Estatística I

Licenciatura em Economia

2.º Ano/2.º Semestre

2023/2024

Aulas Teóricas N.^{os} 22 e 23 (Semana 12)

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 Teóricas (Semanas 1 a 3)	Aulas Teóricas (Semanas 3 a 5)	Aulas Teóricas (Semanas 5 a 7)	Aulas Teóricas (Semanas 8 a 13)
• **Capítulo 2**: Probabilidades	• **Capítulo 3**: Variáveis Aleatórias Unidimensionais	• **Capítulo 4**: Variáveis Aleatórias Multidimensionais	• **Capítulo 5**: Distribuições Teóricas • **Capítulo 6**: Amostragem. Distribuições por Amostragem.

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.

https://cas.iseg.ulisboa.pt

1 Inferência Estatística

Conceitos: Universo, Amostra Aleatória, Distribuição de uma Amostra Aleatória e Estatísticas

Probabilidade vs. Inferência Estatística

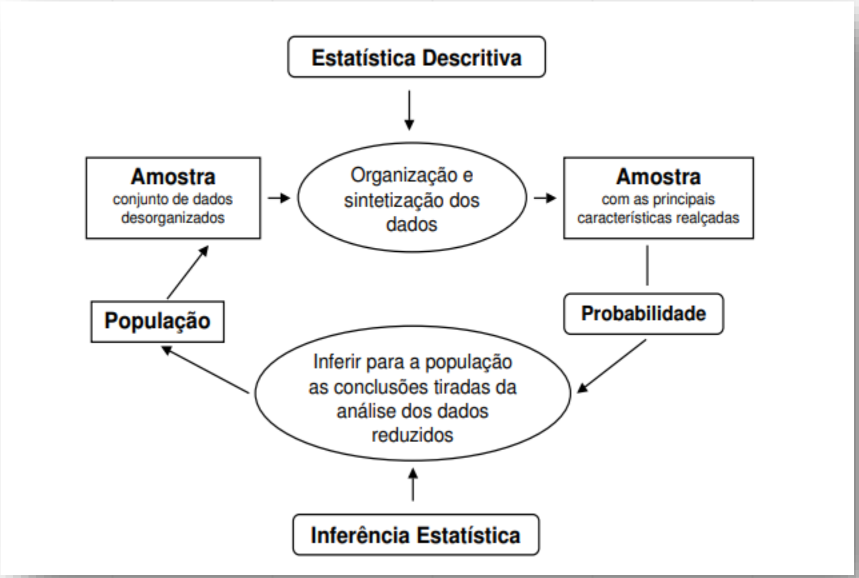
Pode dizer-se que a Probabilidade e a Inferência têm objectivos diferentes: enquanto na Probabilidade se parte de um dado esquema ou modelo para calcular a probabilidade de certos resultados ou acontecimentos se realizarem; na Inferência parte-se de dados ou observações e procura saber-se ou inferir-se algo sobre o modelo.

