Estatística II

Licenciatura em Gestão do Desporto

2.º Ano/2.º Semestre

2024/2025

Aulas Teórico-Práticas N.º 12 e 13 (Semana 7)

Docente: Elisabete Fernandes
E-mail: efernandes@iseg.ulisboa.pt

https://doity.com.br/estatistica-aplicada-a-nutricao

https://basiccode.com.br/produto/informatica-basica/

Conteúdos Programáticos

Aulas Teórico-Práticas (Semanas 1 a 3)	Aulas Teórico-Práticas (Semanas 4 a 7)	Aulas Teórico-Práticas (Semanas 7 a 9)	Aulas Teórico-Práticas (Semanas 10 a 13)
• **Capítulo 1**: Revisões e Distribuições de Amostragem	• **Capítulo 2**: Estimação	• **Capítulo 3**: Testes de Hipóteses	• **Capítulo 4**: Modelo de Regressão Linear Múltipla

Material didático: Exercícios do Livro Murteira et al (2015), Formulário e Tabelas Estatísticas

Bibliografia: B. Murteira, C. Silva Ribeiro, J. Andrade e Silva, C. Pimenta e F. Pimenta; *Introdução à Estatística*, 2ª ed., Escolar Editora, 2015.

Intervalo de Confiança para a Diferença de Valores Médios μ_1-μ_2

1

Intervalo de Confiança para μ_1-μ_2: Variâncias Conhecidas

Portanto, quando as populações são Normais com variâncias conhecidas, o **I. C. para $\boldsymbol{\mu_1 - \mu_2}$ com $\mathbf{100(1-\alpha)\%}$ de confiança** é dado por:

$$\left](\overline{X}_1 - \overline{X}_2) - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}};\ (\overline{X}_1 - \overline{X}_2) + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.\right[$$

ProbabilidadesEstatistica_2019 (uevora.pt)

Intervalo de Confiança para μ_1-μ_2: Variâncias Desconhecidas e Iguais

Portanto, quando as populações são Normais com variâncias desconhecidas, mas iguais, o **I. C. para $\mu_1 - \mu_2$ com $100(1-\alpha)\%$ de confiança** é dado por:

$$\left]\left(\bar{X}_1 - \bar{X}_2\right) - t_{n_1+n_2-2\,;\,1-\frac{\alpha}{2}} S^* \,;\, \left(\bar{X}_1 - \bar{X}_2\right) + t_{n_1+n_2-2\,;\,1-\frac{\alpha}{2}} S^*\right[.$$

ProbabilidadesEstatistica_2019 (uevora.pt)

$$S^* = \sqrt{\frac{(n_1-1)S1'^2 + (n_2-1)S2'^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

Variâncias corrigidas

$$S'^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Intervalo de Confiança para μ_1-μ_2: Variâncias Desconhecidas e Diferentes

Portanto, quando as populações são Normais com variâncias desconhecidas, , o **I. C. para $\mu_1 - \mu_2$ com $100(1-\alpha)\%$ de confiança** é dado por:

$$\left]\left(\bar{X}_1 - \bar{X}_2\right) - t_{v;\,1-\frac{\alpha}{2}}\sqrt{\frac{S1'^2}{n_1} + \frac{S2'^2}{n_2}}\,;\left(\bar{X}_1 - \bar{X}_2\right) + t_{v;\,1-\frac{\alpha}{2}}\sqrt{\frac{S1'^2}{n_1} + \frac{S2'^2}{n_2}}\right[.$$

ProbabilidadesEstatistica_2019 (uevora.pt)

Variâncias corrigidas

$$v = \left[\frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{S_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{S_2^2}{n_2}\right)^2}\right].$$

Variâncias corrigidas

$$S'^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Intervalo de Confiança para μ_1-μ_2 e n ≥ 30: Variâncias Conhecidas e Desconhecidas

- **Parâmetro:** $\mu_1 - \mu_2$
  (populações quaisquer independentes com variâncias finitas)

V. F.

Variâncias Conhecidas

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \overset{a}{\sim} N(0,1)$$

ou

Variâncias Desconhecidas

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1'^2}{m} + \frac{S_2'^2}{n}}} \overset{a}{\sim} N(0,1).$$

I. C.

$$(\bar{x}_1 - \bar{x}_2 - z_{\alpha/2}\sigma^*, \bar{x}_1 - \bar{x}_2 + z_{\alpha/2}\sigma^*)$$

ou

$$(\bar{x}_1 - \bar{x}_2 - z_{\alpha/2}s^*, \bar{x}_1 - \bar{x}_2 + z_{\alpha/2}s^*),$$

com $\sigma^* = \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}$ $\qquad$ $s^* = \sqrt{\frac{s_1'^2}{m} + \frac{s_2'^2}{n}}$

e $z_{\alpha/2}$ a verificar $\Phi(z_{\alpha/2}) = 1 - \alpha/2$.

Murteira et al (2015)

IC para μ_1-μ_2: Formulário

Variância corrigida

$$S'^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

- **POPULAÇÕES NORMAIS**

Média	$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\frac{\bar{X}-\mu}{S'/\sqrt{n}} \sim t(n-1)$
Diferença de médias	$\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}} \sim N(0,1)$ Variâncias Conhecidas $T = \frac{\frac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{\sqrt{\frac{1}{m}+\frac{1}{n}}}}{\sqrt{\frac{(m-1)S_1'^2+(n-1)S_2'^2}{m+n-2}}} \sim t(m+n-2)$	$Z = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1'^2}{m}+\frac{S_2'^2}{n}}} \overset{a}{\sim} t(\nu)$ onde ν é o maior inteiro contido em r, $r = \frac{\left(\frac{s_1'^2}{m}+\frac{s_2'^2}{n}\right)^2}{\frac{1}{m-1}\left(\frac{s_1'^2}{m}\right)^2+\frac{1}{n-1}\left(\frac{s_2'^2}{n}\right)^2}$
Variância	$\frac{nS^2}{\sigma^2} = \frac{(n-1)S'^2}{\sigma^2} \sim \chi^2(n-1)$	
Relação de variâncias	$\frac{S_1'^2}{S_2'^2}\frac{\sigma_2^2}{\sigma_1^2} \sim F(m-1,n-1)$	

Variâncias Desconhecidas e Iguais

Variâncias Desconhecidas e Diferentes

IC para μ_1-μ_2: Formulário

- **GRANDES AMOSTRAS**

Caso geral

Média	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \overset{a}{\sim} N(0,1)$	$\dfrac{\bar{X}-\mu}{S'/\sqrt{n}} \overset{a}{\sim} N(0,1)$
Diferença de médias	$\dfrac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}} \overset{a}{\sim} N(0,1)$	$\dfrac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{S_1'^2}{m}+\dfrac{S_2'^2}{n}}} \overset{a}{\sim} N(0,1)$

Variâncias Conhecidas

Variâncias Desconhecidas

Intervalo de Confiança para a Diferença de Valores Médios μ_1-μ_2: Exercícios

2

Exercício Suplementar que não consta do livro Murteira et al (2015)

Havendo indícios de que o esquema de avaliação e as classificações finais atribuídas diferem fortemente entre duas escolas, decidiu-se comprovar estatisticamente esta hipótese. Os desvios-padrão são conhecidos sendo 2,1 valores na escola A e 1,8 valores na escola B. Assim, retirou-se uma amostra de testes de alunos em cada uma das escolas que levaram aos seguintes resultados:

Escola	n_i	$\bar{x}_i$
A	41	12,9
B	31	14,7

Recorrendo a um intervalo de confiança a 90%, diga se há diferenças entre as classificações médias das escolas A e B. Justifique.

