \sqrt{\frac{0,24}{250} + \frac{0,25}{200}}$$
$$EP = \sqrt{0,00096 + 0,00125} = \sqrt{0,00221} \approx \boxed{0,0470}$$

4. Margem de Erro (E):

$$E = Z_{\alpha/2} \times EP = 1,96 \times 0,0470 \approx \boxed{0,0921}$$

5. Construção e Interpretação:

$$IC(p_N - p_S; 95\%) = [0,10 - 0,0921\,;\ 0,10 + 0,0921] = \boxed{[0,0079\,;\ 0,1921]}$$

(c) Interprete o intervalo. Há evidência de diferença entre as regiões?

Interpretação: O intervalo de 95% obtido é $[0,0079\,;\ 0,1921]$. Como o intervalo construído **não contém o valor 0** e situa-se inteiramente na porção de valores reais positivos, temos evidência estatística de diferença significativa entre as duas proporções. Conclui-se, de forma direta, que a proporção populacional de preferência à marca A na Região Norte é maior do que na Região Sul.

Questão 6

(Retirado de: "Lista Atividade 2 - Intervalo de Confiança para duas populações", Exercício número 6)

Um laboratório testa a precisão de duas balanças. Para isso, pesa-se 15 vezes um mesmo objeto na balança 1 e 12 vezes na balança 2. As variâncias amostrais obtidas foram $s_1^2 = 0{,}45\ \text{g}^2$ e $s_2^2 = 0{,}32\ \text{g}^2$. Suponha que as medições sigam distribuição normal.

Resolução:
Extração de Dados:
Balança 1: $n_1 = 15 \implies gl_1 = 14,\ s_1^2 = 0,45$.
Balança 2: $n_2 = 12 \implies gl_2 = 11,\ s_2^2 = 0,32$.
Nível de confiança $1 - \alpha = 0,90 \implies \alpha = 0,10$.

(a) Construa um intervalo de confiança de 90% para a razão das variâncias σ_1^2/σ_2^2.
1. Valores Críticos: A razão entre variâncias amostrais independentes de populações normais tem distribuição F de Snedecor de $F(gl_1, gl_2)$. Buscamos os valores críticos que acumulam a proporção $\alpha/2 = 0,05$ nas caudas (sendo que o limite inferior dependerá da inversão dos graus de liberdade):

$$F_{\text{sup}} = F_{0{,}05}(gl_1, gl_2) = F_{0{,}05}(14, 11) \approx 2,63$$
$$F_{\text{inf}} = F_{0{,}05}(gl_2, gl_1) = F_{0{,}05}(11, 14) \approx 2,56$$

2. Razão Pontual:

$$\frac{s_1^2}{s_2^2} = \frac{0,45}{0,32} = \boxed{1,40625}$$

3. Construção do IC: A fórmula teórica do intervalo de confiança para razão de variâncias inversamente aplica as tabelas dos fatores críticos (F):

$$IC\left(\frac{\sigma_1^2}{\sigma_2^2}; 90\%\right) = \left[\frac{s_1^2}{s_2^2} \times \frac{1}{F_{0{,}05}(14, 11)}\ ;\ \frac{s_1^2}{s_2^2} \times F_{0{,}05}(11, 14)\right]$$
$$IC\left(\frac{\sigma_1^2}{\sigma_2^2}; 90\%\right) = \left[1,40625 \times \frac{1}{2,63}\ ;\ 1,40625 \times 2,56\right]$$
$$IC\left(\frac{\sigma_1^2}{\sigma_2^2}; 90\%\right) = \left[\frac{1,40625}{2,63}\ ;\ 3,60\right] \approx \boxed{[0,5347\,;\,3,6000]}$$

(b) Com base no intervalo, é possível afirmar que as balanças têm precisões diferentes (isto é, variâncias diferentes)?
Interpretação: O intervalo de confiança obtido é $[0,5347\,;\,3,6000]$. Como o intervalo **contém o valor 1** (que corresponde à igualdade das variâncias, $\sigma_1^2 = \sigma_2^2$), não há evidência estatística suficiente, a 90% de confiança, para afirmar que as balanças possuem precisões diferentes.

(c) Que suposição é fundamental para a validade desse intervalo?
A suposição fundamental é que as medições de ambas as balanças seguem distribuições

Normais e são independentes entre si. Sem essa suposição, a razão entre as variâncias amostrais não segue a distribuição F de Snedecor, invalidando a construção do intervalo.

Questão 7

(Retirado de: "Lista Atividade 2 - Intervalo de Confiança para duas populações", Exercício número 7)

Um Estatístico deseja estimar a variabilidade do diâmetro de peças produzidas por uma máquina. Ele coleta uma amostra de 20 peças e obtém uma variância amostral de 0,0025 mm^2. Suponha que os diâmetros sigam distribuição normal.

Extração de Dados:
$n = 20 \implies gl = 19$, $S^2 = 0{,}0025$, $1 - \alpha = 0{,}95 \implies \alpha = 0{,}05$.

(a) Construa um intervalo de confiança de 95% para a variância populacional σ^2.
1. Valores Críticos: Como a população é normal, utilizamos a distribuição Qui-Quadrado (χ^2) com $gl = n - 1 = 19$. Para $1 - \alpha = 0{,}95$, calculamos primeiro o quantil de ordem $1 - \alpha/2 = 0{,}975$ e depois o quantil de ordem $\alpha/2 = 0{,}025$:

$$\chi^2_{0{,}975;\,19} = \boxed{32{,}852}$$
$$\chi^2_{0{,}025;\,19} = \boxed{8{,}907}$$

2. Construção do IC:

$$IC(\sigma^2; 95\%) = \left[\frac{(n-1)S^2}{\chi^2_{0{,}975}}\,;\,\frac{(n-1)S^2}{\chi^2_{0{,}025}}\right]$$
$$IC(\sigma^2; 95\%) = \left[\frac{(20-1)\times 0{,}0025}{32{,}852}\,;\,\frac{(20-1)\times 0{,}0025}{8{,}907}\right]$$
$$IC(\sigma^2; 95\%) = [0{,}00145\,;\,0{,}00533] \approx \boxed{[0{,}00145\,;\,0{,}00533]}$$

(b) Obtenha o intervalo de confiança para o desvio padrão populacional σ.
1. Construção do IC: Como o desvio padrão é a raiz quadrada positiva da variância, basta extrair a raiz dos limites obtidos em (a):

$$IC(\sigma; 95\%) = \left[\sqrt{0{,}00145}\,;\,\sqrt{0{,}00533}\right] \approx \boxed{[0{,}0380\,;\,0{,}0730]}$$

(c) Interprete os resultados.
Interpretação: Estima-se, com 95% de confiança, que a verdadeira variância populacional dos diâmetros produzidos pela máquina esteja entre 0,00145 mm^2 e 0,00533 mm^2. Consequentemente, o desvio padrão populacional está estimado entre 0,0380 mm e 0,0730 mm.

