Intervalos de Confiança

Estatística II - 2024/2025
ISCTE-IUL

Afonso Moniz Moreira[1,2]

[1] ISCTE-IUL, Departamento de Métodos Quantitativos para a Economia e Gestão
[2] CMVM - Comissão do Mercado de Valores Mobiliários, Departamento de Supervisão de Mercados

Aviso/Disclaimer

- Este conjunto de slides não são, nem pretendem ser uma substituição à bibliografia principal da cadeira de estatística II.
- Este conjunto de slides não são, nem pretendem ser uma fonte rigorosa de estudo dos tópicos da cadeira.
- O único propósito destes slides é ajudar o autor a guiar as aulas da forma mais coloquial possível sem ter de carregar formalismos desnecessários.
- Assim sendo, o formalismo estatístico é eliminado sempre que possível para agilizar uma primeira aprendizagem por parte dos estudantes.

O que são? Para que servem ?

- Um intervalo de confiança, um IC, é uma forma de se fazer inferência, em que se utiliza um estimador pontual e o seu desvio-padrão.

O que são? Para que servem ?

- Um intervalo de confiança, um IC, é uma forma de se fazer inferência, em que se utiliza um estimador pontual e o seu desvio-padrão.
- Ao estimador pontual é adicionado e subtraído o seu desvio-padrão ponderado por um valor crítico, dando origem a um intervalo.

O que são? Para que servem ?

- Um intervalo de confiança, um IC, é uma forma de se fazer inferência, em que se utiliza um estimador pontual e o seu desvio-padrão.
- Ao estimador pontual é adicionado e subtraído o seu desvio-padrão ponderado por um valor crítico, dando origem a um intervalo.
- O cálculo deste valor crítico vai ser detalhado mais adiante.

O que são? Para que servem ?

- Um intervalo de confiança, um IC, é uma forma de se fazer inferência, em que se utiliza um estimador pontual e o seu desvio-padrão.
- Ao estimador pontual é adicionado e subtraído o seu desvio-padrão ponderado por um valor crítico, dando origem a um intervalo.
- O cálculo deste valor crítico vai ser detalhado mais adiante.
- Assim sendo, quando se utiliza uma realização da amostra aleatória (i.e., amostra observada), obtemos um intervalo em vez um número como fizemos até aqui.

O que são? Para que servem ?

- Um intervalo de confiança, um IC, é uma forma de se fazer inferência, em que se utiliza um estimador pontual e o seu desvio-padrão.
- Ao estimador pontual é adicionado e subtraído o seu desvio-padrão ponderado por um valor crítico, dando origem a um intervalo.
- O cálculo deste valor crítico vai ser detalhado mais adiante.
- Assim sendo, quando se utiliza uma realização da amostra aleatória (i.e., amostra observada), obtemos um intervalo em vez um número como fizemos até aqui.
- Ficamos imediatamente a perceber a amplitude de valores que **poderá** conter o parâmetro de interesse.

O que são? Para que servem ?

- A título de exemplo, vamos considerar a média amostral $(\bar{X})$ que é um estimador pontual centrado para a média da população X e o seu desvio-padrão $\sigma_{\bar{X}}$, respetivamente:

O que são? Para que servem ?

- A título de exemplo, vamos considerar a média amostral $(\bar{X})$ que é um estimador pontual centrado para a média da população X e o seu desvio-padrão $\sigma_{\bar{X}}$, respetivamente:

$$\bar{X} = \sum_{i=1}^{N} X_i, \qquad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}, \tag{1}$$

então o intervalo de confiança, em termos gerais será algo deste género:

O que são? Para que servem ?

- A título de exemplo, vamos considerar a média amostral ($\bar{X}$) que é um estimador pontual centrado para a média da população X e o seu desvio-padrão $\sigma_{\bar{X}}$, respetivamente:

$$\bar{X} = \sum_{i=1}^{N} X_i, \qquad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}, \tag{1}$$

então o intervalo de confiança, em termos gerais será algo deste género:

$$]IC[_{\mu} =]\bar{X} - \text{v.c} \times \sigma_{\bar{X}}; \bar{X} + \text{v.c} \times \sigma_{\bar{X}}[\tag{2}$$

O que são? Para que servem ?

- A título de exemplo, vamos considerar a média amostral $(\bar{X})$ que é um estimador pontual centrado para a média da população X e o seu desvio-padrão $\sigma_{\bar{X}}$, respetivamente:

$$\bar{X} = \sum_{i=1}^{N} X_i, \qquad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}, \tag{1}$$

então o intervalo de confiança, em termos gerais será algo deste género:

$$]IC[_{\mu} =]\bar{X} - \text{v.c} \times \sigma_{\bar{X}}; \bar{X} + \text{v.c} \times \sigma_{\bar{X}}[\tag{2}$$

- Portanto, um intervalo de confiança é um estimador pontual com uma margem de erro dada pelo desvio-padrão ponderado pelo valor crítico.