ProbabilidadesEstatistica_2019 (uevora.pt)

Exercício: IC para μ_1-μ_2 (Variâncias Conhecidas)

Sejam:

- X_1 a v.a. que representa a classificação final dos alunos na escola A,
- X_2 a v.a. que representa a classificação final dos alunos na escola B,

com $\sigma_1 = 2{,}1$ e $\sigma_2 = 1{,}8$.

Como as amostras são grandes, o I. C. para $\mu_1 - \mu_2$ a 90% é dado por:

$$\left](\overline{X}_1 - \overline{X}_2) - z_{0,95}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}};(\overline{X}_1 - \overline{X}_2) + z_{0,95}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right[.$$

Substituindo pelos valores conhecidos, sendo $z_{0,95} = 1{,}645$, obtém-se

$$\left](12{,}9 - 14{,}7) \pm 1{,}645\sqrt{\frac{2,1^2}{41} + \frac{1,8^2}{31}}\right[=]-2{,}558;\ -1{,}043[.$$

ProbabilidadesEstatistica_2019 (uevora.pt)

Com 90% de confiança, existe evidência de que as classificações médias são diferentes (0 não está no intervalo). Como ambos os limites do intervalo são negativos então significa que $\mu_1 < \mu_2$, ou seja, a classificação média é superior na escola B do que na escola A[†].

Exercício Suplementar que não consta do livro Murteira et al (2015)

Um determinado método de análise permite determinar o conteúdo de enxofre no petróleo bruto. Os ensaios efetuados em 10 e 8 amostras aleatórias de 1 kg de petróleo bruto, provenientes de furos pertencentes respetivamente aos campos A e B, revelaram os seguintes resultados (em gramas):

Campo A:	111	114	105	112	107	109	112	110	110	106
Campo B:	109	103	101	105	106	108	110	104		

Construa um intervalo de confiança a 90% para a diferença entre os valores esperados da quantidade de enxofre por quilograma de petróleo proveniente de cada campo, considerando que populações são Normais, com variâncias desconhecidas mas iguais.

ProbabilidadesEstatistica_2019 (uevora.pt)

Exercício: IC para μ_1-μ_2 (Variâncias Desconhecidas e Iguais)

Sejam:

- X_1 a v.a. que representa a quantidade de enxofre por quilograma de petróleo do campo A,
- X_2 a v.a. que representa a quantidade de enxofre por quilograma de petróleo do campo B,

com $X_1 \sim N(\mu_1 = ?\,; \sigma_1 = ?)$ e $X_2 \sim N(\mu_2 = ?; \sigma_2 = ?)$, mas $\sigma_1 = \sigma_2$.

$n_1 = 10, \quad \bar{x}_1 = 109{,}6 \quad \text{e} \quad s_1 = 2{,}875,$

$n_2 = 8, \quad \bar{x}_2 = 105{,}75 \quad \text{e} \quad s_2 = 3{,}105.$

O I. C. para $\mu_1 - \mu_2$ a 90% é dado por:

$$\left]\left(\bar{X}_1 - \bar{X}_2\right) - t_{n_1+n_2-2;\,0{,}95} S^* ; \left(\bar{X}_1 - \bar{X}_2\right) + t_{n_1+n_2-2;\,0{,}95} S^* \right[,$$

$$\text{com } S^* = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}.$$

ProbabilidadesEstatistica_2019 (uevora.pt)

Variâncias corrigidas

$$S'^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

Exercício: IC para μ_1-μ_2 (Variâncias Desconhecidas e Iguais)

Substituindo pelos valores conhecidos,

$$s^* = \sqrt{\frac{(10-1)2{,}875^2 + (8-1)3{,}105^2}{10+8-2}}\sqrt{\frac{1}{10}+\frac{1}{8}} = 1{,}413$$

e como $t_{16\,;\,0{,}95} = 1{,}746$, obtém-se

$$](109{,}6 - 105{,}75) \pm 1{,}746 \times 1{,}413; [\ = \ ]1{,}384;\ 6{,}316[.$$

Com 90% de confiança, existe evidência de que o teor médio de enxofre nos campos A e B é diferente (0 não está no intervalo). Uma vez que ambos os limites são positivos, então significa que $\mu_1 > \mu_2$, ou seja, o conteúdo médio de enxofre por quilograma de petróleo extraído do campo A é superior ao registado no campo B[†].

ProbabilidadesEstatistica_2019 (uevora.pt)

Exercício Suplementar que não consta do livro Murteira et al (2015)

Para um estudo sobre a caracterização da altura da população portuguesa, foi recolhida uma amostra de 1861 pessoas, com as seguintes características:

Group Statistics

	Sexo	N	Mean	Std. Deviation
Altura	Masculino	853	168,46	7,617
	Feminino	1007	158,48	6,652

Supondo a normalidade das distribuições e assumindo que as variâncias populacionais são desconhecidas e diferentes, verifique se se pode considerar que as alturas médias dos homens e das mulheres são iguais, com 95% de confiança.

ProbabilidadesEstatistica_2019 (uevora.pt)

Exercício: IC para μ_1-μ_2 (Variâncias Desconhecidas e Diferentes)

Sejam:

- X_1 a v.a. que representa a altura dos indivíduos do sexo masculino,
- X_2 a v.a. que representa a altura dos indivíduos do sexo feminino,

com $X_1 \sim N(\mu_1 = ?\,;\sigma_1 = ?)$ e $X_2 \sim N(\mu_2 = ?; \sigma_2 = ?)$, mas $\sigma_1^2 \neq \sigma_2^2$.

$n_1 = 853, \quad \bar{x}_1 = 168{,}46 \quad$ e $\quad s_1 = 7{,}617,$
$n_2 = 1007, \quad \bar{x}_2 = 158{,}48 \quad$ e $\quad s_2 = 6{,}652.$

O I. C. para $\mu_1 - \mu_2$ a 95% é dado por:

$$\left]\left(\bar{X}_1 - \bar{X}_2\right) - t_{v;\,0{,}975}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}\,;\left(\bar{X}_1 - \bar{X}_2\right) + t_{v;\,0{,}975}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}\right[$$

Variâncias corrigidas

$$S'^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

com $v = \left[\dfrac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{1}{n_1 - 1}\left(\dfrac{S_1^2}{n_1}\right)^2 + \dfrac{1}{n_2 - 1}\left(\dfrac{S_2^2}{n_2}\right)^2}\right].$

ProbabilidadesEstatistica_2019 (uevora.pt)

Exercício: IC para μ_1-μ_2 (Variâncias Desconhecidas e Diferentes)

Substituindo pelos valores conhecidos,

$$v = \left[\frac{\left(\frac{7{,}617^2}{853} + \frac{6{,}652^2}{1007}\right)^2}{\frac{1}{853-1}\left(\frac{7{,}617^2}{853}\right)^2 + \frac{1}{1007-1}\left(\frac{6{,}652^2}{1007}\right)^2} \right] = [1705{,}6] = 1705,$$

e como $t_{1705;\, 0{,}975} = 1{,}96$, obtém-se

$$\left](168{,}46 - 158{,}48) \pm 1{,}96\sqrt{\frac{7{,}617^2}{853} + \frac{6{,}652^2}{1007}}\right[=]9{,}32;\ 10{,}64[.$$

Com 90% de confiança, existe diferença significativa entre as médias das alturas dos homens e das mulheres (0 não está contido do I. C. a 95%). Como ambos os limites do intervalo são positivos então significa que $\mu_H > \mu_M$, ou seja, a altura média dos homens é superior à altura média das mulheres.