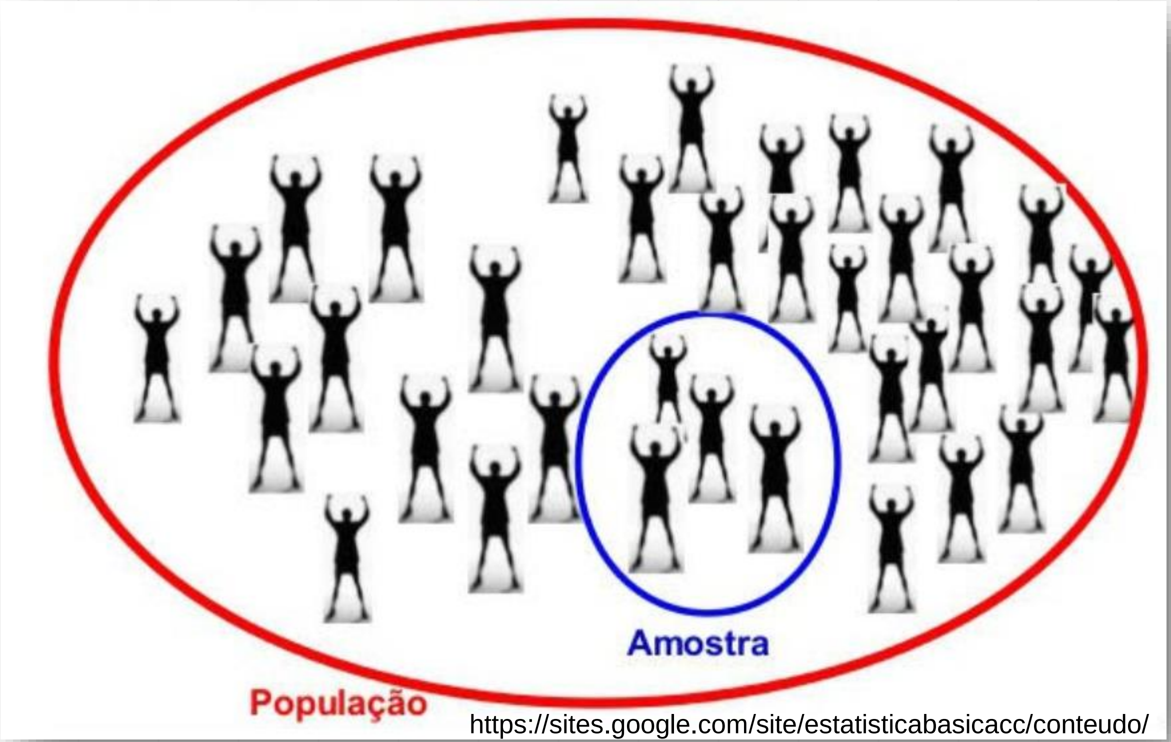
A Inferência é a "passagem do particular ao geral."

A Inferência Estatística tem como objectivo definir procedimentos que aplicados a uma amostra extraída da população, nos permitam estimar parâmetros desconhecidos dessa população ou algo sobre o modelo da população.

De facto uma amostra particular é apenas uma das muitas amostras (em nº infinito se a população for infinita) que se podem obter por um processo de amostragem.

Estatística Descritiva vs. Inferência Estatística





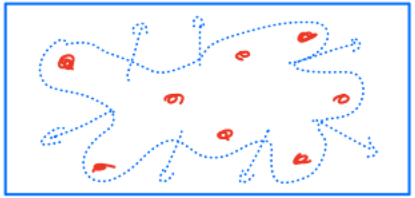
https://sites.google.com/site/estatisticabasicacc/conteudo/

- **POPULAÇÃO:**
  - é uma coleção completa de todos os elementos a serem estudados
- **AMOSTRA:**
  - é um subconjunto da população
- **CENSO:**
  - é uma coleção da dados relativos a todos os elementos de uma população:

https://slideplayer.com.br/slide/2627398

Amostra Aleatória

Perante "falta de informação" sobre a v.a. (população), vamos recolher uma amostra, estudar intensamente esta amostra e tentar extrapolar p/a população.



Ω

• – elementos da população que vão ser estudados

$\Downarrow$

amostra

$(x_1, x_2, .., x_n)$

População vs Amostra Aleatória

população → Recolher um subconjunto representativo da população →

(X)
↓
v.a. interesse

① $(X_1, X_2, \ldots, X_n)$
(amostra aleatória)

② estudar a amostra através da estatística descritiva

→ ③ inferir ("extrapolar") os resultados da amostra p/ a população.