Questão 8

(Retirado de: "Lista Atividade 2 - Intervalo de Confiança para duas populações", Exercício número 8)

Responda de forma justificada:

(a) Qual a principal diferença entre o intervalo de confiança para a diferença de médias quando as variâncias populacionais são conhecidas e quando são desconhecidas, mas iguais?
A principal diferença está na distribuição utilizada para obter o valor crítico e na forma do erro padrão. Quando as variâncias populacionais são **conhecidas** (σ_1^2, σ_2^2), utiliza-se a distribuição Normal Padrão (Z) e o erro padrão é calculado diretamente com os valores populacionais. Quando as variâncias são **desconhecidas, mas assumidas iguais**, utiliza-se a distribuição t de Student, e as variâncias amostrais são combinadas em uma estimativa única (S_p^2), com graus de liberdade $gl = n_1 + n_2 - 2$.

(b) Por que, no caso de amostras pareadas, a análise é feita sobre as diferenças em vez das observações originais?
Em amostras pareadas, as duas medições são realizadas sobre o mesmo sujeito ou unidade experimental (e.g., antes e depois de um tratamento). Tratar essas observações como independentes ignoraria a correlação natural entre elas, inflando artificialmente a variabilidade estimada. Ao calcular as diferenças individuais $d_i = X_{\text{depois},i} - X_{\text{antes},i}$, essa correlação é eliminada e o problema reduz-se à inferência sobre uma única média populacional μ_d, tornando a análise mais precisa e conceitualmente correta.

(c) O que significa, na prática, um intervalo de confiança para a razão de variâncias conter o valor 1?
O valor 1 representa a igualdade entre as variâncias populacionais, pois $\frac{\sigma_1^2}{\sigma_2^2} = 1 \implies \sigma_1^2 = \sigma_2^2$. Quando o intervalo de confiança construído pela distribuição F contém o valor 1 (e.g., $[0{,}8\,;\,1{,}4]$), não há evidência estatística suficiente para rejeitar a hipótese de igualdade das variâncias ($H_0 : \sigma_1^2 = \sigma_2^2$). Na prática, isso indica que as duas populações podem ser consideradas homocedásticas, o que justifica, por exemplo, o uso da variância combinada S_p^2 na comparação de médias.

Questão 9

(Retirado de: Slide "Aula 7.1 - Intervalo de confiança para duas populações", Exercício 1 - pág. 122)

Um estudo deseja comparar o desempenho de alunos submetidos a dois métodos de ensino: Método A (tradicional): 15 alunos. Método B (inovador): 12 alunos. A variável resposta é a nota em um teste padronizado (0 a 100). Os dados resumidos são: Método A: $n_A = 15$, $\bar{x}_A = 72{,}0$, $s_A^2 = 25{,}0$ Método B: $n_B = 12$, $\bar{x}_B = 78{,}0$, $s_B^2 = 30{,}0$

Supondo populações normais e que as variâncias populacionais sejam iguais, construa um intervalo de confiança de 95% para a diferença entre as médias populacionais ($\mu_B - \mu_A$). Interprete o resultado.

Resolução:
Extração de Dados:
$n_A = 15$, $\bar{x}_A = 72{,}0$, $S_A^2 = 25{,}0$.
$n_B = 12$, $\bar{x}_B = 78{,}0$, $S_B^2 = 30{,}0$.
$\sigma_A^2 = \sigma_B^2$ (desconhecidas), $1 - \alpha = 0{,}95 \implies \alpha = 0{,}05$.
1. Variância Combinada (S_p^2): Como assumimos igualdade das variâncias populacionais, combinamos as variâncias amostrais:

$$S_p^2 = \frac{(n_A - 1)S_A^2 + (n_B - 1)S_B^2}{n_A + n_B - 2}$$
$$S_p^2 = \frac{(15 - 1)25{,}0 + (12 - 1)30{,}0}{15 + 12 - 2} = \frac{14 \times 25 + 11 \times 30}{25}$$
$$S_p^2 = \frac{350 + 330}{25} = \frac{680}{25} = \boxed{27{,}2}$$

2. Valor Crítico: Utilizamos a distribuição t de Student com $gl = n_A + n_B - 2 = 25$. Para 95% de confiança ($\alpha/2 = 0{,}025$):

$$t_{0{,}025;\,25} = \boxed{2{,}060}$$

3. Diferença Pontual:

$$\bar{x}_B - \bar{x}_A = 78{,}0 - 72{,}0 = \boxed{6{,}0}$$

4. Erro Padrão (EP):

$$EP = \sqrt{S_p^2 \left(\frac{1}{n_B} + \frac{1}{n_A} \right)}$$
$$EP = \sqrt{27{,}2 \left(\frac{1}{12} + \frac{1}{15} \right)} = \sqrt{27{,}2 \left(\frac{9}{60} \right)} = \sqrt{27{,}2 \times 0{,}15} = \sqrt{4{,}08} \approx \boxed{2{,}020}$$

5. Margem de Erro (E):

$$E = t_{\alpha/2} \times EP = 2{,}060 \times 2{,}020 \approx \boxed{4{,}161}$$

6. Construção e Interpretação:

$$IC(\mu_B - \mu_A; 95\%) = [6{,}0 - 4{,}161\,;\, 6{,}0 + 4{,}161] = \boxed{[1{,}839\,;\, 10{,}161]}$$

Interpretação: Estima-se, com 95% de confiança, que a verdadeira diferença entre as médias populacionais $\mu_B - \mu_A$ esteja entre 1,839 e 10,161 pontos. O intervalo **não contém o valor 0** e abriga apenas valores positivos, logo, conclui-se que existe diferença estatisticamente significativa e que o Método B gera médias superiores às do Método A.

Questão 10

(Retirado de: Slide "Aula 7.1 - Intervalo de confiança para duas populações", Exercício 2 - pág. 123)

Um programa de bem-estar avaliou o nível de estresse (escala de 0 a 50) de 10 funcionários antes e depois de um curso de mindfulness. As diferenças individuais (depois - antes) foram: d_i: $-5, -3, -4, -6, -2, -3, -5, -4, -3, -4$ (valores negativos indicam redução do estresse).

Construa um intervalo de confiança de 95% para a verdadeira redução média do estresse (μ_d) proporcionada pelo programa. Interprete o resultado, verificando se o intervalo contém o valor 0. É necessário fazer alguma suposição?

Resolução:
Extração de Dados:
$d_i = \text{Depois} - \text{Antes}$. $n = 10$ pares amostrais.
$1 - \alpha = 0{,}95 \implies \alpha = 0{,}05$.
1. Estatísticas das Diferenças:

$$\bar{d} = \frac{\sum d_i}{n} = \frac{-39}{10} = \boxed{-3{,}9}$$
$$S_d^2 = \frac{\sum d_i^2 - \frac{(\sum d_i)^2}{n}}{n-1} = \frac{165 - \frac{(-39)^2}{10}}{9} = \frac{165 - 152{,}1}{9} = \frac{12{,}9}{9} \approx \boxed{1{,}4333}$$
$$S_d = \sqrt{1{,}4333} \approx \boxed{1{,}1972}$$

2. Valor Crítico: Utilizamos a distribuição t de Student para $gl = n - 1 = 9$. Para 95% de confiança ($\alpha/2 = 0{,}025$):

$$t_{0{,}025;\,9} = \boxed{2{,}262}$$

3. Erro Padrão (EP):

$$EP = \frac{S_d}{\sqrt{n}} = \frac{1{,}1972}{\sqrt{10}} \approx \boxed{0{,}3786}$$

4. Margem de Erro (E):

$$E = t_{\alpha/2} \times EP = 2{,}262 \times 0{,}3786 \approx \boxed{0{,}8564}$$

5. Construção e Interpretação:

$$IC(\mu_d; 95\%) = [\bar{d} - E\,;\, \bar{d} + E] = [-3{,}9 - 0{,}8564\,;\, -3{,}9 + 0{,}8564]$$
$$IC(\mu_d; 95\%) = \boxed{[-4{,}7564\,;\, -3{,}0436]}$$

Interpretação e Suposições: O intervalo construído **não contém o valor 0** e é estritamente negativo. Isso indica que houve uma redução estatisticamente significativa nos níveis de estresse ($d_i = \text{Depois} - \text{Antes} < 0$). Estima-se, com 95% de confiança, que a

redução média induzida pelo curso seja entre 3,04 e 4,76 pontos.
Como a amostra tem tamanho $n = 10$, a **suposição necessária** para a precisão matemática deste IC é assumir premissa de que a distribuição populacional das diferenças seja Normal.