ProbabilidadesEstatistica_2019 (uevora.pt)

Murteira et al (2015)
Capítulo 7

37. Para comparar dois métodos de ensino, dividiu-se aleatoriamente uma turma de 22 alunos em dois grupos iguais. Cada grupo foi ensinado com um método diferente (método A e método B) e, no fim da formação, foram sujeitos à mesma prova de avaliação. Os resultados da avaliação, numa escala de 0 a 100, foram os seguintes:

$$\text{Método } A\text{: } \bar{x}_A = 74.8\text{; } s_A'^2 = 81.5\text{.}$$
$$\text{Método } B\text{: } \bar{x}_B = 72.1\text{; } s_B'^2 = 110.5\text{.}$$

Admita que as classificações obtidas em cada grupo têm distribuição normal.

a) Construa um intervalo de confiança a 95% para o quociente das variâncias dos dois grupos.

b) Tendo em conta o resultado da alínea anterior, construa um intervalo de confiança a 99% para a diferença de médias. Interprete o resultado obtido.

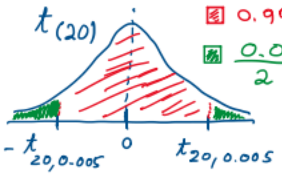
Nota: Vamos considerar que às variâncias são desconhecidas e iguais.

Exercício 37 b)

b) População normal, variâncias desconhecidas mas iguais. Queremos IC 99% para $\mu_A - \mu_B$, logo $\alpha = 0.01$ e $\alpha/2 = 0.005$.

Variável fulcral: $$Z = \frac{\bar{X}_A - \bar{X}_B - (\mu_A - \mu_B)}{\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}\sqrt{\frac{(n_A-1)S_1'^2 + (n_B-1)S_2'^2}{n_A + n_B - 2}}}$$

$$Z \sim t(n_A + n_B - 2) = t(20)$$



0.99

$\frac{0.01}{2} = 0.005$

tabela 7 $\rightarrow$ $t_{20,\,0.005} = 2.845$

Exercício 37 b)

$$P\left(-t_{20,0.005} < Z < t_{20,0.005}\right) = 0.99 \iff$$

$$\iff P\left(-2.845 < \frac{\bar{X}_A - \bar{X}_B - (\mu_A - \mu_B)}{\sqrt{\frac{1}{n_A}+\frac{1}{n_B}}\sqrt{\frac{(n_A-1)S_1'^2+(n_B-1)S_2'^2}{n_A+n_B-2}}} < 2.845\right) = 0.99 \iff$$

$$\iff \dots \iff$$

$$\iff P\left(\bar{X}_A - \bar{X}_B - 2.845\sqrt{\frac{1}{n_A}+\frac{1}{n_B}}\sqrt{\frac{(n_A-1)S_1'^2+(n_B-1)S_2'^2}{n_A+n_B-2}} < \mu_A - \mu_B < \bar{X}_A - \bar{X}_B + 2.845\sqrt{\frac{1}{n_A}+\frac{1}{n_B}}\sqrt{\frac{(n_A-1)S_1'^2+(n_B-1)S_2'^2}{n_A+n_B-2}}\right) = 0.99$$

$$I.C._{99\%}(\mu_A - \mu_B) = \bar{x}_A - \bar{x}_B \pm t_{20,0.005}\sqrt{\frac{1}{n_A}+\frac{1}{n_B}}\sqrt{\frac{(n_A-1)s_A'^2+(n_B-1)s_2'^2}{n_A+n_B-2}}$$

$$= 74.8 - 72.1 \pm 2.845\sqrt{\frac{2}{11}}\sqrt{\frac{10\times 81.5+10\times 110.5}{20}} =$$

$$= (-9.186, 14.586)$$

Uma vez que "0" pertence ao IC, a informação presente na amostra não nos leva a crer (com este nível de confiança) que um dos métodos tenha resultados superiores ao outro. Com base neste IC e a este nível de confiança não se rejeita a hipótese de igualdade de médias.

Murteira et al (2015)
Capítulo 7

46. Com o objectivo de comparar a autonomia (duração da carga da bateria, em horas) de dois modelos de telemóveis, recolheu-se a seguinte informação:

	N.º de aparelhos	Duração da carga da bateria	
Modelo	**Observados**	**Média**	**Desvio padrão corrigido**
A	40	6.5	3.0
B	50	5.5	2.5

Com base num intervalo de confiança a 95% poder-se-á afirmar que não existem diferenças significativas entre os modelos testados, no que respeita à duração média da carga da bateria.

Exercício 46

$X_A \equiv$ autonomia do telemóvel A (em horas).

$X_B \equiv$ " " " B (em horas).

$n_A = 40 \quad \bar{x}_A = 6.5 \quad s'_A = 3$

$n_B = 50 \quad \bar{x}_B = 5.5 \quad s'_B = 2.5$

I.C.$_{95\%}$ para $\mu_A - \mu_B$

$1-\alpha = 0.95 \Leftrightarrow \alpha = 0.05 \Leftrightarrow \frac{\alpha}{2} = 0.025$

Duas populações desconhecidas: amostras grandes (caso geral)

Variável fulcral: $Z = \dfrac{\bar{X}_A - \bar{X}_B - (\mu_A - \mu_B)}{\sqrt{\dfrac{S'^2_A}{n_A} + \dfrac{S'^2_B}{n_B}}} \overset{a}{\sim} N(0,1)$

Exercício 46

$$P\left(-z_{0.025} < Z < z_{0.025}\right) = 0.95 =$$

$$= P\left(-z_{0.025} < \frac{\bar{X}_A - \bar{X}_B - (\mu_A - \mu_B)}{\sqrt{\frac{S_A'^2}{n_A} + \frac{S_B'^2}{n_B}}} < z_{0.025}\right) =$$

$$= P\left(\bar{X}_A - \bar{X}_B - z_{0.025}\sqrt{\frac{S_A'^2}{n_A} + \frac{S_B'^2}{n_B}} < \mu_A - \mu_B < \bar{X}_A - \bar{X}_B + z_{0.025}\sqrt{\frac{S_A'^2}{n_A} + \frac{S_B'^2}{n_B}}\right)$$

Exercício 46

$$I.A._{95\%}: \bar{X}_A - \bar{X}_B \pm z_{0.025}\sqrt{\frac{S_A'^2}{n_A} + \frac{S_B'^2}{n_B}}$$

$$I.C._{95\%}: \bar{x}_A - \bar{x}_B \pm z_{0.025}\sqrt{\frac{s_A'^2}{n_A} + \frac{s_B'^2}{n_B}} =$$

$$= 6.5 - 5.5 \pm 1.96\sqrt{\frac{3^2}{40} + \frac{2.5^2}{50}} =$$

$$= (-0.1596, 2.1596)$$

Resposta: Sim, porque o valor $\mu_A - \mu_B = 0$ pertence ao $I.C._{95\%}(\mu_A - \mu_B)$.