Slides Professora Claúdia Nunes

Amostragem

① Amostragem (Regras p/ assegurar a "bondade" dos resultados)

Amostra é representativa da população se tiver as seguintes características:

$(X_1, X_2, \ldots, X_n)$: conjunto de variáveis aleatórias independentes entre si e têm todas a mesma distribuição, que é a distribuição da v.a. (população) X

a.a. = amostra aleatória
i.i.d. = independentes e identicamente distribuídas

Distribuição de uma a.a.: V.a.´s Discretas

Qual é a distribuição da a.a. $(X_1, \ldots, X_n)$, sendo as v.a.´s X_i´s ($i = 1, \ldots, n$) iid a uma v.a. discreta X?

Distribuição de uma a.a.: V.a.´s Discretas

1ª Consequência: Qual a distribuição de uma a.a.?

i) se X for uma v.a. discreta

$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$ [função de prob. conjunta da a.a.]

$$\underset{\text{independentes}}{=} P(X_1 = x_1)\,P(X_2 = x_2)\ldots P(X_n = x_n)$$

$$\underset{\text{ident. distribuídas}}{=} P(X = x_1)\,P(X = x_2)\ldots P(X = x_n)$$

$$= \prod_{i=1}^{n} P(X = x_i)$$

Slides Professora Claúdia Nunes

Distribuição de uma a.a.: V.a.´s Poisson

Determine a seguinte função de probabilidade conjunta da a.a. (X_1, X_2, X_3): $P(X_1=4, X_2=3, X_3=5)$, sendo X_i´s $(i = 1, 2, 3)$ iid a $X \cap P(\lambda)$.

Distribuição de uma a.a.: V.a.´s Poisson

exemplo: $X \sim Poisson(\lambda)$ $\lambda = ?$

$$P(X_1 = 4, X_2 = 3, X_3 = 5) = \prod_{i=1}^{3} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!}$$

$$= \frac{e^{-\lambda}\lambda^4}{4!} \frac{e^{-\lambda}\lambda^3}{3!} \frac{e^{-\lambda}\lambda^5}{5!} = \frac{e^{-3\lambda}\lambda^{12}}{4!\,3!\,5!}$$

Slides Professora Claúdia Nunes

Distribuição de uma a.a.: V.a.´s Contínuas

Qual é a distribuição da a.a. $(X1, \ldots, Xn)$, sendo Xi´s $(i = 1, \ldots, n)$ iid a uma v.a. Contínua X?

Distribuição de uma a.a.: V.a.´s Contínuas

ii) se X for uma v.a. contínua

$$f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n)$$ [função de densidade de probabilidade conjunta]

$$\underset{\text{independentes}}{=} f_{X_1}(x_1) f_{X_2}(x_2) \ldots f_{X_n}(x_n)$$

$$\underset{\text{identicamente distribuidas}}{=} f_X(x_1) f_X(x_2) \ldots f_X(x_n)$$

$$= \prod_{i=1}^{n} f_X(x_i)$$

Slides Professora Claúdia Nunes

Distribuição de uma a.a.: V.a.´s Exponenciais

Determine a seguinte função densidade de probabilidade conjunta da a.a. $(X1, X2, X3, X4)$: $f(1.3, 0.5, 2.5, 0.78)$, sendo $Xi´s$ $(i = 1, 2, 3, 4)$ iid a $X \cap Exp(\lambda)$.

Distribuição de uma a.a.: V.a.´s Exponenciais

exemplo: $X \sim Exp(\mu)$

$$x_1 = 1.3 \; ; \; x_2 = 0.5 \; ; \; x_3 = 2.5 \; ; \; x_4 = 0.78$$

$$f_{X_1, X_2, X_3, X_4}(1.3, 0.5, 2.5, 0.78) = \mu e^{-\mu \times 1.3} \times \mu e^{-\mu \times 0.5}$$

$$\times \mu e^{-\mu \times 2.5} \times \mu e^{-\mu \times 0.78}$$

$$= \mu^4 e^{-\mu(1.3 + 0.5 + 2.5 + 0.78)} \quad //$$

Inferência Estatística

③ Inferência estatística

- X ✓ (v.a. de interesse)
- com distribuição conhecida D
- Todos ou alguns parâmetros da distribuição são desconhecidos

parâmetros desconhecidos (θ)

Objectivo: "adivinhar" o valor de θ, com base na amostra

i.e

estimar θ, com base na amostra

↓

Estimação
- Pontual (cap. 6)
- Intervalar (cap. 7)

Slides Professora Claúdia Nunes

Estatística

i) <u>ESTATÍSTICA</u>: uma estatística é uma função (qualquer) que depende da amostra e não depende de nenhum parâmetro desconhecido.