Questão 11

(Retirado de: Slide "Aula 7.1 - Intervalo de confiança para duas populações", Exercício 3 - pág. 124)

Uma montadora de veículos quer comparar o consumo de combustível (km/l) de dois modelos de carro. Por especificações técnicas do fabricante, as variâncias populacionais são conhecidas. Dados: Modelo A: $\sigma_A = 1{,}2$ km/l, $n_A = 35$, $\bar{x}_A = 12{,}8$ km/l Modelo B: $\sigma_B = 1{,}5$ km/l, $n_B = 30$, $\bar{x}_B = 11{,}9$ km/l Nível de confiança: 95% ($\alpha = 0{,}05$)

Resolução:
Extração de Dados:
$n_A = 35$, $\bar{x}_A = 12{,}8$, $\sigma_A = 1{,}2$.
$n_B = 30$, $\bar{x}_B = 11{,}9$, $\sigma_B = 1{,}5$.
Variâncias populacionais **conhecidas**. $1 - \alpha = 0{,}95 \implies \alpha = 0{,}05$.

(a) Construa o intervalo de confiança para $\mu_A - \mu_B$.
1. Valor Crítico: Como as variâncias populacionais são formalmente conhecidas, utilizamos estritamente a distribuição Normal Padrão (Z).

$$Z_{0{,}025} = \boxed{1{,}96}$$

2. Diferença Pontual:

$$\bar{x}_A - \bar{x}_B = 12{,}8 - 11{,}9 = \boxed{0{,}9}$$

3. Erro Padrão (EP): Utilizamos os desvios paramétricos reais:

$$EP = \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}} = \sqrt{\frac{1{,}44}{35} + \frac{2{,}25}{30}} = \sqrt{0{,}04114 + 0{,}0750} \approx \sqrt{0{,}11614} \approx \boxed{0{,}3408}$$

4. Margem de Erro (E):

$$E = Z_{\alpha/2} \times EP = 1{,}96 \times 0{,}3408 \approx \boxed{0{,}6680}$$

5. Construção do IC:

$$IC(\mu_A - \mu_B; 95\%) = [0{,}9 - 0{,}6680\,;\, 0{,}9 + 0{,}6680] = \boxed{[0{,}2320\,;\, 1{,}5680]}$$

(b) Interprete o resultado.
Interpretação: Estima-se, com 95% de confiança, que o modelo A consome em média entre 0,232 km/l e 1,568 km/l a mais de combustível na média em comparação ao modelo

B.

(c) É possível afirmar que os modelos têm consumos médios diferentes?
Conclusão: Sim. Como o intervalo de confiança [0,2320 ; 1,5680] **não contém o valor 0** e situa-se integralmente no eixo positivo, há evidência estatística sólida de que existe diferença significativa no consumo e que o consumo médio real do Modelo A (μ_A) é maior que o do Modelo B (μ_B).

Questão 12

(Retirado de: Slide "Aula 7.1 - Intervalo de confiança para duas populações", Exercício 4 - pág. 125)

Um agrônomo quer comparar a produtividade (em toneladas por hectare) de duas variedades de soja cultivadas em diferentes regiões. Devido às características do solo, suspeita-se que as variabilidades sejam diferentes. Dados coletados após a colheita: Variedade A (transgênica): $n_1 = 9$ fazendas, $\bar{x}_1 = 4{,}2$ t/ha, $s_1 = 0{,}85$ t/ha Variedade B (convencional): $n_2 = 11$ fazendas, $\bar{x}_2 = 3{,}6$ t/ha, $s_2 = 0{,}42$ t/ha Como as amostras são pequenas ($n_1 = 9, n_2 = 11$), o agrônomo verificou a normalidade com gráficos e testes.

Estime com 95% de confiança um intervalo para a diferença das médias.

Resolução:
Extração de Dados:
$n_1 = 9$, $\bar{x}_1 = 4{,}2$, $S_1 = 0{,}85 \implies S_1^2 = 0{,}7225$.
$n_2 = 11$, $\bar{x}_2 = 3{,}6$, $S_2 = 0{,}42 \implies S_2^2 = 0{,}1764$.
Variâncias amostrais, populações com variâncias **desconhecidas e desiguais** (Behrens-Fisher). $1 - \alpha = 0{,}95 \implies \alpha = 0{,}05$.
1. Valor Crítico (Aproximação de Welch): Calculamos os graus de liberdade equivalentes (ν):

$$\nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1-1} + \frac{(S_2^2/n_2)^2}{n_2-1}} = \frac{\left(\frac{0{,}7225}{9} + \frac{0{,}1764}{11}\right)^2}{\frac{(0{,}7225/9)^2}{8} + \frac{(0{,}1764/11)^2}{10}}$$

$$\nu = \frac{(0{,}08028 + 0{,}01604)^2}{\frac{(0{,}08028)^2}{8} + \frac{(0{,}01604)^2}{10}} = \frac{0{,}00928}{0{,}000806 + 0{,}000026} = \frac{0{,}00928}{0{,}000832} \approx 11{,}15$$

Arredondando ν conservadoramente para 11. Utilizando a distribuição t com $gl = 11$:

$$t_{0{,}025;\,11} = \boxed{2{,}201}$$

2. Diferença Pontual:

$$\bar{x}_1 - \bar{x}_2 = 4{,}2 - 3{,}6 = \boxed{0{,}6}$$

3. Erro Padrão (EP):

$$EP = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = \sqrt{0{,}08028 + 0{,}01604} = \sqrt{0{,}09632} \approx \boxed{0{,}3104}$$

4. Margem de Erro (E):

$$E = t_{\alpha/2} \times EP = 2{,}201 \times 0{,}3104 \approx \boxed{0{,}6832}$$

5. Construção e Interpretação:

$$IC(\mu_1 - \mu_2; 95\%) = [0{,}6 - 0{,}6832\,;\, 0{,}6 + 0{,}6832] = \boxed{[-0{,}0832\,;\, 1{,}2832]}$$

Interpretação: Estima-se, com 95% de confiança, que a verdadeira diferença média na produtividade ($\mu_1 - \mu_2$) varie de $-0{,}0832$ t/ha a 1,2832 t/ha. Como o intervalo **contém o valor 0**, conclui-se que não há evidência estatística suficiente para afirmar categoricamente que há distinção na média populacional de produtividade final atestada entre as duas variedades de soja nestas condições amostrais limitantes.