Murteira et al (2015)
Capítulo 7

49. Com o objectivo de avaliar o efeito da dimensão da turma sobre o aproveitamento dos alunos recolheu-se a seguinte informação (amostras casuais de alunos da mesma disciplina):

Classificação (0-20)	**< 8**	**[8, 10)**	**[10, 12)**	**[12, 14)**	**≥ 14**	**Total**
Turmas grandes	12	42	30	11	5	100
Turmas pequenas	4	30	40	21	5	100

a) Construa um intervalo de confiança a 90% para a nota média dos alunos de turmas pequenas.

b) Construa um intervalo de confiança a 90% para a diferença na classificação média das duas populações e comente o resultado obtido.

c) Com base num intervalo de confiança a 90%, diga se pode afirmar que a proporção de alunos com nota positiva é maior nas turmas pequenas.

Exercício 49 a)

Ponto médio	4	9	11	13	17	
Classificação (0–20)	< 8	[8, 10)	[10, 12)	[12, 14)	≥ 14	Total
Turmas grandes	12	42	30	11	5	100 = n_G
Turmas pequenas	4	30	40	21	5	100 = n_p

$X_p \equiv$ Nota de um aluno (de turma pequena)

$X_G \equiv$ 〃 〃 〃 〃 (〃 〃 grande)

a) $\bar{x}_p = \dfrac{(4\times4)+(9\times30)+(11\times40)+(13\times21)+(17\times5)}{100} = 10.84$

$$s_p^2 = \frac{\sum_{i=1}^{100} x_{p_i}^2}{100} - \bar{x}^2 = \frac{(4^2\times4)+(9^2\times30)+(11^2\times40)+(13^2\times21)+(17^2\times5)}{100} - 10.84^2$$

$$= 123.28 - 10.84^2 = 5.7744$$

$s_p'^2 = \dfrac{n}{n-1} s_p'^2 \qquad s_p' = \sqrt{\dfrac{n_p}{n_p-1}}\sqrt{s_p^2} \quad \sqrt{\dfrac{100}{99}}\sqrt{5.7744} \simeq 2.4151$

Exercício 49 a)

Pop. desconhecida (média): $\dfrac{\bar{X}_P - \mu_P}{S'_P / \sqrt{n_P}} \sim N(0,1)$ $z_{0.05} = 1.645$

$$IC_{90\%}(\mu_P) = \bar{x}_P \pm z_{0.05} \frac{s'}{\sqrt{n}} = 10.84 \pm 1.645 \times \frac{2.4151}{\sqrt{100}} =$$

$$= (10.443 ; 11.237)$$

Exercício 49 b)

b) Duas populações desconhecidas (diferença de médias):

$$\frac{\bar{X}_G - \bar{X}_P - (\mu_G - \mu_P)}{\sqrt{\frac{S'^2_G}{n_G} + \frac{S'^2_P}{n_P}}} \stackrel{a}{\sim} N(0,1)$$

$$\bar{x}_G = \frac{(4\times 12) + (9\times 42) + (11\times 30) + (13\times 11) + (17\times 5)}{100} = 9.84$$

$$s_G^2 = \frac{\sum_{i=1}^{100} x_{G_i}^2}{100} - \bar{x}_G^2 =$$

$$= \frac{(4^2\times 12) + (9^2\times 42) + (11^2\times 30) + (13^2\times 11) + 17^2\times 5}{100} - 9.84^2 =$$

$$= 105.28 - 9.84^2 = 8.4544$$

$$s'^2_G = \frac{n_G}{n_G - 1} s_G^2 = \frac{100}{99} \times 8.4544 \qquad s'_G = \sqrt{\frac{100}{99} \times 8.4544} \simeq 2.922$$

Exercício 49 b)

$$IC_{90\%}(\mu_G - \mu_P) = \bar{x}_G - \bar{x}_P \pm z_{0.05}\sqrt{\frac{s'^2_G}{n_G} + \frac{s'^2_P}{n_P}}$$

$$= 9.84 - 10.84 \pm 1.645\sqrt{\frac{2.922^2}{100} + \frac{2.4151^2}{100}} =$$

$$= (-1.6236; -0.3764)$$

$$\mu_G - \mu_P = 0 \notin IC_{90\%}(\mu_G - \mu_P)$$

Logo, com base neste IC, não decidimos favoravelmente à igualdade de classificações médias em turmas de tamanhos diferentes.

Como todos o valores de μ_g - μ_p pertencentes ao IC são negativos conclui-se que a classificação média deve ser superior nas turmas pequenas $\mu_p > \mu_g$.

Intervalo de Confiança para a Variância σ^2

3

Intervalo de Confiança para σ^2

Intervalo a $(1-\alpha)\times 100\%$ de confiança para σ^2 numa população normal

$$\left]\frac{(n-1)\,S'^2}{\chi^2_{n-1\,;\,1-\frac{\alpha}{2}}}\,;\,\frac{(n-1)\,S'^2}{\chi^2_{n-1\,;\,\frac{\alpha}{2}}}\right[.$$

$$\frac{(n-1)\,S'^2}{\chi^2_{1-\alpha/2,(n-1)}} < \sigma^2 < \frac{(n-1)\,S'^2}{\chi^2_{\alpha/2,(n-1)}}$$

Variância corrigida

$$S'^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

[moduloI_aula5.pdf]

Intervalo de Confiança para σ^2

Para obter o Intervalo de Confiança para σ^2

Variável usada	Condições	Distribuição
$\frac{(n-1)\ S'^2}{\sigma^2}$ $S'^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$	$X_i \frown N(\mu, \sigma)$ $i = 1, 2, \cdots, n$	$\chi^2_{(n-1)}$

Definição da distribuição χ^2

Se $Z_1, Z_2, \cdots, Z_n$ são v.'s a.'s $N(0, 1)$ independentes

$$\Downarrow$$

a v.a. $X = Z_1^2 + \cdots + Z_n^2$ é tal que $X \frown \chi^2_{(n)}$

Tem-se $E[X] = n$; $Var[X] = 2n$

moduloI_aula5.pdf

IC para σ²: Formulário

- **POPULAÇÕES NORMAIS**

Média	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\dfrac{\bar{X}-\mu}{S'/\sqrt{n}} \sim t(n-1)$
Diferença de médias	$\dfrac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}} \sim N(0,1)$ $T=\dfrac{\dfrac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}}{\sqrt{\dfrac{(m-1)S_1'^2+(n-1)S_2'^2}{m+n-2}}} \sim t(m+n-2)$	$Z=\dfrac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{S_1'^2}{m}+\dfrac{S_2'^2}{n}}} \overset{a}{\sim} t(\nu)$ onde ν é o maior inteiro contido em r, $r=\dfrac{\left(\dfrac{s_1'^2}{m}+\dfrac{s_2'^2}{n}\right)^2}{\dfrac{1}{m-1}\left(\dfrac{s_1'^2}{m}\right)^2+\dfrac{1}{n-1}\left(\dfrac{s_2'^2}{n}\right)^2}$
Variância	$\dfrac{nS^2}{\sigma^2}=\dfrac{(n-1)S'^2}{\sigma^2} \sim \chi^2(n-1)$	Variância corrigida $S'^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$
Relação de variâncias	$\dfrac{S_1'^2}{S_2'^2}\dfrac{\sigma_2^2}{\sigma_1^2} \sim F(m-1,n-1)$	

Intervalo de Confiança para a Variância σ^2: Exercícios

4

Exercício Suplementar que não consta do livro Murteira et al (2015)

7.2 Admita que a densidade de construção, X, num projecto de urbanização tem distribuição normal. Uma amostra aleatória de 50 lotes desse projecto conduziu a

n

$$\sum_{i=1}^{50} x_i = 227.2\,;\ \sum_{i=1}^{50} x_i^2 = 2242.6$$

~~Assumindo que o desvio padrão de X é igual a~~ 4, construa um intervalo de confiança a 95% para a densidade média de construção. Que dimensão deveria ter a amostra para que a amplitude desse intervalo fosse reduzida a metade?