<u>Exemplo</u>: $X \sim \text{Poisson}(\lambda)$ $\qquad \lambda = ?$

a.a. $(X_1, X_2, \ldots, X_{10})$

$T_1 = \overline{X} = \frac{1}{10} \sum_{i=1}^{n} X_i$: É UMA ESTATÍSTICA

$T_2 = 3$: É UMA ESTATÍSTICA

$T_3 = \frac{1}{10} \sum_{i=1}^{10} (X_i - \lambda)^2$ NÃO É ESTATÍSTICA!!

Slides Professora Claúdia Nunes

Estatística

DEF. 4: ESTATÍSTICA:

Uma estatística é uma função das variáveis observáveis, e é por si própria uma variável observável que não contém qualquer parâmetro desconhecido.

EXEMPLO 5: Seja a a.a. $[X_1, X_2, \ldots, X_n]$. A média amostral $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ é uma estatística.

EXEMPLO 6: Seja a v.a $X \sim N(\mu, \sigma^2)$ com μ e σ^2 desconhecidos, então $x - \mu$ não é estatística porque não é observável, é função do parâmetro desconhecido μ.

EXEMPLO 6: Seja uma v.a. com distribuição $N(\mu;\sigma^2)$. Quais são Estatísticas?

a) $X^2-\mu$

b) $\frac{X}{\sigma^2}$

c) X^2-3

d) $X-4$

e) $X - \log X^3$

Quais destas funções são estatísticas?

apostila.PDF (ufpr.br)

Média e Variância Amostrais

(2) Duas características da amostra frequentemente calculadas são:

média amostral: $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i$

variância amostral: $s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$

$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2$ (*)

(*) porque: $\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 =$

$= \frac{1}{n} \sum_{i=1}^{n} (x_i^2 + \bar{x}^2 - 2x_i\bar{x}) = \frac{1}{n} \left[\sum_{i=1}^{n} x_i^2 + n\bar{x}^2 - 2\bar{x} \underbrace{\sum_{i=1}^{n} x_i}_{n\bar{x}} \right]$

$= \frac{1}{n} \left[\sum_{i=1}^{n} x_i^2 + n\bar{x}^2 - 2n\bar{x}^2 \right]$

$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2$

variância amostral corrigida: $s^2_{n-1} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

2

Média Amostral

Distribuições de Amostragem

Média Amostral

Seja $X_1, X_2, \dots, X_n$, uma a. a. duma dada população com média μ e variância σ^2. A **média amostral** é definida por:

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

Propriedades:

- $\mu_{\overline{X}} = E(\overline{X}) = \mu$;
- $\sigma^2_{\overline{X}} = Var(\overline{X}) = \frac{1}{n}\sigma^2$.

ProbabilidadesEstatistica2019.pdf

Média Amostral: Variância Conhecida

Se a distribuição da população é Normal com desvio padrão σ conhecido, então $\overline{X} \sim N\left(\mu; \frac{\sigma}{\sqrt{n}}\right)$, ou seja,