O que são? Para que servem ?

- Então afinal como se obtém o valor crítico ?

O que são? Para que servem ?

- Então afinal como se obtém o valor crítico ?
- O valor crítico é um quantil da distribuição do estimador pontual que estamos a usar na inferência.

O que são? Para que servem ?

- Então afinal como se obtém o valor crítico ?
- O valor crítico é um quantil da distribuição do estimador pontual que estamos a usar na inferência.
- Este é determinado pela escolha do nível de confiança do intervalo, correntemente denominado por λ, e vice-versa.

O que são? Para que servem ?

- Então afinal como se obtém o valor crítico ?
- O valor crítico é um quantil da distribuição do estimador pontual que estamos a usar na inferência.
- Este é determinado pela escolha do nível de confiança do intervalo, correntemente denominado por λ, e vice-versa.
- Quanto maior o nível de confiança (λ), maior é valor crítico (v.c$^{\lambda}$) porque estamos a considerar uma maior possibilidade de valores da distribuição do estimador.

O que são? Para que servem ?

- Então afinal como se obtém o valor crítico ?
- O valor crítico é um quantil da distribuição do estimador pontual que estamos a usar na inferência.
- Este é determinado pela escolha do nível de confiança do intervalo, correntemente denominado por λ, e vice-versa.
- Quanto maior o nível de confiança (λ), maior é valor crítico (v.c$^{\lambda}$) porque estamos a considerar uma maior possibilidade de valores da distribuição do estimador.
- Os níveis de confiança mais frequentes são 90%, 95% e 99%.

O que são? Para que servem ?

- Então afinal como se obtém o valor crítico ?
- O valor crítico é um quantil da distribuição do estimador pontual que estamos a usar na inferência.
- Este é determinado pela escolha do nível de confiança do intervalo, correntemente denominado por λ, e vice-versa.
- Quanto maior o nível de confiança (λ), maior é valor crítico ($v.c^{\lambda}$) porque estamos a considerar uma maior possibilidade de valores da distribuição do estimador.
- Os níveis de confiança mais frequentes são 90%, 95% e 99%.
- Estes dão origem a valores críticos distintos, respetivamente $v.c^{90\%}$, $v.c^{95\%}$ e $v.c^{99\%}$, que são diferentes consoante a distribuição do estimador pontual que estamos a utilizar.

O que são? Para que servem ?

- No entanto, os quantis da distribuição do estimador pontual nunca são dados diretamente pelo nível de confiança λ mas como uma função deste.

O que são? Para que servem ?

- No entanto, os quantis da distribuição do estimador pontual nunca são dados diretamente pelo nível de confiança λ mas como uma função deste.
- Os quantis da distribuição do estimador são assim dados por:

$$\text{v.c}_a^\lambda = Q_{\frac{1-\lambda}{2}} = Q_{\frac{\alpha}{2}} \quad \text{e} \quad \text{v.c}_b^\lambda = Q_{\lambda+\frac{1-\lambda}{2}} = Q_{\lambda+\frac{\alpha}{2}}. \qquad (3)$$

O que são? Para que servem ?

- No entanto, os quantis da distribuição do estimador pontual nunca são dados diretamente pelo nível de confiança λ mas como uma função deste.
- Os quantis da distribuição do estimador são assim dados por:

$$\mathrm{v.c}_a^{\lambda} = Q_{\frac{1-\lambda}{2}} = Q_{\frac{\alpha}{2}} \quad \text{e} \quad \mathrm{v.c}_b^{\lambda} = Q_{\lambda+\frac{1-\lambda}{2}} = Q_{\lambda+\frac{\alpha}{2}}. \tag{3}$$

- Se a distribuição do estimador for simétrica, então $\mathrm{v.c}_a^{\lambda} = -\mathrm{v.c}_b^{\lambda} \Longleftrightarrow Q_{\frac{1-\lambda}{2}} = -Q_{\lambda+\frac{1-\lambda}{2}} \Longleftrightarrow Q_{\frac{\alpha}{2}} = -Q_{\lambda+\frac{\alpha}{2}}.$

O que são? Para que servem ?