Exercício 7.2: Admitindo que se pretende construir um intervalo de confiança a 90% p/ σ^2.

Exercício 7.2: IC para σ^2

<u>Passo 0</u> : $X \sim N(\mu, \sigma^2)$ $\quad \mu, \sigma^2 = ?$

$I.C._{0,90}(\sigma^2)$

<u>Passo 1</u> : v. fulcral

$$\frac{49\, S^2}{\sigma^2} \sim \chi^2_{(49)}$$

Slides Professora Claúdia Nunes

Cálculo dos Quantis da Distribuição Qui-Quadrado de Probabilidade 1-α/2 e α/2 com n-1 g.l.´s

Nível de confiança (1-α=0,90)

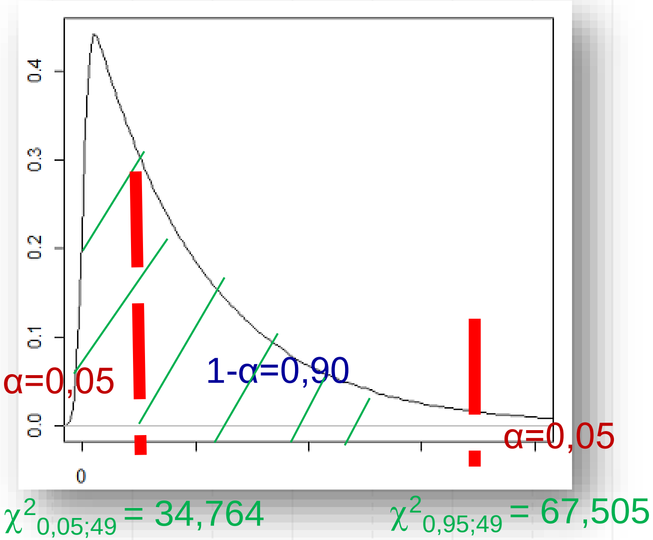
Nível de significância (α=0,10)

Área total é igual a 1

O nível de confiança é 1-α = 0,90 ⇔ α = 0,10, então tem-se 1-α/2 = 0,95 e α/2 = 0,05

Logo, pretende-se calcular o quantil da distribuição Qui-Quadrado de probabilidade 0,05
$\chi^2_{0,05;49} = \chi^{2*}_{0,95;49}$ = 34,764 (ver tabela a seguir)

Logo, pretende-se calcular o quantil da distribuição Qui-Quadrado de probabilidade 0,95
$\chi^2_{0,95;49} = \chi^2_{0,05;49}$= 67,505 (ver tabela a seguir)



Cálculo dos Quantis da Distribuição Qui-Quadrado de Probabilidade 1-α/2 e α/2 com n-1 g.l.´s

Tabela do Qui-Quadrado

$$\chi^2_{n.\varepsilon} : P(X > \chi^2_{n.\varepsilon}) = \varepsilon$$

n \ ε	.995	.990	.975	.950	.900	.750	.500	.250	.100	.050	.025	.010	.005	.001
1	.000	.000	.001	.004	.016	.102	.455	1.323	2.706	3.841	5.024	6.635	7.879	10.827
2	.010	.020	.051	.103	.211	.575	1.386	2.773	4.605	5.991	7.378	9.210	10.597	13.815
3	.072	.115	.216	.352	.584	1.213	2.366	4.108	6.251	7.815	9.348	11.345	12.838	16.266
4	.207	.297	.484	.711	1.064	1.923	3.357	5.385	7.779	9.488	11.143	13.277	14.860	18.466
5	.412	.554	.831	1.145	1.610	2.675	4.351	6.626	9.236	11.070	12.832	15.086	16.750	20.515
6	.676	.872	1.237	1.635	2.204	3.455	5.348	7.841	10.645	12.592	14.449	16.812	18.548	22.457
7	.989	1.239	1.690	2.167	2.833	4.255	6.346	9.037	12.017	14.067	16.013	18.475	20.278	24.321
8	1.344	1.647	2.180	2.733	3.490	5.071	7.344	10.219	13.362	15.507	17.535	20.090	21.955	26.124
9	1.735	2.088	2.700	3.325	4.168	5.899	8.343	11.389	14.684	16.919	19.023	21.666	23.589	27.877
10	2.156	2.558	3.247	3.940	4.865	6.737	9.342	12.549	15.987	18.307	20.483	23.209	25.188	29.588
11	2.603	3.053	3.816	4.575	5.578	7.584	10.341	13.701	17.275	19.675	21.920	24.725	26.757	31.264
12	3.074	3.571	4.404	5.226	6.304	8.438	11.340	14.845	18.549	21.026	23.337	26.217	28.300	32.909
13	3.565	4.107	5.009	5.892	7.041	9.299	12.340	15.984	19.812	22.362	24.736	27.688	29.819	34.527
14	4.075	4.660	5.629	6.571	7.790	10.165	13.339	17.117	21.064	23.685	26.119	29.141	31.319	36.124
15	4.601	5.229	6.262	7.261	8.547	11.037	14.339	18.245	22.307	24.996	27.488	30.578	32.801	37.698
16	5.142	5.812	6.908	7.962	9.312	11.912	15.338	19.369	23.542	26.296	28.845	32.000	34.267	39.252
17	5.697	6.408	7.564	8.672	10.085	12.792	16.338	20.489	24.769	27.587	30.191	33.409	35.718	40.791
18	6.265	7.015	8.231	9.390	10.865	13.675	17.338	21.605	25.989	28.869	31.526	34.805	37.156	42.312
19	6.844	7.633	8.907	10.117	11.651	14.562	18.338	22.718	27.204	30.144	32.852	36.191	38.582	43.819
20	7.434	8.260	9.591	10.851	12.443	15.452	19.337	23.828	28.412	31.410	34.170	37.566	39.997	45.314
21	8.034	8.897	10.283	11.591	13.240	16.344	20.337	24.935	29.615	32.671	35.479	38.932	41.401	46.796
22	8.643	9.542	10.982	12.338	14.041	17.240	21.337	26.039	30.813	33.924	36.781	40.289	42.796	48.268
23	9.260	10.196	11.689	13.091	14.848	18.137	22.337	27.141	32.007	35.172	38.076	41.638	44.181	49.728
24	9.886	10.856	12.401	13.848	15.659	19.037	23.337	28.241	33.196	36.415	39.364	42.980	45.558	51.179
25	10.520	11.524	13.120	14.611	16.473	19.939	24.337	29.339	34.382	37.652	40.646	44.314	46.928	52.619
26	11.160	12.198	13.844	15.379	17.292	20.843	25.336	30.435	35.563	38.885	41.923	45.642	48.290	54.051
27	11.808	12.878	14.573	16.151	18.114	21.749	26.336	31.528	36.741	40.113	43.195	46.963	49.645	55.475
28	12.461	13.565	15.308	16.928	18.939	22.657	27.336	32.620	37.916	41.337	44.461	48.278	50.994	56.892
29	13.121	14.256	16.047	17.708	19.768	23.567	28.336	33.711	39.087	42.557	45.722	49.588	52.335	58.301
30	13.787	14.953	16.791	18.493	20.599	24.478	29.336	34.800	40.256	43.773	46.979	50.892	53.672	59.702
40	20.707	22.164	24.433	26.509	29.051	33.660	39.335	45.616	51.805	55.758	59.342	63.691	66.766	73.403
50	27.991	29.707	32.357	34.764	37.689	42.942	49.335	56.334	63.167	67.505	71.420	76.154	79.490	86.660
60	35.534	37.485	40.482	43.188	46.459	52.294	59.335	66.981	74.397	79.082	83.298	88.379	91.952	99.608
70	43.275	45.442	48.758	51.739	55.329	61.698	69.334	77.577	85.527	90.531	95.023	100.425	104.215	112.317
80	51.172	53.540	57.153	60.391	64.278	71.145	79.334	88.130	96.578	101.879	106.629	112.329	116.321	124.839
90	59.196	61.754	65.647	69.126	73.291	80.625	89.334	98.650	107.565	113.145	118.136	124.116	128.299	137.208
100	67.328	70.065	74.222	77.929	82.358	90.133	99.334	109.141	118.498	124.342	129.561	135.807	140.170	149.449