Questão 13

(Retirado de: Slide "Aula 7.1 - Intervalo de confiança para duas populações", Exercício 5 - pág. 139)

Uma pesquisa de satisfação comparou a proporção de clientes satisfeitos entre duas lojas. Dados: Loja A: $n_A = 150$, $x_A = 120$ clientes satisfeitos Loja B: $n_B = 180$, $x_B = 126$ clientes satisfeitos

Extração de Dados:
$n_A = 150$, $x_A = 120 \implies \hat{p}_A = \frac{120}{150} = 0{,}80$.
$n_B = 180$, $x_B = 126 \implies \hat{p}_B = \frac{126}{180} = 0{,}70$.
$1 - \alpha = 0{,}95 \implies \alpha = 0{,}05$.

(a) Verifique as condições para usar o IC (aproximação Normal).
Para garantir o uso da aproximação Normal em casos de proporções, os sucessos e falhas contabilizados nas amostras devem ser ≥ 5:

$$\begin{aligned} n_A\hat{p}_A &= 120 \geq 5 \qquad & n_A(1-\hat{p}_A) &= 30 \geq 5 \\ n_B\hat{p}_B &= 126 \geq 5 \qquad & n_B(1-\hat{p}_B) &= 54 \geq 5 \end{aligned}$$

Conclusão: Como todas as contagens são manifestamente maiores que 5, as condições para a aproximação Normal estão satisfeitas.

(b) Construa um IC de 95% para $p_A - p_B$.
1. Valor Crítico: Utilizamos a distribuição Normal Padrão (Z).

$$Z_{0{,}025} = \boxed{1{,}96}$$

2. Diferença Pontual:

$$\hat{p}_A - \hat{p}_B = 0{,}80 - 0{,}70 = \boxed{0{,}10}$$

3. Erro Padrão (EP):

$$EP = \sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A} + \frac{\hat{p}_B(1-\hat{p}_B)}{n_B}}$$
$$EP = \sqrt{\frac{0{,}80 \times 0{,}20}{150} + \frac{0{,}70 \times 0{,}30}{180}} = \sqrt{\frac{0{,}16}{150} + \frac{0{,}21}{180}}$$
$$EP = \sqrt{0{,}001067 + 0{,}001167} = \sqrt{0{,}002234} \approx \boxed{0{,}0473}$$

4. Margem de Erro (E):

$$E = Z_{\alpha/2} \times EP = 1{,}96 \times 0{,}0473 \approx \boxed{0{,}0927}$$

5. Construção do IC:

$$IC(p_A - p_B; 95\%) = [0{,}10 - 0{,}0927\,;\, 0{,}10 + 0{,}0927] = \boxed{[0{,}0073\,;\, 0{,}1927]}$$

(c) Interprete o resultado.
Interpretação: Estima-se, com 95% de confiança, que a verdadeira diferença percentual na taxa de satisfação populacional (Loja A - Loja B) caia entre 0,73% e 19,27%. Como o intervalo **não contém o zero** e abrange inteiramente a região positiva, conclui-se formalmente que a proporção originária de clientes satisfeitos na Loja A é estatisticamente superior à da Loja B.

Questão 14

(Retirado de: Slide "Aula 7.1 - Intervalo de confiança para duas populações", Exercício 6 - pág. 140)

Um laboratório testa a precisão de uma balança, pesando 12 vezes o mesmo objeto. Dados: $n = 12$ medições. Variância amostral: $s^2 = 0{,}36$ g^2

Extração de Dados:
$n = 12 \implies gl = 11$, $S^2 = 0{,}36$ g^2.
$1 - \alpha = 0{,}90 \implies \alpha = 0{,}10$.

(a) Construa um IC de 90% para a variância verdadeira σ^2.
1. Valores Críticos: Assumindo a premissa de normalidade populacional, balizamos via distribuição Qui-Quadrado (χ^2) em $gl = 11$. Para as caudas de $\alpha/2 = 0{,}05$ e $1 - \alpha/2 = 0{,}95$:

$$\chi^2_{0{,}95;\,11} = \boxed{4{,}575}$$
$$\chi^2_{0{,}05;\,11} = \boxed{19{,}675}$$

2. Construção do IC:

$$IC(\sigma^2; 90\%) = \left[\frac{(n-1)S^2}{\chi^2_{0,05}}; \frac{(n-1)S^2}{\chi^2_{0,95}}\right]$$
$$IC(\sigma^2; 90\%) = \left[\frac{11 \times 0{,}36}{19{,}675}; \frac{11 \times 0{,}36}{4{,}575}\right]$$
$$IC(\sigma^2; 90\%) = \left[\frac{3{,}96}{19{,}675}; \frac{3{,}96}{4{,}575}\right] \approx \boxed{[0{,}2013\,;\,0{,}8656]}$$

(b) Construa um IC de 90% para o desvio padrão verdadeiro σ.
1. Construção do IC: Extraindo a raiz quadrada matemática exata dos limites delineados para o erro da variância na alínea (a):

$$IC(\sigma; 90\%) = [\sqrt{0{,}2013}\,;\,\sqrt{0{,}8656}] \approx \boxed{[0{,}4487\,;\,0{,}9304]}$$

(c) Interprete os resultados.
Interpretação: Estima-se, com 90% de confiança, que a verdadeira variância da balança esteja entre 0,2013 g^2 e 0,8656 g^2. Consequentemente, o verdadeiro desvio padrão está estimado entre 0,4487 g e 0,9304 g.

Questão 15

(Retirado de: Slide "Aula 7.1 - Intervalo de confiança para duas populações", Exercício 7 - pág. 141)

Dois operadores medem a mesma peça várias vezes para comparar a precisão de suas medições. Dados: Operador 1: $n_1 = 15$, $s_1^2 = 2{,}5$ mm^2 Operador 2: $n_2 = 12$, $s_2^2 = 1{,}8$ mm^2

Resolução:
Extração de Dados:
Operador 1: $n_1 = 15 \implies gl_1 = 14$, $s_1^2 = 2{,}5$.
Operador 2: $n_2 = 12 \implies gl_2 = 11$, $s_2^2 = 1{,}8$.
$1 - \alpha = 0{,}95 \implies \alpha = 0{,}05$.

(a) Construa um IC de 95% para σ_1^2/σ_2^2.
1. Valores Críticos: Pela definição do intervalo para razão de variâncias, utilizamos a distribuição F de Snedecor. Buscamos as abscissas que alocam a área $\alpha/2 = 0{,}025$ nas caudas.

$$F_{\text{sup}} = F_{0,025}(14, 11) \approx 3{,}359$$
$$F_{\text{inf}} = F_{0,025}(11, 14) \approx 3{,}095$$

2. Razão Pontual:

$$\frac{s_1^2}{s_2^2} = \frac{2{,}5}{1{,}8} \approx \boxed{1{,}3889}$$

3. Construção do IC:

$$IC\left(\frac{\sigma_1^2}{\sigma_2^2}; 95\%\right) = \left[\frac{s_1^2}{s_2^2} \times \frac{1}{F_{0,025}(14, 11)}\,;\, \frac{s_1^2}{s_2^2} \times F_{0,025}(11, 14)\right]$$
$$IC\left(\frac{\sigma_1^2}{\sigma_2^2}; 95\%\right) = \left[1{,}3889 \times \frac{1}{3{,}359}\,;\, 1{,}3889 \times 3{,}095\right]$$
$$IC\left(\frac{\sigma_1^2}{\sigma_2^2}; 95\%\right) = \left[\frac{1{,}3889}{3{,}359}\,;\, 4{,}2986\right] \approx \boxed{[0{,}4135\,;\, 4{,}2986]}$$

(b) Os operadores têm precisões diferentes?
Interpretação: Como o intervalo de confiança [0,4135 ; 4,2986] **contém o valor 1**, que marca a razão de variâncias estatisticamente iguais ($\sigma_1^2 = \sigma_2^2$), não há evidência estatística, com 95% de confiança, para afirmar que os operadores possuam precisões diferentes.