- No entanto, os quantis da distribuição do estimador pontual nunca são dados diretamente pelo nível de confiança λ mas como uma função deste.
- Os quantis da distribuição do estimador são assim dados por:

$$\text{v.c}_a^\lambda = Q_{\frac{1-\lambda}{2}} = Q_{\frac{\alpha}{2}} \quad \text{e} \quad \text{v.c}_b^\lambda = Q_{\lambda+\frac{1-\lambda}{2}} = Q_{\lambda+\frac{\alpha}{2}}. \tag{3}$$

- Se a distribuição do estimador for simétrica, então $\text{v.c}_a^\lambda = -\text{v.c}_b^\lambda \iff Q_{\frac{1-\lambda}{2}} = -Q_{\lambda+\frac{1-\lambda}{2}} \iff Q_{\frac{\alpha}{2}} = -Q_{\lambda+\frac{\alpha}{2}}$.
- A relação $\alpha = 1 - \lambda$ será clarificada quando abordarmos os ensaios de hipóteses, por agora é simplesmente uma equação.

O que são? Para que servem ?

- Para o exemplo anterior, sabemos que a distribuição da média amostral ($\bar{X}$) é a Gaussiana (i.e., Normal), logo se quisermos um nível de confiança de 95% temos de calcular o quantil 97.5 da Gaussiana (i.e., tem de sobrar 2.5% para cada cauda) que para a Gaussiana padrão será v.c$^{95\%} = Z_{97.5} = 1.96$.

O que são? Para que servem ?

- Para o exemplo anterior, sabemos que a distribuição da média amostral ($\bar{X}$) é a Gaussiana (i.e., Normal), logo se quisermos um nível de confiança de 95% temos de calcular o quantil 97.5 da Gaussiana (i.e., tem de sobrar 2.5% para cada cauda) que para a Gaussiana padrão será $\text{v.c}^{95\%} = Z_{97.5} = 1.96$.
- Então podemos concretizar mais o intervalo (2):

$$]IC_{95\%}[_{\mu} = \left] \underbrace{\sum_{i=1}^{N} \frac{1}{N} X_i}_{\bar{X}} - \underbrace{1.96}_{\text{v.c}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}} ; \underbrace{\sum_{i=1}^{N} \frac{1}{N} X_i}_{\bar{X}} + \underbrace{1.96}_{\text{v.c}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}} \right[\tag{4}$$

O que são? Para que servem ?

- Para o exemplo anterior, sabemos que a distribuição da média amostral ($\bar{X}$) é a Gaussiana (i.e., Normal), logo se quisermos um nível de confiança de 95% temos de calcular o quantil 97.5 da Gaussiana (i.e., tem de sobrar 2.5% para cada cauda) que para a Gaussiana padrão será $\text{v.c}^{95\%} = Z_{97.5} = 1.96$.
- Então podemos concretizar mais o intervalo (2):

$$]IC_{95\%}[_{\mu} = \left] \underbrace{\sum_{i=1}^{N} \frac{1}{N} X_i}_{\bar{X}} - \underbrace{1.96}_{\text{v.c}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}} ; \underbrace{\sum_{i=1}^{N} \frac{1}{N} X_i}_{\bar{X}} + \underbrace{1.96}_{\text{v.c}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}} \right[\tag{4}$$

- Como podemos interpretar este intervalo?

O que são? Para que servem ?

- É mais fácil de perceber a interpretação se escrevermos o intervalo da seguinte maneira:

O que são? Para que servem ?

- É mais fácil de perceber a interpretação se escrevermos o intervalo da seguinte maneira:

$$\mathbb{P}\left[\underbrace{\sum_{i=1}^{N}\frac{1}{N}X_i}_{\bar{X}} - \underbrace{1.96}_{\text{v.c}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}} < \mu < \underbrace{\sum_{i=1}^{N}\frac{1}{N}X_i}_{\bar{X}} + \underbrace{1.96}_{\text{v.c}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}}\right] = 0.95 \tag{5}$$

O que são? Para que servem ?

- É mais fácil de perceber a interpretação se escrevermos o intervalo da seguinte maneira:

$$\mathbb{P}\left[\underbrace{\sum_{i=1}^{N}\frac{1}{N}X_i}_{\bar{X}} - \underbrace{1.96}_{\text{v.c}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}} < \mu < \underbrace{\sum_{i=1}^{N}\frac{1}{N}X_i}_{\bar{X}} + \underbrace{1.96}_{\text{v.c}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}}\right] = 0.95 \tag{5}$$

- A probabilidade do verdadeiro valor da média da população (μ) estar contido no intervalo (5) é de 95%.

O que são? Para que servem ?