$\chi^2_{0,05;49} = \chi^{2*}_{0,95;49} = 34{,}764$

$\chi^2_{0,95;49} = \chi^{2*}_{0,05;49} = 67{,}505$

É a variância amostral corrigida.

NOTA: Saber a variância da amostra não significa que se saiba a variância da população!!

$$S'^2 = \frac{1}{50-1}\, 2242.6 - \frac{50}{50-1}\left(\frac{227.2}{50}\right)^2 = 24.698$$

$$\left]\frac{(n-1)\,S'^2}{\chi^2_{n-1;1-\frac{\alpha}{2}}};\frac{(n-1)\,S'^2}{\chi^2_{n-1;\frac{\alpha}{2}}}\right[.$$

Limite inferior = (50-1)*24,698/67,505 = 7,928

Limite superior = (50-1)*24,698/34,764 = 34,811

$IC_{90\%}(\sigma^2)$: (7,928; 34,811)

Exercício 7.2: IC para σ^2

NOTA:

1. v. fulcral: $\frac{49 s^2}{\sigma^2} \sim \chi^2_{(49)}$

   v. fulcral $\frac{49 \times 24.698}{\sigma^2} \sim \chi^2_{(49)}$

2. Interpretação: $I.C.A_{1-\alpha}(\theta) =]\, L(x_1, \ldots, x_n);\ U(x_1, \ldots, x_n)$

   significa que se recolhermos muitas amostras e calcularmos, p/ cada uma delas, o correspondente intervalo, então em cerca de $(1-\alpha)100\%$, os intervalos assim obtidos contêm o verdadeiro valor do parâmetro desconhecido θ.

3. A amplitude do intervalo é tanto menor, em geral, quanto maior for a dimensão da amostra.

Murteira et al (2015)
Capítulo 7

30. Com base numa amostra casual com 16 observações, retirada de uma população normal, construiu-se, segundo o processo habitual, o seguinte intervalo de confiança para o valor esperado: (7.398, 12.602).

a) Sabendo que, com a informação da amostra, se obteve $s = 3.872$, qual o grau de confiança que pode atribuir ao intervalo atrás referido?

b) Com base na mesma amostra, construa um intervalo de confiança a 95% para a variância da população.

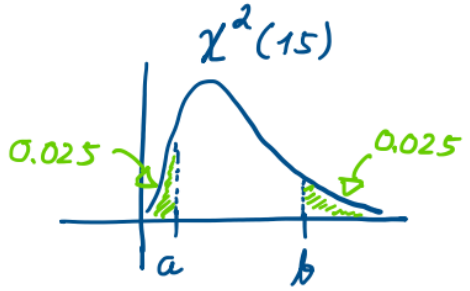
c) Suponha que a verdadeira variância é 36. Se pretender construir um intervalo de confiança a 95% para a média, de modo que a respectiva amplitude não exceda 6.5, qual é a dimensão mínima da amostra a recolher?

Nesta alínea não se considerou a variância amostral, mas neste caso é igual como se verifica por esta igualdade:

$$\frac{nS^2}{\sigma^2} = \frac{(n-1)S'^2}{\sigma^2}$$

Exercício 30 b)

b) $\dfrac{n s^2}{\sigma^2} \sim \chi^2(n-1) = \chi^2(15)$



$$P\left(a < \frac{n s^2}{\sigma^2} < b\right) = 0.95 \iff$$

$$\iff P\left(\frac{n s^2}{a} > \sigma^2 > \frac{n s^2}{b}\right) = 0.95 \iff$$

$$\iff P\left(\frac{n s^2}{b} < \sigma^2 < \frac{n s^2}{a}\right) = 0.95$$

$a = \chi^2_{15,\,0.975} = 6.262$

$b = \chi^2_{15,\,0.025} = 27.488$

Exercício 30 b)

$$\left]\frac{(n-1)\,S'^2}{\chi^2_{n-1\,;\,1-\frac{\alpha}{2}}}\,;\,\frac{(n-1)\,S'^2}{\chi^2_{n-1\,;\,\frac{\alpha}{2}}}\right[.$$

Nota: Podemos considerar n*variância não corrigida ou (n-1)*variância corrigida… O resultado é o mesmo…

$$IC_{95\%}(\sigma^2) = \left(\frac{n s^2}{b}\,;\,\frac{n s^2}{a}\right) =$$

$$= \left(\frac{16 \times 3.872^2}{27.488}\,;\,\frac{16 \times 3.872^2}{6.262}\right) =$$

$$= (8.72665\,;\,38.30695)$$

Murteira et al (2015)
Capítulo 7

31. Considere uma população com distribuição normal de parâmetros desconhecidos. Dessa população foi retirada uma amostra casual de dimensão 25. Suponha que a amostra forneceu os seguintes resultados:

$$\sum_{i=1}^{25} x_i = 75 \text{ e } \sum_{i=1}^{25} x_i^2 = 321.$$

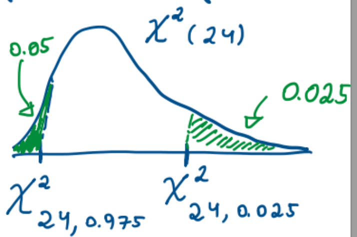
a) Construa um intervalo de confiança a 95% para a média.
b) Construa um intervalo de confiança a 95% para o desvio padrão.