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0;1).$$

Se a distribuição da população não for Normal, mas a amostra é de grande dimensão então, pelo corolário do T. L. C., vem

$$Z \overset{\circ}{\sim} N(0;1).$$

ProbabilidadesEstatistica2019.pdf

Média Amostral: Variância Desconhecida

Se a população é Normal mas o desvio padrão σ é desconhecido, e não rejeitando a hipótese de independência das distribuições por amostragem da média e da variância da amostra, então tem-se que:

$$T = \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}.$$

Se a distribuição da população não for Normal, mas a amostra for de grande dimensão então, por extensão do corolário do T. L. C.,

$$\overline{X} \overset{\circ}{\sim} N\left(\mu; \frac{S}{\sqrt{n}}\right), \text{ ou seja, } Z = \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \overset{\circ}{\sim} N(0; 1).$$

Observação: Para valores elevados de n, a distribuição *t*-Student toma valores muito próximos aos da $N(0; 1)$. Existem alguns programas estatísticos (por exemplo, SPSS) que, nestas condições, não utilizam a distribuição Normal mas a *t*-Student.

ProbabilidadesEstatistica2019.pdf

Distribuições de Amostragem

Formulário

AMOSTRAGEM. DISTRIBUIÇÕES POR AMOSTRAGEM

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \; ; \qquad S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n} = \frac{\sum_{i=1}^{n} X_i^2}{n} - \overline{X}^2 \; ; \qquad (n-1)S'^2 = n\, S^2$$

$$E(\overline{X}) = \mu \; ; \qquad \mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n} \; ; \qquad E(S^2) = \frac{n-1}{n}\sigma^2 \; ; \qquad E(S'^2) = \sigma^2$$

Média Amostral: Exercícios

Distribuições de Amostragem

3

Um fabricante de automóveis defende que o novo modelo que vai ser lançado no próximo mês gasta em média 9,7 litros aos 100 km, em circuito urbano, com desvio padrão de 1 litro. Admita que o consumo segue uma distribuição Normal.

a) Qual a probabilidade de numa amostra aleatória de 20 automóveis o gasto médio ser superior a 10 litros.

b) Qual a deverá ser a dimensão da amostra para obter, com pelo menos 90% probabilidade, um gasto médio inferior a 10 litros.

ProbabilidadesEstatistica2019.pdf

Exercício: Média Amostral com Variância Conhecida

Seja X a v.a. que representa o número de litros consumidos pelo automóvel aos 100 km, em circuito urbano, com $X \sim N(\mu = 9{,}7;\ \sigma = 1)$.

a) $n = 20$.

$$\left.\begin{array}{l} X \text{ dist. Normal} \\ \sigma \text{ conhecido} \end{array}\right| \Rightarrow Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0;1).$$

$$P(\overline{X} > 10) = 1 - P(\overline{X} \leq 10) = 1 - P\left(Z \leq \frac{10 - 9{,}7}{\frac{1}{\sqrt{20}}}\right) = 1 - P(Z \leq 1{,}34) = 1 - \Phi(1{,}34)$$

$$= 1 - 0{,}9099 = 0{,}0901.$$

b) $n = ?$

$$P(\overline{X} < 10) \geq 0{,}9 \Leftrightarrow P\left(Z < \frac{10 - 9{,}7}{\frac{1}{\sqrt{n}}}\right) \geq 0{,}9 \Leftrightarrow \Phi(0{,}3\sqrt{n}) \geq 0{,}9$$

como $\Phi(1{,}282) = 0{,}9$

$$\Leftrightarrow 0{,}3\sqrt{n} \geq 1{,}282 \Leftrightarrow \sqrt{n} \geq 4{,}2733 \Rightarrow n \geq 4{,}2733^2 = 18{,}3 \Rightarrow n \geq 19.$$

ProbabilidadesEstatistica2019.pdf

Um fabricante de automóveis defende que o novo modelo que vai ser lançado no próximo mês gasta em média 9,7 litros aos 100 km, em circuito urbano, e o desvio padrão é desconhecido.
Através de um esquema de amostragem estimou-se tal desvio padrão como sendo $s = 1$ litro. Admitindo que o consumo segue uma distribuição Normal, qual a probabilidade de, numa amostra aleatória de 20 automóveis, o consumo médio amostral ser superior a 10 litros? E inferior a 8,9 litros?