- É mais fácil de perceber a interpretação se escrevermos o intervalo da seguinte maneira:

$$\mathbb{P}\left[\underbrace{\sum_{i=1}^{N}\frac{1}{N}X_i}_{\bar{X}} - \underbrace{1.96}_{\text{v.c}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}} < \mu < \underbrace{\sum_{i=1}^{N}\frac{1}{N}X_i}_{\bar{X}} + \underbrace{1.96}_{\text{v.c}^{95\%}} \times \underbrace{\frac{\sigma}{\sqrt{N}}}_{\sigma_{\bar{X}}}\right] = 0.95 \tag{5}$$

- A probabilidade do verdadeiro valor da média da população (μ) estar contido no intervalo (5) é de 95%.
- Ou seja, se substituirmos 100 realizações da amostra aleatória no intervalo anterior, 95 deles vão conter o verdadeiro valor de μ.

O que são? Para que servem ?

- No entanto, **não** podemos afirmar em quais dos 100 intervalos concretos calculados está o verdadeiro μ, **porque ?**

O que são? Para que servem ?

- No entanto, **não** podemos afirmar em quais dos 100 intervalos concretos calculados está o verdadeiro μ, **porque ?**
- O intervalo (4) e o anterior (5), são chamados de intervalos aleatórios porque dependem das variáveis aleatórias $(X_1, ..., X_N)$, ainda não foram substituídos pelos valores da realização da amostra aleatória.

O que são? Para que servem ?

- No entanto, **não** podemos afirmar em quais dos 100 intervalos concretos calculados está o verdadeiro μ, **porque ?**
- O intervalo (4) e o anterior (5), são chamados de intervalos aleatórios porque dependem das variáveis aleatórias $(X_1, ..., X_N)$, ainda não foram substituídos pelos valores da realização da amostra aleatória.
- Atenção!!! A interpretação anterior é para o IC aleatório e não para o concreto.

O que são? Para que servem ?

- No entanto, **não** podemos afirmar em quais dos 100 intervalos concretos calculados está o verdadeiro μ, **porque ?**
- O intervalo (4) e o anterior (5), são chamados de intervalos aleatórios porque dependem das variáveis aleatórias $(X_1, ..., X_N)$, ainda não foram substituídos pelos valores da realização da amostra aleatória.
- Atenção!!! A interpretação anterior é para o IC aleatório e não para o concreto.
- Quando substituímos uma amostra observada passa a ser considerado um intervalo de confiança concreto, porque tem valores concretos/observados, as variáveis aleatórias $(X_1, ..., X_n)$ dão lugar às suas realizações $(x_1, ..., x_N)$.

O que são? Para que servem ?

- No entanto, **não** podemos afirmar em quais dos 100 intervalos concretos calculados está o verdadeiro μ, **porque ?**
- O intervalo (4) e o anterior (5), são chamados de intervalos aleatórios porque dependem das variáveis aleatórias $(X_1, ..., X_N)$, ainda não foram substituídos pelos valores da realização da amostra aleatória.
- Atenção!!! A interpretação anterior é para o IC aleatório e não para o concreto.
- Quando substituímos uma amostra observada passa a ser considerado um intervalo de confiança concreto, porque tem valores concretos/observados, as variáveis aleatórias $(X_1, ..., X_n)$ dão lugar às suas realizações $(x_1, ..., x_N)$.
- Quando nos é apresentado um IC concreto apenas se pode esperar que seja um dos que contêm o verdadeiro valor de μ.

Exemplos Reais do Uso de Intervalos de Confiança I

- Eleições Legislativas e Presidenciais:

Exemplos Reais do Uso de Intervalos de Confiança I

- Eleições Legislativas e Presidenciais:
  - Ainda que sejam apresentado uma estimativa pontual para cada candidato XPTO, de seguida é dito que estes valores têm em conta um intervalo de confiança de $x\%$, normalmente de 95%

Exemplos Reais do Uso de Intervalos de Confiança I

- Eleições Legislativas e Presidenciais:
  - Ainda que sejam apresentado uma estimativa pontual para cada candidato XPTO, de seguida é dito que estes valores têm em conta um intervalo de confiança de x%, normalmente de 95%
  - O ponto médio do intervalo corresponde à estimativa pontual. No entanto um estimador tem variância e esta tem de ser considerada na análise do fenómeno. Portanto um intervalo de confiança tem em conta uma margem de erro que têm subjacente um determinado nível de confiança.

Exemplos Reais do Uso de Intervalos de Confiança I

- Eleições Legislativas e Presidenciais:
    - Ainda que sejam apresentado uma estimativa pontual para cada candidato XPTO, de seguida é dito que estes valores têm em conta um intervalo de confiança de $x\%$, normalmente de 95%
    - O ponto médio do intervalo corresponde à estimativa pontual. No entanto um estimador tem variância e esta tem de ser considerada na análise do fenómeno. Portanto um intervalo de confiança tem em conta uma margem de erro que têm subjacente um determinado nível de confiança.
    - É precisamente com estas bandas de variação que se consegue ir fazendo projeções ao longo da noite eleitoral.