Exercício 31 b)

b) I.C. a 95% para σ^2

$1 - \alpha = 0.95 \qquad \alpha = 0.05 \qquad \alpha/2 = 0.025$

Variável fulcral: $Q = \dfrac{(n-1)S'^2}{\sigma^2} \sim \chi^2_{(n-1)} = \chi^2_{(24)}$



$P\left(\chi^2_{24,\,0.975} < Q < \chi^2_{24,\,0.025}\right) = 0.95 \Leftrightarrow$

$\Leftrightarrow P\left(12.401 < Q < 39.364\right) = 0.95 \Leftrightarrow$

$\Leftrightarrow P\left(12.401 < \dfrac{(n-1)S'^2}{\sigma^2} < 39.364\right) = 0.95 \Leftrightarrow$

$\Leftrightarrow P\left(\dfrac{1}{12.401} > \dfrac{\sigma^2}{(n-1)S'^2} > \dfrac{1}{39.364}\right) = 0.95 \Leftrightarrow$

$\Leftrightarrow P\left(\dfrac{(n-1)S'^2}{39.364} < \sigma^2 < \dfrac{(n-1)S'^2}{12.401}\right) = 0.95$

Exercício 31 b)

$$\text{I.A.}: \left[\frac{24\, S'^2}{39.364} ; \frac{24\, S'^2}{12.401} \right]$$

$$\text{I.C.}: \left[\frac{24\, s'^2}{39.364} ; \frac{24\, s'^2}{12.401} \right] = \qquad (s'^2 = 4)$$

$$= \left[\frac{24 \times 4}{39.364} ; \frac{24 \times 4}{12.401} \right] = [2.439, 7.741]$$

Nota: exercício pede I.C. para σ mas as soluções apresentam o resultado aqui obtido, ou seja, o I.C. para σ^2.

Se quisermos o IC para σ em vez de σ^2 basta fazer $IC = (\sqrt{2.439}, \sqrt{7.741})$.

Intervalo de Confiança para a Relação de Variâncias σ_1^2/σ_2^2

5

Intervalo de Confiança para σ_1^2/σ_2^2

Quando as populações são Normais, o **I. C. para** $\frac{\sigma_1^2}{\sigma_2^2}$ **com** $100(1-\alpha)\%$ **de confiança** é dado por:

$$\left]\frac{S_1^2}{S_2^2}\frac{1}{F_{n_1-1,n_2-1,1-\frac{\alpha}{2}}};\frac{S_1^2}{S_2^2}\frac{1}{F_{n_1-1,n_2-1,\frac{\alpha}{2}}}\right[.$$

Variável Fulcral

$$\frac{S_1'^2}{S_2'^2}\frac{\sigma_2^2}{\sigma_1^2} \sim F(m-1,n-1)$$

Como $F_{n_1-1,n_2-1;\frac{\alpha}{2}} = \frac{1}{F_{n_2-1,n_1-1;1-\frac{\alpha}{2}}}$ este intervalo pode ser escrito da seguinte forma:

$$\left]\frac{S_1^2}{S_2^2}\frac{1}{F_{n_1-1,n_2-1;1-\frac{\alpha}{2}}};\frac{S_1^2}{S_2^2}F_{n_2-1,n_1-1;1-\frac{\alpha}{2}}\right[.$$

Variâncias corrigidas

$$S'^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

IC para σ_1^2/σ_2^2: Formulário

- **POPULAÇÕES NORMAIS**

Média	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\dfrac{\bar{X}-\mu}{S'/\sqrt{n}} \sim t(n-1)$
Diferença de médias	$\dfrac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}} \sim N(0,1)$	$Z=\dfrac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{S_1'^2}{m}+\dfrac{S_2'^2}{n}}} \overset{a}{\sim} t(\nu)$ onde ν é o maior inteiro contido em r, $r=\dfrac{\left(\dfrac{s_1'^2}{m}+\dfrac{s_2'^2}{n}\right)^2}{\dfrac{1}{m-1}\left(\dfrac{s_1'^2}{m}\right)^2+\dfrac{1}{n-1}\left(\dfrac{s_2'^2}{n}\right)^2}$
	$T=\dfrac{\dfrac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}}{\sqrt{\dfrac{(m-1)S_1'^2+(n-1)S_2'^2}{m+n-2}}} \sim t(m+n-2)$	
Variância	$\dfrac{nS^2}{\sigma^2}=\dfrac{(n-1)S'^2}{\sigma^2} \sim \chi^2(n-1)$	
Relação de variâncias	$\dfrac{S_1'^2}{S_2'^2}\dfrac{\sigma_2^2}{\sigma_1^2} \sim F(m-1,n-1)$	

Variâncias corrigidas

$$S'^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

Intervalo de Confiança para a Relação de Variâncias σ_1^2/σ_2^2: Exercícios

6

Exercício Suplementar que não consta do livro Murteira et al (2015)

Para substituir uma máquina antiquada encaram-se 2 alternativas: o equipamento A ou equipamento B. Dado tratar-se de uma decisão que envolve custos consideráveis uma vez que os equipamentos são bastante dispendiosos, resolveu-se testar os dois equipamentos durante um período experimental.
No final do período experimental selecionaram-se 31 e 61 peças da produção dos equipamentos A e B, respetivamente, tendo-se registado os seguintes valores relativamente à característica de interesse na avaliação da qualidade do trabalho das máquinas:

$$\sum_{i=1}^{31} x_{iA} = 43{,}4; \qquad \sum_{1=1}^{31} x_{iA}^2 = 123{,}76; \qquad \sum_{i=1}^{61} x_{iB} = 91{,}5; \qquad \sum_{i=1}^{61} x_{iB}^2 = 269{,}25$$

Utilizando um intervalo de confiança a 95% diga se há razões para crer que com a máquina A se consegue uma menor variabilidade da característica de avaliação do que com a máquina B. Admita a normalidade das distribuições.

ProbabilidadesEstatistica_2019 (uevora.pt)

Exercício: IC para σ_1^2/σ_2^2

Sejam:

- X_1 a v.a. que representa o valor da característica de interesse no equipamento A,
- X_2 a v.a. que representa o valor da característica de interesse no equipamento B,

Com $X_1 \sim N(\mu_1 = ?\,; \sigma_1 = ?\,)$ e $X_2 \sim N(\mu_2 = ?; \sigma_2 = ?\,)$.

$n_1 = 31$ e $n_2 = 61$.