(c) Que suposição é necessária para a validade do IC?
É pressuposto obrigatório que as medidas coletadas pelos dois operadores sejam variáveis (População 1 e População 2) provenientes de distribuições normais, contínuas e independentes entre si.

Questão 16

(Retirado de: Slide "Aula 8 - Teste de Hipotese", Exercício 4 - pág. 233)

Um estudo compara dois tipos de ração para ganho de peso em animais. Para a ração A, com 40 animais, obteve-se ganho médio de 2,5 kg, com desvio padrão populacional conhecido de 0,8 kg. Para a ração B, com 36 animais, ganho médio de 2,2 kg, com desvio padrão populacional conhecido de 0,7 kg.

Teste se há diferença significativa entre as rações, com $\alpha = 0{,}05$.

Resolução:
Extração de Dados:
Ração A: $n_A = 40$, $\bar{x}_A = 2{,}5$, $\sigma_A = 0{,}8$.
Ração B: $n_B = 36$, $\bar{x}_B = 2{,}2$, $\sigma_B = 0{,}7$.
Desvios padrão populacionais **conhecidos**.

1. Estabelecer as hipóteses:

$$H_0 : \mu_A - \mu_B = 0$$
$$H_1 : \mu_A - \mu_B \neq 0$$

2. Fixar o nível de significância: Valor expresso $\alpha = 0{,}05$.
3. Definir a Estatística de Teste: Como os desvios padrão populacionais (σ_A, σ_B) são perfeitamente conhecidos, a estatística modelada exata é Z (Normal Padrão).

$$Z = \frac{(\bar{x}_A - \bar{x}_B) - \Delta}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}}$$

4. Definir a região de rejeição: Para um teste bilateral de igualdade com nível $\alpha = 0{,}05$, observamos $\alpha/2 = 0{,}025$ nas caudas. O escore crítico provém da distribuição Normal (Z):

$$Z_{crit} = \pm 1{,}96$$

Rejeita-se H_0 se $Z_{calc} < -1{,}96$ ou $Z_{calc} > 1{,}96$.
5. Calcular a estatística do teste: Assumindo diferença nula $\Delta = 0$:

$$\bar{x}_A - \bar{x}_B = 2{,}5 - 2{,}2 = \boxed{0{,}3}$$
$$EP = \sqrt{\frac{0{,}8^2}{40} + \frac{0{,}7^2}{36}} = \sqrt{\frac{0{,}64}{40} + \frac{0{,}49}{36}} = \sqrt{0{,}016 + 0{,}013611} \approx \sqrt{0{,}029611} \approx \boxed{0{,}1721}$$
$$Z_{calc} = \frac{0{,}3 - 0}{0{,}1721} \approx \boxed{1{,}743}$$

6. Tomar a decisão estatística: Como $-1{,}96 \leq 1{,}743 \leq 1{,}96$, o escore recai na região de não rejeição. Portanto, a decisão objetiva é: **Não rejeitar** H_0.
7. Interpretar a decisão: Não há evidências suficientes para afirmar, ao nível de significância de 5%, que exista diferença estatisticamente significativa nos ganhos de peso acarretados pelas rações A e B. As eficácias médias em engorda comportam-se de modo estatisticamente equivalente.

Questão 17

(Retirado de: Slide "Aula 8 - Teste de Hipotese", Exercício 5 - pág. 234)

Dois métodos de ensino são comparados. No método tradicional, 18 alunos tiveram média 68 e desvio padrão 6. No método novo, 15 alunos tiveram média 72 e desvio padrão 5.

Assumindo variâncias populacionais iguais, teste se o método novo é melhor com $\alpha = 0{,}05$. É necessário verificar alguma suposição? Qual?

Resolução:
Extração de Dados:
Método Tradicional (1): $n_1 = 18$, $\bar{x}_1 = 68$, $s_1 = 6 \implies s_1^2 = 36$.
Método Novo (2): $n_2 = 15$, $\bar{x}_2 = 72$, $s_2 = 5 \implies s_2^2 = 25$.
Variâncias populacionais **desconhecidas e assumidas iguais** ($\sigma_1^2 = \sigma_2^2$).

1. Estabelecer as hipóteses: A intenção é provar que o método novo (2) é superior ao tradicional (1), ou seja, $\mu_2 > \mu_1$.

$$H_0 : \mu_2 - \mu_1 \leq 0$$
$$H_1 : \mu_2 - \mu_1 > 0$$

2. Fixar o nível de significância: Valor expresso $\alpha = 0{,}05$.
3. Definir a Estatística de Teste: Como utilizamos amostras empíricas e variâncias

desconhecidas assumidas iguais, a estatística recomendada é t de Student com variância combinada (S_p^2) e graus de liberdade $gl = n_1 + n_2 - 2 = 18 + 15 - 2 = 31$.

$$t = \frac{(\bar{x}_2 - \bar{x}_1) - \Delta}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

4. Definir a região de rejeição: Para um teste unilateral superior com $\alpha = 0{,}05$ e $gl = 31$, inferimos o valor limite diretivo na cauda direita da distribuição t:

$$t_{crit} = \boxed{1{,}6955}$$

Rejeita-se H_0 se $t_{calc} > 1{,}6955$.
5. Calcular a estatística do teste: Sendo a diferença paramétrica testada $\Delta = 0$:

$$S_p^2 = \frac{(18-1)36 + (15-1)25}{18+15-2} = \frac{612+350}{31} = \frac{962}{31} \approx \boxed{31{,}0323}$$
$$EP = \sqrt{31{,}0323\left(\frac{1}{18} + \frac{1}{15}\right)} = \sqrt{31{,}0323\left(\frac{11}{90}\right)} \approx \boxed{1{,}9475}$$
$$t_{calc} = \frac{72-68}{1{,}9475} = \frac{4}{1{,}9475} \approx \boxed{2{,}0539}$$

6. Tomar a decisão estatística: Como $2{,}0539 > 1{,}6955$, a métrica recai ativamente na região crítica. Portanto, a decisão objetiva é: **Rejeitar** H_0.
7. Interpretar a decisão e suposições: Há evidências estatísticas suficientes, a 5% de significância, para afirmar que o método novo resultou em desempenho superior ao tradicional.
Suposição necessária: Como as amostras são pequenas ($n_1, n_2 < 30$), é necessário assumir que as notas provêm de populações com **distribuição Normal** (além da homocedasticidade já assumida no enunciado).

Questão 18

(Retirado de: Slide "Aula 8 - Teste de Hipotese", Exercício 6 - pág. 234)

Um programa de redução de peso avaliou 12 participantes antes e depois do programa. A média das diferenças (depois - antes) foi de -3,2 kg (ou seja, redução de 3,2 kg em média), com desvio padrão das diferenças de 2,5 kg.