ProbabilidadesEstatistica2019.pdf

Exercício: Média Amostral com Variância Desconhecida

Seja X a v.a. que representa o número de litros consumidos pelo automóvel aos 100 km, em circuito urbano, com $X \sim N(\mu = 9{,}7;\ \sigma = ?)$.

$n = 20.$

$$\left.\begin{array}{l} X \text{ dist. Normal} \\ \sigma \text{ desconhecido} \end{array}\right| \Rightarrow T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1=19}.$$

$$P(\bar{X} > 10) = 1 - P(\bar{X} \leq 10) = 1 - P\left(T \leq \frac{10 - 9{,}7}{\frac{1}{\sqrt{20}}}\right) = 1 - P(T \leq 1{,}342) \approx 1 - 0{,}9 = 0{,}1.$$

$$P(\bar{X} < 8{,}9) = P\left(T < \frac{8{,}9 - 9{,}7}{\frac{1}{\sqrt{20}}}\right) = P(T < -3{,}578) = 1 - P(T < 3{,}578) \approx 1 - 0{,}999 = 0{,}001.$$

ProbabilidadesEstatistica2019.pdf

Diferença de Médias Amostrais

4

Distribuições de Amostragem

Diferença de Médias Amostrais

Sejam $X_{11}, X_{12}, \dots, X_{1n_1}$ e $X_{21}, X_{22}, \dots, X_{2n_2}$ duas a. a. independentes, de dimensão n_1 e n_2 retiradas de duas populações com médias μ_1 e μ_2 e desvios padrão σ_1 e σ_2, respetivamente, e

$$\overline{X}_1 = \sum_{i=1}^{n_1} \frac{X_{1i}}{n_1} \text{ e } \overline{X}_2 = \sum_{i=1}^{n_2} \frac{X_{2i}}{n_2}.$$

ProbabilidadesEstatistica2019.pdf

Diferença de Médias: Variâncias Conhecidas

Se as populações forem Normais, sendo $X_1 \sim N(\mu_1; \sigma_1)$ e $X_2 \sim N(\mu_2; \sigma_2)$, com σ_1 e σ_2 conhecidos, então, pelo Teorema da aditividade da distribuição Normal, tem-se que:

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0; 1).$$

Se as populações não forem Normais, mas as amostras forem de grande dimensão então, pelo corolário do T. L. C., vem $Z \overset{\circ}{\sim} N(0; 1)$, considerando o já anteriormente exposto em situação análoga.

ProbabilidadesEstatistica2019.pdf

Diferença de Médias: Variâncias Desconhecidas e Iguais

Se as populações forem Normais, sendo $X_1 \sim N(\mu_1; \sigma_1)$ e $X_2 \sim N(\mu_2; \sigma_2)$, com σ_1 e σ_2 desconhecidos mas iguais $(\sigma_1 = \sigma_2)$, então demonstra-se que:

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}.$$

Se as populações não forem Normais, mas as amostras forem de grande dimensão então, por extensão do T. L. C., a expressão anterior segue aproximadamente uma $N(0; 1)$.

ProbabilidadesEstatistica2019.pdf

Diferença de Médias: Variâncias Desconhecidas e Diferentes

Se as populações forem Normais, sendo $X_1 \sim N(\mu_1; \sigma_1)$ e $X_2 \sim N(\mu_2; \sigma_2)$, com σ_1 e σ_2 desconhecidos e diferentes ($\sigma_1 \neq \sigma_2$), então, pela aproximação de Welch tem-se que (Murteira *et al.*, 2007):

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \mathrel{\overset{\circ}{\sim}} t_v, \text{ onde } v = \left[\frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{S_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{S_2^2}{n_2}\right)^2} \right].$$

sendo $[r]$ a parte inteira de r, ou seja, arredonda-se por defeito o valor obtido.