Exemplos Reais do Uso de Intervalos de Confiança II

- Ensaios clínicos para a eficácia de um novo tratamento médico para correção da miopia:

Exemplos Reais do Uso de Intervalos de Confiança II

- Ensaios clínicos para a eficácia de um novo tratamento médico para correção da miopia:
  - Queremos inferir a verdadeira média da redução de diopetrias nos pacientes que se submeteram ao novo tratamento.

Exemplos Reais do Uso de Intervalos de Confiança II

- Ensaios clínicos para a eficácia de um novo tratamento médico para correção da miopia:
  - Queremos inferir a verdadeira média da redução de diopetrias nos pacientes que se submeteram ao novo tratamento.
  - Atentem que é diferente afirmar que a média desta redução é um determinado valor ξ de afirmar que o verdadeiro valor estará entre ζ e η com uma determinada probabilidade.
  - Estamos a considerar uma margem de erro, que corresponde a considerar a variabilidade do estimador para um determinado nível de confiança.

Passos para se construir um Intervalo de Confiança

- 1º Passo: Encontrar um estimador pontual propriedades adequadas.
- 2º Passo: Estabelecer um nível de confiança λ.

Passos para se construir um Intervalo de Confiança

- 1º Passo: Encontrar um estimador pontual propriedades adequadas.
- 2º Passo: Estabelecer um nível de confiança λ.
- 3º Passo: Conhecer a dimensão da amostra n.

Passos para se construir um Intervalo de Confiança

- 1º Passo: Encontrar um estimador pontual propriedades adequadas.
- 2º Passo: Estabelecer um nível de confiança λ.
- 3º Passo: Conhecer a dimensão da amostra n.
- 4º Passo: Conhecer as distribuições amostrais das variáveis aletórias (i.e., do estimador pontual) utilizadas para estimar o parâmetro.
- Este passos culminam no que se chama de método da variável fulcral, onde o cerne da questão como vamos verificar é a escolha do estimador pontual (i.e., variável fulcral) para se construir o intervalo.

Quais as são as condições de escolha da variável fulcral ?

- A escolha da variável fulcral (i.e., do estimador pontual) bem sucedida obdece aos seguintes pontos:

Quais as são as condições de escolha da variável fulcral ?

- A escolha da variável fulcral (i.e., do estimador pontual) bem sucedida obdece aos seguintes pontos:
- 1º - Tem de conter o parâmetro a estimar na sua expressão.

Quais as são as condições de escolha da variável fulcral ?

- A escolha da variável fulcral (i.e., do estimador pontual) bem sucedida obdece aos seguintes pontos:
- 1º - Tem de conter o parâmetro a estimar na sua expressão.
- 2º - Não contenha na sua expressão quaisquer outros parâmetros desconhecidos.

Quais as são as condições de escolha da variável fulcral ?

- A escolha da variável fulcral (i.e., do estimador pontual) bem sucedida obdece aos seguintes pontos:
- 1º - Tem de conter o parâmetro a estimar na sua expressão.
- 2º - Não contenha na sua expressão quaisquer outros parâmetros desconhecidos.
- 3º - A sua distribuição não pode depender do parâmetro a estimar, nem de quaisquer outros valores que se desconheçam.

Quais são as possibilidades de variáveis fulcrais ?

- Esta tabela de variáveis fulcrais vai ser dada nos momentos de avaliação. Neste sentido só têm de saber o que procurar.

Quais são as possibilidades de variáveis fulcrais ?

- Esta tabela de variáveis fulcrais vai ser dada nos momentos de avaliação. Neste sentido só têm de saber o que procurar.
  ▸ Link Tabelas no GitHub
- A escolha da variável fulcral depende do problema e dos dados disponíveis no mesmo.

O Intervalo de Confiança mais Simples

- Exemplo de um intervalo de confiança para a média de uma população com variância conhecida.

O Intervalo de Confiança mais Simples

- Exemplo de um intervalo de confiança para a média de uma população com variância conhecida.
- 2º Passo - Estabelecer um nível de confiança λ.

O Intervalo de Confiança mais Simples

- Exemplo de um intervalo de confiança para a média de uma população com variância conhecida.
- 2º Passo - Estabelecer um nível de confiança λ.
- Vamos considerar um nível de confiança de 95%.