Teste se o programa é eficaz na redução de peso, com $\alpha = 0{,}01$. A suposição de normalidade das diferenças foi verificada.

Resolução:
Extração de Dados:
$d_i = \text{Depois} - \text{Antes}$. Variável base empírica é a **redução** ($d_i < 0$ implica que o indivíduo perdeu peso).
$n = 12 \implies gl = 11$, $\bar{d} = -3{,}2$, $s_d = 2{,}5$.

1. Estabelecer as hipóteses: A eficácia almejada do programa depende estritamente de uma regressão contínua no peso dos avaliados (d_i negativo).

$$H_0 : \mu_d \geq 0$$
$$H_1 : \mu_d < 0$$

2. Fixar o nível de significância: Valor nominal $\alpha = 0{,}01$.
3. Definir a Estatística de Teste: Como os dados são pareados e o desvio padrão populacional é desconhecido, emprega-se o teste t de Student para amostras dependentes com $gl = n - 1 = 11$:

$$t = \frac{\bar{d} - \Delta_0}{\frac{s_d}{\sqrt{n}}}$$

4. Definir a região de rejeição: Sendo o teste formatado correlacionado sob diretriz restritiva ($\alpha = 0{,}01$), o eixo de rejeição restringe-se à cauda esquerda rigorosa:

$$t_{crit} = \boxed{-2{,}7181}$$

Rejeita-se H_0 se $t_{calc} < -2{,}7181$.
5. Calcular a estatística do teste: Validando numericamente a referencial nula ($\Delta_0 = 0$):

$$EP = \frac{s_d}{\sqrt{n}} = \frac{2{,}5}{\sqrt{12}} \approx \boxed{0{,}7217}$$
$$t_{calc} = \frac{-3{,}2 - 0}{0{,}7217} \approx \boxed{-4{,}4341}$$

6. Tomar a decisão estatística: Como $-4{,}4341 < -2{,}7181$, o valor calculado cai na região de rejeição. Portanto, a decisão é: **Rejeitar** H_0.
7. Interpretar a decisão: Há evidências estatísticas suficientes, a 1% de significância, para afirmar que o programa é eficaz na redução de peso, resultando em perda média significativa nos participantes avaliados.

Questão 19

(Retirado de: Slide "Aula 8 - Teste de Hipotese", Exercício 7 - pág. 235)

Um estudo comparou dois tipos de tratamento para redução da pressão arterial. Os resultados (redução em mmHg) foram:

	Tratamento A	**Tratamento B**
Tamanho da amostra	$n_1 = 15$	$n_2 = 12$
Média amostral	$\bar{x}_1 = 12{,}5$	$\bar{x}_2 = 8{,}3$
Desvio padrão amostral	$s_1 = 4{,}2$	$s_2 = 2{,}8$

Considerando que as variâncias populacionais são desconhecidas e diferentes e as populações são normais, teste se o tratamento A é mais eficaz que o tratamento B, com nível de significância $\alpha = 0{,}05$.

Resolução:
Extração de Dados:
Tratamento A (1): $n_1 = 15$, $\bar{x}_1 = 12{,}5$, $s_1 = 4{,}2 \implies s_1^2 = 17{,}64$.
Tratamento B (2): $n_2 = 12$, $\bar{x}_2 = 8{,}3$, $s_2 = 2{,}8 \implies s_2^2 = 7{,}84$.
Variâncias populacionais **desconhecidas e diferentes** ($\sigma_1^2 \neq \sigma_2^2$). Populações normais assumidas.

1. Estabelecer as hipóteses: Pretende-se provar que o Tratamento A é mais eficaz (gera maior redução aferida), ou seja, $\mu_1 > \mu_2$.

$$H_0 : \mu_1 - \mu_2 \leq 0$$
$$H_1 : \mu_1 - \mu_2 > 0$$

2. Fixar o nível de significância: Valor expresso $\alpha = 0{,}05$.
3. Definir a Estatística de Teste: Como utilizamos amostras empíricas com variâncias populacionais assumidamente diferentes, recorre-se ao teste t de Student sob a **aproximação de Welch** para graus de liberdade.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

4. Definir a região de rejeição: Calculando os graus de liberdade aferidos por Welch (ν):

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} = \frac{\left(\frac{17{,}64}{15} + \frac{7{,}84}{12}\right)^2}{\frac{(17{,}64/15)^2}{14} + \frac{(7{,}84/12)^2}{11}} \approx \boxed{24{,}3223}$$

Abaixando conservadoramente ν para 24, o limite crítico na cauda direita da distribuição t será:

$$t_{crit} = \boxed{1{,}7109}$$

Rejeita-se H_0 se $t_{calc} > 1{,}7109$.
5. Calcular a estatística do teste: Testando frente à hipótese de diferença nula $\Delta = 0$:

$$EP = \sqrt{\frac{17{,}64}{15} + \frac{7{,}84}{12}} \approx \boxed{1{,}3525}$$
$$t_{calc} = \frac{12{,}5 - 8{,}3}{1{,}3525} = \frac{4{,}2}{1{,}3525} \approx \boxed{3{,}1053}$$

6. Tomar a decisão estatística: Como $3{,}1053 > 1{,}7109$, o escore recai de forma consolidada dentro da região crítica superior. Portanto, a decisão objetiva é: **Rejeitar**

H_0.
7. Interpretar a decisão: Existem evidências estatísticas significativas, ao nível de 5%, para certificar na prática que o Tratamento A exibe uma eficácia média de redução da pressão arterial validamente superior ao induzido pelo Tratamento B.

Questão 20

(Retirado de: Slide "Aula 8 - Teste de Hipotese", Exercício 8 - pág. 239)

Uma pesquisa eleitoral entrevistou eleitores em duas cidades sobre a intenção de voto no candidato X. Os resultados foram:

	Cidade A	**Cidade B**
Número de eleitores entrevistados	$n_A = 200$	$n_B = 180$
Número que votariam no candidato X	$x_A = 110$	$x_B = 85$

Teste se a proporção de eleitores do candidato X é maior na cidade A do que na cidade B, com nível de significância $\alpha = 0{,}05$.

Resolução:
Extração de Dados:
Cidade A: $n_A = 200$, sucessos $x_A = 110 \implies \hat{p}_A = \frac{110}{200} = 0{,}5500$.
Cidade B: $n_B = 180$, sucessos $x_B = 85 \implies \hat{p}_B = \frac{85}{180} \approx 0{,}4722$.

1. Estabelecer as hipóteses: Deseja-se atestar que a proporção na cidade A é maior que na cidade B ($p_A > p_B$).

$$H_0 : p_A - p_B \leq 0$$
$$H_1 : p_A - p_B > 0$$

2. Fixar o nível de significância: Valor expresso $\alpha = 0{,}05$.
3. Definir a Estatística de Teste: Tratando referencial de comparação de proporções populacionais autônomas compostas de amostras consolidadas ($n > 30$), aplicamos a estatística Normal Padrão (Z). Assumindo a igualdade parametrizada sob a hipótese nula ($p_A = p_B$), extrai-se a proporção amostral combinada $\hat{p}$:

$$\hat{p} = \frac{x_A + x_B}{n_A + n_B} = \frac{110 + 85}{200 + 180} = \frac{195}{380} \approx \boxed{0{,}5132}$$
$$Z = \frac{(\hat{p}_A - \hat{p}_B) - \Delta}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}}$$

4. Definir a região de rejeição: Dado um escopo avaliativo unilateral de direção superior para margem de tolerância $\alpha = 0{,}05$, o escore limite de distribuição Z provém do corte à direita:

$$Z_{crit} = \boxed{1{,}6449}$$

Rejeita-se H_0 se $Z_{calc} > 1{,}6449$.