Também neste caso, se as populações não forem Normais, mas as amostras forem de grande dimensão então, por extensão do T. L. C., a expressão anterior segue aproximadamente uma $N(0; 1)$, sendo válidas as observações anteriores.

ProbabilidadesEstatistica2019.pdf

Diferença de Médias Amostrais: Exercícios

5

Distribuições de Amostragem

Uma dada empresa farmacêutica lançou no mercado um novo medicamento, para dormir, que tem estado a ser utilizado nos hospitais. Constatou-se que os doentes não sujeitos a este medicamento em média dormiam 7,5 horas, com desvio padrão de 1,4 horas, ao passo que os doentes aos quais se administrou este medicamento dormiam em média 8 horas com desvio padrão de 2 horas.
Num determinado hospital observaram-se 31 doentes não sujeitos ao referido medicamento e 61 sob a referida medicação. Qual a probabilidade de os doentes do primeiro grupo observado dormirem em média mais do que os do segundo grupo? Assuma a normalidade das distribuições.

ProbabilidadesEstatistica2019.pdf

Exercício: Diferença de Médias Amostrais e Variâncias Conhecidas

Sejam:

- X_1 a v.a. que representa o número de horas de sono dos doentes não sujeitos ao medicamento,
- X_2 a v.a. que representa o número de horas de sono dos doentes sujeitos ao medicamento,

com

$$X_1 \sim N(\mu_1 = 7{,}5; \sigma_1 = 1{,}4) \text{ e } X_2 \sim N(\mu_2 = 8; \sigma_2 = 2).$$

$$\left.\begin{array}{r} X_1 \text{ e } X_2 \text{ dist. Normal} \\ \sigma_1\ (= 1{,}4) \text{ e } \sigma_2\ (= 2) \text{ conhecidos} \end{array}\right| \Rightarrow Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0; 1).$$

$$P(\overline{X}_1 > \overline{X}_2) = P(\overline{X}_1 - \overline{X}_2 > 0) = 1 - P(\overline{X}_1 - \overline{X}_2 \leq 0) = 1 - P\left(Z \leq \frac{0 - (7{,}5 - 8)}{\sqrt{\frac{1{,}4^2}{31} + \frac{2^2}{61}}}\right) = 1 - \Phi(1{,}39)$$

$$= 1 - 0{,}9177 = 0{,}0823.$$

ProbabilidadesEstatistica2019.pdf

Uma dada empresa farmacêutica lançou no mercado um novo medicamento, para dormir, que tem estado a ser utilizado nos hospitais. Constatou-se que os doentes não sujeitos a este medicamento em média dormiam 7,5 horas, enquanto os doentes aos quais se administrou este medicamento dormiam em média 8 horas.

Num determinado hospital observaram-se n_1 doentes não sujeitos ao referido medicamento e n_2 sob a referida medicação tendo-se obtido, respetivamente, os seguintes desvios-padrão: 1,4 horas e 2 horas. Determine a probabilidade de os doentes do primeiro grupo dormirem em média menos do que os do segundo grupo, quando $n_1 = 20$ e $n_2 = 27$, quando se verifica a normalidade das distribuições e se considera:

a) A igualdade das variâncias populacionais.

b) A desigualdade das variâncias populacionais.

ProbabilidadesEstatistica2019.pdf

Exercício: Diferença de Médias Amostrais e Variâncias Desconhecidas e Iguais

Sejam:

- X_1 a v.a. que representa o número de horas de sono dos doentes não sujeitos ao medicamento,
- X_2 a v.a. que representa o número de horas de sono dos doentes sujeitos ao medicamento,

Com $X_1 \sim N(\mu_1 = 7{,}5; \sigma_1 =?)$ e $X_2 \sim N(\mu_2 = 8; \sigma_2 =?)$.