O Intervalo de Confiança mais Simples

- Exemplo de um intervalo de confiança para a média de uma população com variância conhecida.
- 2º Passo - Estabelecer um nível de confiança λ.
- Vamos considerar um nível de confiança de 95%.
- 3º Passo - Conhecer a dimensão da amostra n.

O Intervalo de Confiança mais Simples

- Exemplo de um intervalo de confiança para a média de uma população com variância conhecida.
- 2º Passo - Estabelecer um nível de confiança λ.
- Vamos considerar um nível de confiança de 95%.
- 3º Passo - Conhecer a dimensão da amostra n.
- Vamos considerar uma amostra de dimensão $n = 100$.

O Intervalo de Confiança mais Simples

- Exemplo de um intervalo de confiança para a média de uma população com variância conhecida.
- 2º Passo - Estabelecer um nível de confiança λ.
- Vamos considerar um nível de confiança de 95%.
- 3º Passo - Conhecer a dimensão da amostra n.
- Vamos considerar uma amostra de dimensão $n = 100$.
- Agora vamos ao quadro para encontrar uma variável fulcral que obedeça aos Passos seguintes:

O Intervalo de Confiança mais Simples

- Exemplo de um intervalo de confiança para a média de uma população com variância conhecida.
- 2º Passo - Estabelecer um nível de confiança λ.
- Vamos considerar um nível de confiança de 95%.
- 3º Passo - Conhecer a dimensão da amostra n.
- Vamos considerar uma amostra de dimensão $n = 100$.
- Agora vamos ao quadro para encontrar uma variável fulcral que obedeça aos Passos seguintes:
- 1º Passo - Encontrar um estimador pontual propriedades adequadas.

O Intervalo de Confiança mais Simples

- Exemplo de um intervalo de confiança para a média de uma população com variância conhecida.
- 2º Passo - Estabelecer um nível de confiança λ.
- Vamos considerar um nível de confiança de 95%.
- 3º Passo - Conhecer a dimensão da amostra n.
- Vamos considerar uma amostra de dimensão $n = 100$.
- Agora vamos ao quadro para encontrar uma variável fulcral que obedeça aos Passos seguintes:
- 1º Passo - Encontrar um estimador pontual propriedades adequadas.
- 4º Passo - Conhecer as distribuições amostrais das variáveis aletórias (i.e., do estimador pontual) utilizadas para estimar o parâmetro.

O Intervalo de Confiança mais Simples

- Assim sendo, ao analisar a tabela, a variável fulcral que vamos utilizar é:

O Intervalo de Confiança mais Simples

- Assim sendo, ao analisar a tabela, a variável fulcral que vamos utilizar é:

$$\bar{X} = \sum_{i=1}^{n} \frac{1}{n} X_i \sim \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right), \tag{6}$$

O Intervalo de Confiança mais Simples

- Assim sendo, ao analisar a tabela, a variável fulcral que vamos utilizar é:

$$\bar{X} = \sum_{i=1}^{n} \frac{1}{n} X_i \sim \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right), \tag{6}$$

- Assim sendo também é verdade que:

O Intervalo de Confiança mais Simples

- Assim sendo, ao analisar a tabela, a variável fulcral que vamos utilizar é:

$$\bar{X} = \sum_{i=1}^{n} \frac{1}{n} X_i \sim \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right), \qquad (6)$$

- Assim sendo também é verdade que:

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1). \qquad (7)$$

O Intervalo de Confiança mais Simples

- Assim sendo, ao analisar a tabela, a variável fulcral que vamos utilizar é:

$$\bar{X} = \sum_{i=1}^{n} \frac{1}{n} X_i \sim \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right), \quad (6)$$

- Assim sendo também é verdade que:

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1). \quad (7)$$

- Esta variável fulcral cumpre todos os critérios anteriores.

O Intervalo de Confiança mais Simples

- Por definição de intervalo de confiança simétrico, temos que:

O Intervalo de Confiança mais Simples

- Por definição de intervalo de confiança simétrico, temos que:

$$\mathbb{P}\left[\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| < \text{c.v}^{95\%}\right] = 0.95 \tag{8}$$

O Intervalo de Confiança mais Simples

- Por definição de intervalo de confiança simétrico, temos que:

$$\mathbb{P}\left[\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| < \text{c.v}^{95\%}\right] = 0.95 \tag{8}$$

- Concretizar a probabilidade anterior só é possivel conhecendo a distribuição da variável fulcral.