5. Calcular a estatística do teste: Arbitrando o referencial isolado nulo testificado $\Delta = 0$:

$$EP = \sqrt{0{,}5132 \times (1 - 0{,}5132) \times \left(\frac{1}{200} + \frac{1}{180}\right)} \approx \boxed{0{,}0514}$$
$$Z_{calc} = \frac{0{,}5500 - 0{,}4722}{0{,}0514} = \frac{0{,}0778}{0{,}0514} \approx \boxed{1{,}5146}$$

6. Tomar a decisão estatística: Como $1{,}5146 \leq 1{,}6449$, o valor calculado não ultrapassa o limite crítico e permanece na região de não rejeição. Portanto, a decisão é: **Não rejeitar** H_0.

7. Interpretar a decisão: Não há evidências estatísticas suficientes, ao nível de 5% de significância, para afirmar que a proporção de eleitores favoráveis ao candidato X é maior na cidade A do que na cidade B.

Questão 21

(Retirado de: Slide "Aula 8 - Teste de Hipotese", Exercício 9 - pág. 245)

Um laboratório testa dois métodos de análise química para verificar se apresentam a mesma precisão (variabilidade). Os resultados de 10 análises com cada método foram:

Método	**Método 1**	**Método 2**
Tamanho da amostra	$n_1 = 10$	$n_2 = 10$
Variância amostral	$s_1^2 = 4{,}8$	$s_2^2 = 2{,}3$

O cientista de dados fez um teste e verificou normalidade dos dados. Teste se o método 1 apresenta maior variabilidade que o método 2, com nível de significância $\alpha = 0{,}05$.

Resolução:

Extração de Dados:

Método 1: $n_1 = 10 \implies gl_1 = 9$, $s_1^2 = 4{,}8$.

Método 2: $n_2 = 10 \implies gl_2 = 9$, $s_2^2 = 2{,}3$.

Distribuições populacionais atestadamente normais.

1. Estabelecer as hipóteses: A suspeita motriz prega que a variância do método 1 desponte estritamente maior que a segunda ($\sigma_1^2 > \sigma_2^2$, isto é, $\frac{\sigma_1^2}{\sigma_2^2} > 1$).

$$H_0 : \sigma_1^2 \leq \sigma_2^2$$
$$H_1 : \sigma_1^2 > \sigma_2^2$$

2. Fixar o nível de significância: Valor expresso $\alpha = 0{,}05$.
3. Definir a Estatística de Teste: Como o objetivo é comparar variâncias de duas populações normais independentes, utiliza-se a distribuição F de Snedecor:

$$F = \frac{s_1^2}{s_2^2}$$

4. Definir a região de rejeição: Para um teste estritamente unilateral atrelado à cauda direita sob cota estanque de probabilidade $\alpha = 0{,}05$ transpassando graus de liberdade conjunturais $(gl_1 = 9, gl_2 = 9)$:

$$F_{crit} = F_{0{,}05}(9,9) = \boxed{3{,}1789}$$

Rejeita-se H_0 se $F_{calc} > 3{,}1789$.
5. Calcular a estatística do teste: Substituindo pontualmente em fração as variâncias amostrais isoladas e documentadas:

$$F_{calc} = \frac{4{,}8}{2{,}3} \approx \boxed{2{,}0870}$$

6. Tomar a decisão estatística: Como $2{,}0870 \leq 3{,}1789$, o valor calculado não ultrapassa o limite crítico e permanece na região de não rejeição. Portanto, a decisão é: **Não rejeitar** H_0.
7. Interpretar a decisão: Não há evidências estatísticas suficientes, ao nível de 5% de significância, para afirmar que o Método 1 apresenta maior variabilidade que o Método 2.

Questão 22

(Retirado de: Slide "Aula 9 - Alguns Testes Especiais", Exercício 1 - pág. 269)

Uma empresa de telefonia quer saber se existe relação entre o plano contratado (Básico, Intermediário, Premium) e a satisfação do cliente (Satisfeito, Insatisfeito). Foram entrevistados 300 clientes. (Os dados geraram uma tabela de contingência a ser analisada, totalizando: Básico = 60 satisfeitos e 40 insatisfeitos; Intermediário = 75 satisfeitos e 25 insatisfeitos; Premium = 85 satisfeitos e 15 insatisfeitos).

Resolução:
Extração de Dados:
Variables Nominais Cruzadas: Plano Contratado (Básico, Intermediário, Premium) e Satisfação (Satisfeito, Insatisfeito).
Total Geral Contabilizado: $N = 300$.

(a) Construa a tabela de contingência entre plano e satisfação.
2. Construir Tabela de Contingência (Frequências Observadas):

(c) Calcule as frequências esperadas. Alguma célula tem valor esperado < 5?
3. Calcular Totais Marginais e Frequências Esperadas: Usando a fórmula global $E_{ij} = \frac{(\text{Total Linha } i)\times(\text{Total Coluna } j)}{N}$, observamos simetria de linhas, gerando os aportes para

Plano	Satisfeito	Insatisfeito	Total (T_{linha})
Básico	60	40	100
Intermediário	75	25	100
Premium	85	15	100
Total (T_{coluna})	220	80	$N =$ **300**

todos os planos fixos:

$$E_{\text{Qualquer Plano, Satisfeito}} = \frac{100 \times 220}{300} \approx \boxed{73{,}3333}$$
$$E_{\text{Qualquer Plano, Insatisfeito}} = \frac{100 \times 80}{300} \approx \boxed{26{,}6667}$$

Conclusão Analítica Auxiliar: Não existem absolutamente células contendo aporte projetado inferior à baliza métrica. Todas excedem folgadamente o valor < 5, blindando categoricamente a presunção aplicável para a inferência paramétrica do teste Qui-Quadrado.

(b) Existe associação entre o tipo de plano e a satisfação do cliente? (use $\alpha = 0{,}05$).

1. Definir as Hipóteses:

H_0 : O tipo de plano e o perfil de satisfação são estocasticamente independentes.

H_1 : Há dependência categórica e associação entre plano e índice de satisfação.

4. Calcular a estatística χ^2_{calc}: Consiste na somatória do diferencial modular ponderado nas 6 células base da matriz associativa $\sum \frac{(Obs-Esp)^2}{Esp}$:

$$\chi^2_{calc} = \frac{(60 - 73{,}3333)^2}{73{,}3333} + \frac{(40 - 26{,}6667)^2}{26{,}6667} + \cdots + \frac{(15 - 26{,}6667)^2}{26{,}6667}$$
$$\chi^2_{calc} \approx \boxed{16{,}1932}$$

5. Determinar os Graus de Liberdade:

$$gl = (\text{linhas} - 1) \times (\text{colunas} - 1) = (3 - 1) \times (2 - 1) = \boxed{2}$$

6. Comparar e concluir: Para $\alpha = 0{,}05$ estanque com sobrecorte de liberação em escopo $gl = 2$, o limiar fronteiriço impeditivo referencial é extraído como $\chi^2_{critico} = \boxed{5{,}9915}$. Fato notório, como $\chi^2_{calc}\,(16{,}1932) > \chi^2_{critico}\,(5{,}9915)$, o preceito analítico orienta textualmente à exclusão sumária. A decisão ponta consiste em: **Rejeitar** H_0.