$n_1 = 20;\ \ s_1 = 20;\ \ n_2 = 27$ e $s_2 = 2.$

a)

X_1 e X_2 dist. Normal
σ_1 e σ_2 desconhecidos, mas iguais

$$\Rightarrow T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2=45}.$$

$$P(\overline{X}_1 < \overline{X}_2) = P(\overline{X}_1 - \overline{X}_2 < 0) = P\left(T < \frac{0 - (7{,}5 - 8)}{\sqrt{\frac{(20-1)1{,}4^2 + (27-1)2^2}{20 + 27 - 2}}\sqrt{\frac{1}{20} + \frac{1}{27}}}\right) = P(T < 0{,}957)$$

$$\approx 0{,}83.$$

ProbabilidadesEstatistica2019.pdf

Exercício: Diferença de Médias Amostrais e Variâncias Desconhecidas e Diferentes

b)

X_1 e X_2 dist. Normal
σ_1 e σ_2 desconhecidos e diferentes

$$\Rightarrow T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \overset{\circ}{\sim} t_{v=44},$$

pois

$$v = \left[\frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}\right] = \left[\frac{\left(\frac{1,4^2}{20} + \frac{2^2}{27}\right)^2}{\frac{1}{20 - 1}\left(\frac{1,4^2}{20}\right)^2 + \frac{1}{27 - 1}\left(\frac{2^2}{27}\right)^2}\right] = [44{,}9] = 44.$$

$$P(\overline{X}_1 < \overline{X}_2) = P(\overline{X}_1 - \overline{X}_2 < 0) = P\left(T < \frac{0 - (7{,}5 - 8)}{\sqrt{\frac{1,4^2}{20} + \frac{2^2}{27}}}\right) = P(T < 1{,}008) \approx 0{,}84.$$

ProbabilidadesEstatistica2019.pdf

6 Variância Amostral

Distribuições de Amostragem

Variância Amostral

Seja $X_1, X_2, \ldots, X_n$, uma a. a. duma dada população com variância σ^2. A **variância amostral** é definida por:

$$S^2 = \sum_{i=1}^{n} \frac{\left(X_i - \overline{X}\right)^2}{n-1}.$$

Propriedades:

- $\mu_{S^2} = E(S^2) = \sigma^2.$
- Se a distribuição da população for Normal então $\sigma^2_{S^2} = Var(S^2) = \frac{2\sigma^4}{n-1}.$

Se a distribuição da população for Normal, com variância σ^2, então:

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

ProbabilidadesEstatistica2019.pdf

Variância Amostral: Exercícios

7

Distribuições de Amostragem

Uma determinada empresa farmacêutica lançou no mercado um novo medicamento, para dormir, que tem estado a ser utilizado nos hospitais. Constatou-se que os doentes sujeitos a este medicamento em média dormiam 8 horas, sendo o desvio padrão de 2 horas, e que a distribuição do número de horas de sono podia ser considerada Normal.

Qual a probabilidade de, numa amostra aleatória de 31 doentes sujeitos ao referido medicamento:

a) A variância amostral ser superior a 5 horas?

b) A variância amostral ser inferior a 2,25 horas?

ProbabilidadesEstatistica2019.pdf

Exercício: Variância Amostral

Seja X a v.a. que representa o número de horas de sono dos doentes sujeitos ao medicamento, com $X \sim N(\mu = 8;\ \sigma = 2)$.

$$\left.\begin{array}{l} X \text{ dist. Normal} \\ n = 31 \end{array}\right| \Rightarrow \chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1=30}.$$

a) $P(S^2 > 5) = 1 - P(S^2 \leq 5) = 1 - P\left(\chi^2 \leq \frac{(31-1)\times 5}{2^2}\right) = 1 - P(\chi^2 \leq 37{,}5) \approx 1 - 0{,}84 = 0{,}16.$

b) $P(S^2 < 2{,}25) = 1 - P\left(\chi^2 < \frac{(31-1)\times 2{,}25}{2^2}\right) = P(\chi^2 \leq 16{,}875) \approx 0{,}026.$

ProbabilidadesEstatistica2019.pdf

Obrigada!

Questões?