O Intervalo de Confiança mais Simples

- Por definição de intervalo de confiança simétrico, temos que:

$$\mathbb{P}\left[\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| < \text{c.v}^{95\%}\right] = 0.95 \tag{8}$$

- Concretizar a probabilidade anterior só é possivel conhecendo a distribuição da variável fulcral.
- Massajando a inequação dentro da probabilidade, obtemos:

O Intervalo de Confiança mais Simples

- Por definição de intervalo de confiança simétrico, temos que:

$$\mathbb{P}\left[\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| < \mathrm{c.v}^{95\%}\right] = 0.95 \quad (8)$$

- Concretizar a probabilidade anterior só é possivel conhecendo a distribuição da variável fulcral.
- Massajando a inequação dentro da probabilidade, obtemos:

$$\mathbb{P}\left[\bar{X} - \mathrm{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + \mathrm{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}}\right] = 0.95 \quad (9)$$

O Intervalo de Confiança mais Simples

- A outra maneira de escrever o intervalo **aleatório** é:

O Intervalo de Confiança mais Simples

- A outra maneira de escrever o intervalo **aleatório** é:

$$]IC_{95\%}[_{\mu} = \left]\bar{X} - \text{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}}; \bar{X} + \text{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}}\right[\quad (10)$$

O Intervalo de Confiança mais Simples

- A outra maneira de escrever o intervalo **aleatório** é:

$$]IC_{95\%}[_{\mu} = \left] \bar{X} - \mathrm{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}} ; \bar{X} + \mathrm{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}} \right[\qquad (10)$$

- Neste momento, a interpretação é a probabilidade de o verdadeiro valor da média da população estar contido no intervalo anterior é de 95%.

O Intervalo de Confiança mais Simples

- A outra maneira de escrever o intervalo **aleatório** é:

$$]IC_{95\%}[_{\mu} = \left] \bar{X} - \text{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}} ; \bar{X} + \text{c.v}^{95\%} \times \frac{\sigma}{\sqrt{n}} \right[\qquad (10)$$

- Neste momento, a interpretação é a probabilidade de o verdadeiro valor da média da população estar contido no intervalo anterior é de 95%.
- Este intervalo é aleatório porque depende da amostra aleatória $(X_1, ..., X_n)$ contida na média amostral $\bar{X}$.

O Intervalo de Confiança mais Simples

- Se agora quiser-mos calcular um intervalo de confiança concreto, basta substituir a realização da amostra aleatória.

O Intervalo de Confiança mais Simples

- Se agora quiser-mos calcular um intervalo de confiança concreto, basta substituir a realização da amostra aleatória.
- Vamos considerar que a média amostra é $\bar{x} = 10$ e o desvio-padrão da população é $\sigma = 2$, então o intervalo de confiança concreto é:

$$]IC_{95\%}[^{*}_{\mu} = \left]10 - 1.96 \times \frac{2}{\sqrt{100}}; 10 + 1.96 \times \frac{2}{\sqrt{100}}\right[\tag{11}$$

O Intervalo de Confiança mais Simples

- Se agora quiser-mos calcular um intervalo de confiança concreto, basta substituir a realização da amostra aleatória.
- Vamos considerar que a média amostra é $\bar{x} = 10$ e o desvio-padrão da população é $\sigma = 2$, então o intervalo de confiança concreto é:

$$]IC_{95\%}[^{*}_{\mu} = \left]10 - 1.96 \times \frac{2}{\sqrt{100}}; 10 + 1.96 \times \frac{2}{\sqrt{100}}\right[\quad (11)$$

- Mais especificamente:

$$]IC_{95\%}[^{*}_{\mu} =]10 - 0.392; 10 + 0.392[=]9.608; 10.392[\quad (12)$$

- A única coisa que podemos afirmar é que este intervalo sejam um dos 95 que contêm o verdadeiro valor de μ.

Exercício Nº1 I

Uma estação de rádio quer estimar o tempo médio que uma família dedica, por dia, a ouvir essa rádio. Foi recolhida uma amostra aleatória de 81 famílias, tendo sido calculados uma média diária de audição de 2,4 horas e um desvio-padrão de 0,7 horas. Naquela cidade, quanto tempo dedica, em média, por dia uma família a ouvir aquela rádio? Responda, fornecendo uma estimativa pontual e um intervalo de confiança a 90%. Interprete o significado dos valores encontrados.

Exercício Nº4 I

Na estimação da média de uma população normal por meio de um intervalo de confiança a 90%, qual deve ser a dimensão mínima da amostra para que a amplitude daquele intervalo seja inferior a $\frac{\sigma}{9}$, sendo σ conhecido ?