Aqui são variâncias amostrais corrigidas:

$$\bar{x}_1 = \sum_{i=1}^{31} \frac{x_{1i}}{31} = \frac{43{,}4}{31} = 1{,}4; \qquad s_1^2 = \frac{1}{30}\left(\sum_{i=1}^{31} x_{1i}^2 - 31 \times \bar{x}_1^2\right) = \frac{123{,}76 - 31 \times 1{,}4^2}{30} = 2{,}1;$$

$$\bar{x}_2 = \sum_{i=1}^{61} \frac{x_{2i}}{61} = \frac{91{,}5}{61} = 1{,}5; \qquad s_2^2 = \frac{1}{60}\left(\sum_{i=1}^{61} x_{2i}^2 - 61 \times \bar{x}_2^2\right) = \frac{269{,}25 - 61 \times 1{,}5^2}{60} = 2{,}2.$$

O I. C. a 95% para $\frac{\sigma_1^2}{\sigma_2^2}$ é dado por:

$$\left]\frac{S_1^2}{S_2^2}\frac{1}{f_{n_1-1;n_2-1\,;\,1-\frac{\alpha}{2}}}\,;\,\frac{S_1^2}{S_2^2}f_{n_2-1;n_1-1\,;\,1-\frac{\alpha}{2}}\right[.$$

ProbabilidadesEstatistica_2019 (uevora.pt)

$f_{0,025;30,60} = f^*_{0,975;30,60} = 1/\, f^*_{0,025;60,30} = 1/1,94 = 0,5154$

$f_{0,975;30,60} = f^*_{0,025;30,60} = 1,82$

Exercício: IC para σ_1^2/σ_2^2

Tabela da F-Snedcor

$$F_{m,n,\varepsilon} : P(X > F_{m,n,\varepsilon}) = \varepsilon$$

n – graus de liberdade do denominador	ε	m – graus de liberdade do numerador: 1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
20	.100	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
	.050	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
	.025	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09
	.010	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
22	.100	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
	.050	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
	.025	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08	2.00
	.010	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
24	.100	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
	.050	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
	.025	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01	1.94
	.010	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
26	.100	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
	.050	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
	.025	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.49	2.39	2.28	2.22	2.16	2.09	2.03	1.95	1.88
	.010	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
28	.100	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
	.050	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
	.025	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91	1.83
	.010	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
30	.100	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
	.050	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
	.025	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87	1.79
	.010	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	.100	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
	.050	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
	.025	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72	1.64
	.010	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	.100	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
	.050	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
	.025	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58	1.48
	.010	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60

Exercício: IC para σ_1^2/σ_2^2

A probabilidade da variável fulcral "estar entre" os quantis 0,5154 e 1,82 é igual a 0,95. Posteriormente, desenvolve-se essas desigualdades e obtêm-se os limites:

$$P(0{,}5154 < s'^2_1/s'^2_2 \cdot \sigma_2^2/\sigma_1^2 < 1{,}82) = 0{,}95$$

Aqui são variâncias amostrais corrigidas:

Limite inferior = (2,1/2,2)*(1/1,82) = 0,526

Limite superior = = (2,1/2,2)*(1/0,5154) = 1,852

$IC_{95\%}(\sigma 1^2/\ \sigma 2^2) = (0{,}526;\ 1{,}852)$

Com 95% confiança não há razões para crer que exista diferença na variabilidade da característica de avaliação obtida com as duas máquinas (a igualdade das variâncias), uma vez que o valor 1 está presente no intervalo.

ProbabilidadesEstatistica_2019 (uevora.pt)

Murteira et al (2015)
Capítulo 7

37. Para comparar dois métodos de ensino, dividiu-se aleatoriamente uma turma de 22 alunos em dois grupos iguais. Cada grupo foi ensinado com um método diferente (método A e método B) e, no fim da formação, foram sujeitos à mesma prova de avaliação. Os resultados da avaliação, numa escala de 0 a 100, foram os seguintes:

$$\text{Método } A\text{: } \bar{x}_A = 74.8;\quad s_A'^2 = 81.5.$$
$$\text{Método } B\text{: } \bar{x}_B = 72.1;\quad s_B'^2 = 110.5.$$

Admita que as classificações obtidas em cada grupo têm distribuição normal.

a) Construa um intervalo de confiança a 95% para o quociente das variâncias dos dois grupos.

b) Tendo em conta o resultado da alínea anterior, construa um intervalo de confiança a 99% para a diferença de médias. Interprete o resultado obtido.

Exercício 37 a)

$X_A \equiv$ Classificação de um aluno (0 a 100) ensinado pelo método A.

$X_B \equiv$ " " " " " " " " " " B.

$n_A = n_B = 11$

$X_A \sim N(\mu_A, \sigma_A^2)$ $\quad \bar{x}_A = 74.8$ $\quad s_A'^2 = 81.5$

$X_B \sim N(\mu_B, \sigma_B^2)$ $\quad \bar{x}_B = 72.1$ $\quad s_B'^2 = 110.5$

Exercício 37 a)

a) I.C. $_{95\%}$ para $\dfrac{\sigma_B^2}{\sigma_A^2}$ $\qquad \alpha = 0.05 \qquad \dfrac{\alpha}{2} = 0.025$

Variável fulcral : $\dfrac{S_A'^2}{S_B'^2}\,\dfrac{\sigma_B^2}{\sigma_A^2} \sim F(n_A-1, n_B-1)$

$F \sim F(10, 10)$

$$P\left(F_{10,10,\,0.975} < F < F_{10,10,\,0.025}\right) = 0.95 \iff$$

$$P\left(F_{10,10,\,0.975} < \frac{S_A'^2}{S_B'^2}\cdot\frac{\sigma_B^2}{\sigma_A^2} < F_{10,10,\,0.025}\right) = 0.95$$

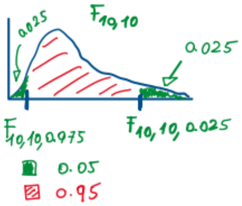


Tabela 8 : $F_{10,\,10,\,0.025} = 3.72$

$$P\left(F < F_{10,10,\,0.975}\right) = 0.025 \iff$$

$$\iff P\left(\frac{1}{F} > \frac{1}{F_{10,10,\,0.975}}\right) = 0.025 \iff$$

$$\iff P\left(F > \frac{1}{F_{10,10,\,0.975}}\right) = 0.025$$

$\hookrightarrow F_{10,10,\,0.025} = 3.72$

Porque $F \sim (10,10)$ e $\frac{1}{F} \sim F(10,10)$

logo $F_{10,10,\,0.975} = \dfrac{1}{F_{10,10,\,0.025}} = \dfrac{1}{3.72}$

$$\iff P\left(\frac{1}{3.72} < \frac{S_A'^2}{S_B'^2}\cdot\frac{\sigma_B^2}{\sigma_A^2} < 3.72\right) = 0.95 \iff$$

$$\iff P\left(\frac{1}{3.72}\,\frac{S_B'^2}{S_A'^2} < \frac{\sigma_B^2}{\sigma_A^2} < 3.72\,\frac{S_B'^2}{S_A'^2}\right) = 0.95$$

Exercício 37 a)

$$IC_{95\%} \text{ para } \frac{\sigma_B^2}{\sigma_A^2} = \left(\frac{1}{3.72} \frac{s_B'^2}{s_A'^2} ; 3.72 \frac{s_B'^2}{s_A'^2} \right) =$$

$$= \left(\frac{1}{3.72} \times \frac{110.5}{81.5} ; 3.72 \times \frac{110.5}{81.5} \right) =$$

$$= (0.36 ; 5.04)$$

Conclusão: Se $\sigma_A^2 = \sigma_B^2$ então $\sigma_B^2 / \sigma_A^2 = 1$. Um possível teste para a hipótese de igualdade de variâncias consiste em construir um I.C. para σ_B^2 / σ_A^2 e verificar se "1" pertence ao intervalo. Como $\sigma_A^2/\sigma_B^2 = 1 \in IC_{95\%}(\sigma_A^2/\sigma_B^2) = (0.36, 5.04)$ não rejeitamos a hipótese de igualdade de variâncias.

Obrigada!

Questões?