Conclusão da Associação: Há evidências estatísticas suficientes, ao nível de 5% de significância, para afirmar que existe associação entre o tipo de plano contratado e o grau de satisfação do cliente.

Questão 23

(Retirado de: Slide "Aula 9 - Alguns Testes Especiais", Exercício 2 - pág. 270)

Um supermercado testou três diferentes embalagens para um mesmo produto e quis verificar se a preferência dos consumidores é influenciada pela embalagem. Os dados coletados mostram a quantidade de pessoas que escolheram cada embalagem: Observado = c(45, 60, 35) para as Embalagens: A, B, C.

Resolução:
Extração de Dados:
Valores observados: $O_A = 45$, $O_B = 60$, $O_C = 35$.
Total de consumidores avaliados: $N = 45 + 60 + 35 = 140$.

(a) Teste a hipótese de que as embalagens são igualmente preferidas (distribuição uniforme).

1. Estabelecer as hipóteses:

$$H_0 : \text{As embalagens são igualmente preferidas } (p_A = p_B = p_C = 1/3).$$
$$H_1 : \text{Pelo menos uma das proporções difere das demais.}$$

2 e 3. Organizar Frequências Esperadas (E_i): Assumindo distribuição uniforme sob H_0:

$$E_A = E_B = E_C = \frac{140}{3} \approx \boxed{46{,}6667}$$

4. Calcular a estatística χ^2_{calc}: Invocando o Teste Qui-Quadrado de Aderência (Goodness-of-Fit), processamos a somatória basilar:

$$\begin{aligned}
\chi^2_{calc} &= \frac{(45 - 46{,}6667)^2}{46{,}6667} + \frac{(60 - 46{,}6667)^2}{46{,}6667} + \frac{(35 - 46{,}6667)^2}{46{,}6667} \\
&= \frac{(-1{,}6667)^2}{46{,}6667} + \frac{(13{,}3333)^2}{46{,}6667} + \frac{(-11{,}6667)^2}{46{,}6667} \\
&= \frac{2{,}7779}{46{,}6667} + \frac{177{,}7769}{46{,}6667} + \frac{136{,}1119}{46{,}6667} \\
&= 0{,}0595 + 3{,}8095 + 2{,}9167 \approx \boxed{6{,}7857}
\end{aligned}$$

(b) Qual é o p-valor? O que você conclui?

5. Determinar Graus de Liberdade e p-valor: $gl = k - 1 = 3 - 1 = \boxed{2}$.
A área à direita de $\chi^2 = 6{,}7857$ com $gl = 2$ fornece o p-valor:

$$\text{p-valor} \approx \boxed{0{,}0336}$$

6 e 7. Conclusão Conjunta: Como $0{,}0336 < 0{,}05$, **rejeitamos** H_0. Há evidências estatísticas de que as embalagens não são igualmente preferidas pelos consumidores.

(c) Se houver diferença, qual embalagem parece ser a preferida?

Inspecionando as frequências observadas ($O_A = 45$, $O_B = 60$, $O_C = 35$), a **Embalagem B** foi a mais escolhida pelos consumidores.

(d) Que recomendações você faria para o supermercado?
Como a Embalagem B obteve a maior frequência de escolhas, recomenda-se ao supermercado priorizar o uso dessa embalagem, pois ela demonstrou maior aceitação entre os consumidores entrevistados.

Questão 24

(Retirado de: Slide "Aula 9 - Alguns Testes Especiais", Exercício 3 - pág. 271)

Um estudo clínico avaliou a eficácia de três tratamentos diferentes (A, B, C) para uma doença. Os pacientes foram classificados como "Curados"ou "Não Curados"após o tratamento. Contagens observadas: A (Curado=30, Não Curado=20); B (Curado=40, Não Curado=10); C (Curado=25, Não Curado=25).

Resolução:
Extração de Dados:
Variável Linha (i): Tipo de Tratamento (A, B, C).
Variável Coluna (j): Resultado (Curado, Não Curado).
Total geral: $N = 150$.

(a) Existe associação entre o tipo de tratamento e a cura?
1. Estabelecer as hipóteses:

$$H_0 : \text{A taxa de cura independe do tipo de tratamento aplicado.}$$
$$H_1 : \text{Existe associação entre o tipo de tratamento e a taxa de cura.}$$

2 e 3. Frequências Esperadas (E_{ij}): Como cada tratamento foi aplicado em 50 pacientes ($T_{linha} = 50$), e os totais de coluna são $T_{Curado} = 95$ e $T_{NoCurado} = 55$:

$$E_{\text{Curado}} = \frac{50 \times 95}{150} \approx 31{,}6667$$
$$E_{\text{Não Curado}} = \frac{50 \times 55}{150} \approx 18{,}3333$$

Como todos os valores esperados são maiores que 5, a condição para uso do Qui-Quadrado está satisfeita.

4. Calcular a estatística χ^2_{calc}:

$$\begin{aligned}
\chi^2_{calc} &= \frac{(30-31{,}6667)^2}{31{,}6667} + \frac{(20-18{,}3333)^2}{18{,}3333} \\
&+ \frac{(40-31{,}6667)^2}{31{,}6667} + \frac{(10-18{,}3333)^2}{18{,}3333} \\
&+ \frac{(25-31{,}6667)^2}{31{,}6667} + \frac{(25-18{,}3333)^2}{18{,}3333} \\
&= 0{,}0881 + 0{,}1521 + 2{,}1881 + 3{,}7812 + 1{,}4070 + 2{,}4306 \\
&\approx \boxed{10{,}0471}
\end{aligned}$$

5. Graus de liberdade e valor crítico:

$$\begin{aligned}
gl &= (3-1)\times(2-1) = \boxed{2} \\
\chi^2_{crit} &= \boxed{5{,}9915} \quad (\alpha = 0{,}05)
\end{aligned}$$

6 e 7. Decisão e conclusão: Como $\chi^2_{calc} = 10{,}0471 > \chi^2_{crit} = 5{,}9915$, **rejeitamos** H_0. Há evidências estatísticas suficientes, ao nível de 5% de significância, para afirmar que existe associação entre o tipo de tratamento e a taxa de cura.

(b) Qual tratamento apresenta a maior taxa de cura?
Com base nas porcentagens calculadas no item (c), o **Tratamento B** apresenta a maior taxa de cura, com 80%.

(c) Calcule as porcentagens de cura por tratamento.

- **Tratamento A:** $\frac{30}{50} = \boxed{60\%}$
- **Tratamento B:** $\frac{40}{50} = \boxed{80\%}$
- **Tratamento C:** $\frac{25}{50} = \boxed{50\%}$

(d) Se um novo paciente chegar, qual tratamento você recomendaria? Por quê?
Recomenda-se o **Tratamento B**, pois apresentou a maior taxa de cura (80%) entre os três tratamentos avaliados, sendo superior ao Tratamento A (60%) e ao Tratamento C (50%).