Exercício Nº6 I

A fábrica de bolachas BEMBOM pensa adquirir uma nova máquina de empacotamento automático que, segundo os técnicos da empresa vendedora, asseguram em média o embalamento de 60 pacotes por minuto. A empresa vendedora colocou a máquina à consignação nas instalações da BEMBOM. Em 25 períodos de um minuto cada, contaram-se o número de pacotes embalados, tendo-se obtido os seguintes resultados:

$$\sum_{i=1}^{25} x_i = 1400 \text{ pacotes} \qquad \sum_{i=1}^{25} x_i^2 = 81000 \text{ pacotes}^2$$

(a) Construa um intervalo d e confiança a 95% para o número médio de pacotes embalados por minuto. Será de aceitar a afirmação dos técnicos da empresa vendedora? Justifique.

Exercício Nº6 II

(b) Qual deverá ser a dimensão da amostra, se pretendermos reduzir a amplitude do intervalo para metade?

Exercício Nº9 I

A quantidade de vinho por garrafa, medida em centilitros, é uma variável aleatória com distribuição normal cuja verdadeira média e variância se desconhecem. Recolheu-se uma amostra aleatória de 25 garrafas que forneceu o valor de 0,0625 centilitros^2 como estimativa para a variância amostral corrigida.

(a) Estime o desvio-padrão desta população através de um intervalo de confiança a 90%.

No passado recente têm surgido diversas queixas de consumidores que reclamam não ser correta a informação contida nos rótulos das garrafas que indicam que cada garrafa contém 75 centilitros do precioso néctar. Tal afirmação é desmentida pelo responsável pelo

Exercício Nº9 II

controlo de qualidade da empresa que, com base numa amostra de 25 garrafas, avançou com o seguinte intervalo de confiança:

$$[I_\lambda]^*_\mu =]74,8968; 75,1032[$$

(b) Que valor proporia como estimativa para a quantidade média de vinho por garrafa? Justifique adequadamente a sua resposta.

(c) Qual a margem de erro associada ao intervalo de confiança obtido?

(d) Determine o nível de confiança $\lambda = 1 - \alpha$ utilizado pelo responsável pelo controlo de qualidade na construção daquele intervalo.

Exercício Nº9 III

(e) Comente a veracidade da seguinte afirmação da associação dos direitos dos consumidores ao serem confrontados com o intervalo de confiança apresentado: ≪É nossa convicção que só metade das garrafas comercializadas respeita o limite inferior e superior daquele intervalo≫.

Exercício Nº10 I

Se numa operação STOP na estrada nacional EN1, em 600 carros, 114 tinham o sistema elétrico com deficiências graves, construa um intervalo de confiança a 95% para a verdadeira proporção de carros com deficiências graves no sistema elétrico viajando nessa estrada. Faça os pressupostos que julgar necessários.

Exercício Nº11 I

Uma empresa pretende lançar um novo produto numa cidade de um milhão de habitantes. No estudo de mercado realizado, foram inquiridas 1000 pessoas tendo 800 delas afirmado que muito dificilmente iriam utilizar aquele novo produto.

(a) Quantos habitantes s e espera venham a utilizar o novo produto? Justifique a sua reposta.

(b) Construa um intervalo de confiança a 95% para a verdadeira proporção de habitantes que utilizarão o novo produto. Interprete o resultado a que chegou.

(c) Quantas pessoas serão necessárias acrescentar à amostra recolhida inicialmente se pretender que a margem de erro seja no máximo de 1%?

Exercício Nº17 I

Em dois Centros de Medicina Desportiva, pretendeu-se comparar os pesos médios dos atletas, tendo para isso sido recolhidas duas amostras aleatórias de 10 atletas cada, que forneceram os seguintes resultados: médias de 77 Kg e 68 Kg, respetivamente, com desvios-padrões de 6 Kg e 10 Kg. Determine um intervalo de confiança a 90% para a diferença entre os pesos médios dos atletas nos dois Centros.

Exercício Nº20 I

No âmbito do plano de formação estrutural de uma instituição financeira foram usados dois tipos de métodos de aprendizagem no aperfeiçoamento da língua inglesa. O método A, designado por convencional, foi utilizado em 100 trabalhadores e o método B, designado por moderno, foi utilizado em 120 trabalhadores. No final do curso, os trabalhadores foram sujeitos a uma prova de avaliação tendo-se obtido os seguintes resultados:

	Nº de Aprovações	Nº de Reprovações
Método A	88	12
Método B	105	15

Face a estes resultados, o responsável pelo plano de formação afirma que ≪o método moderno evidenciou ser mais eficaz≫.

Exercício Nº20 II

Comente esta afirmação, utilizando um intervalo de confiança com $\lambda = 0,90$.

Exercícios de Interpretação de Outputs

- No link seguinte temos os exercícios de intervalos de confiança com interpretação de outputs: Link Exercícios de Outputs no GitHub