Ensaios de Hipóteses

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.

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Para que serve um ensaio de hipóteses?

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Para que serve um ensaio de hipóteses?
  - Servem para credibilizar ou descredibilizar afirmações sobre os parâmetros (de uma população).

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Para que serve um ensaio de hipóteses?
  - Servem para credibilizar ou descredibilizar afirmações sobre os parâmetros (de uma população).
  - Então... mas nós conseguimos fazer isso com os intervalos de confiança...

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Para que serve um ensaio de hipóteses?
  - Servem para credibilizar ou descredibilizar afirmações sobre os parâmetros (de uma população).
  - Então... mas nós conseguimos fazer isso com os intervalos de confiança...
  - Verdade, mas não com o nível de granularidade (ou precisão) de um teste de hipóteses.

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Para que serve um ensaio de hipóteses?
  - Servem para credibilizar ou descredibilizar afirmações sobre os parâmetros (de uma população).
  - Então... mas nós conseguimos fazer isso com os intervalos de confiança...
  - Verdade, mas não com o nível de granularidade (ou precisão) de um teste de hipóteses.
- Que tipo de perguntas conseguimos responder?

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Para que serve um ensaio de hipóteses?
  - Servem para credibilizar ou descredibilizar afirmações sobre os parâmetros (de uma população).
  - Então... mas nós conseguimos fazer isso com os intervalos de confiança...
  - Verdade, mas não com o nível de granularidade (ou precisão) de um teste de hipóteses.
- Que tipo de perguntas conseguimos responder?

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Para que serve um ensaio de hipóteses?
  - Servem para credibilizar ou descredibilizar afirmações sobre os parâmetros (de uma população).
  - Então... mas nós conseguimos fazer isso com os intervalos de confiança...
  - Verdade, mas não com o nível de granularidade (ou precisão) de um teste de hipóteses.
- Que tipo de perguntas conseguimos responder?
  - "O Ministério da Saúde afirma que, com os meios agora postos à disposição dos hospitais civis, o número médio de dias de internamento é, no máximo, oito."

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Para que serve um ensaio de hipóteses?
  - Servem para credibilizar ou descredibilizar afirmações sobre os parâmetros (de uma população).
  - Então... mas nós conseguimos fazer isso com os intervalos de confiança...
  - Verdade, mas não com o nível de granularidade (ou precisão) de um teste de hipóteses.
- Que tipo de perguntas conseguimos responder?
  - "O Ministério da Saúde afirma que, com os meios agora postos à disposição dos hospitais civis, o número médio de dias de internamento é, no máximo, oito."
  - Seja $X =$ Número de dias que um doente fica internado no hospital.

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Para que serve um ensaio de hipóteses?
  - Servem para credibilizar ou descredibilizar afirmações sobre os parâmetros (de uma população).
  - Então... mas nós conseguimos fazer isso com os intervalos de confiança...
  - Verdade, mas não com o nível de granularidade (ou precisão) de um teste de hipóteses.
- Que tipo de perguntas conseguimos responder?
  - "O Ministério da Saúde afirma que, com os meios agora postos à disposição dos hospitais civis, o número médio de dias de internamento é, no máximo, oito."
  - Seja $X =$ Número de dias que um doente fica internado no hospital.
  - $X \sim \mathcal{N}(\mu, \sigma)$, assumindo uma população normal.

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Para que serve um ensaio de hipóteses?
  - Servem para credibilizar ou descredibilizar afirmações sobre os parâmetros (de uma população).
  - Então... mas nós conseguimos fazer isso com os intervalos de confiança...
  - Verdade, mas não com o nível de granularidade (ou precisão) de um teste de hipóteses.
- Que tipo de perguntas conseguimos responder?
  - "O Ministério da Saúde afirma que, com os meios agora postos à disposição dos hospitais civis, o número médio de dias de internamento é, no máximo, oito."
  - Seja $X =$ Número de dias que um doente fica internado no hospital.
  - $X \sim \mathcal{N}(\mu, \sigma)$, assumindo uma população normal.
  - Portanto, matematicamente, será que $\mu \leq 8$?

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Uma outra situação:

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Uma outra situação:
  - "Pretendem comparar-se dois processos de fabrico do mesmo produto. Adopta-se a seguinte regra de decisão: com base numa amostra de 100 unidades para cada processo, eliminar-se-á aquele processo que conduza a uma proporção observada de produtos defeituosos superior à do outro, em pelo menos 2%"

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Uma outra situação:
  - "Pretendem comparar-se dois processos de fabrico do mesmo produto. Adopta-se a seguinte regra de decisão: com base numa amostra de 100 unidades para cada processo, eliminar-se-á aquele processo que conduza a uma proporção observada de produtos defeituosos superior à do outro, em pelo menos 2%"
  - $Y_i =$ Unidades defeituosas do processo de fabrico i, $i \in \{1, 2\}$

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Uma outra situação:
  - "Pretendem comparar-se dois processos de fabrico do mesmo produto. Adopta-se a seguinte regra de decisão: com base numa amostra de 100 unidades para cada processo, eliminar-se-á aquele processo que conduza a uma proporção observada de produtos defeituosos superior à do outro, em pelo menos 2%"
  - $Y_i =$ Unidades defeituosas do processo de fabrico i, $i \in \{1, 2\}$
  - $Y_i \sim B(p_i)$, p_i é a proporção de produtos defeituosos.

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Uma outra situação:
  - "Pretendem comparar-se dois processos de fabrico do mesmo produto. Adopta-se a seguinte regra de decisão: com base numa amostra de 100 unidades para cada processo, eliminar-se-á aquele processo que conduza a uma proporção observada de produtos defeituosos superior à do outro, em pelo menos 2%"
  - $Y_i =$ Unidades defeituosas do processo de fabrico i, $i \in \{1, 2\}$
  - $Y_i \sim B(p_i)$, p_i é a proporção de produtos defeituosos.
  - Portanto, matematicamente, será que $p_1 - p_2 > 0.02$ ou $p_2 - p_1 > 0.02$?

Teoria dos Ensaios de Hipóteses

Qual é a utilidade?

- Uma outra situação:
  - "Pretendem comparar-se dois processos de fabrico do mesmo produto. Adopta-se a seguinte regra de decisão: com base numa amostra de 100 unidades para cada processo, eliminar-se-á aquele processo que conduza a uma proporção observada de produtos defeituosos superior à do outro, em pelo menos 2%"
  - Y_i = Unidades defeituosas do processo de fabrico i, $i \in \{1, 2\}$
  - $Y_i \sim B(p_i)$, p_i é a proporção de produtos defeituosos.
  - Portanto, matematicamente, será que $p_1 - p_2 > 0.02$ ou $p_2 - p_1 > 0.02$?
- São apenas dois exemplos onde os intervalos de confiança não conseguem responder com a mesma eficácia. Conseguem dar uma ideia mas não permitem fazer a pergunta desta forma tão direta.

As Hipóteses e os Erros

Quais são as hipóteses?

- Considerem um tribunal em que uma dada acusação contra um réu é credibilizada ou descredibilizada mediante a apresentação de provas por parte da defesa e da acusação.

As Hipóteses e os Erros

Quais são as hipóteses?

- Considerem um tribunal em que uma dada acusação contra um réu é credibilizada ou descredibilizada mediante a apresentação de provas por parte da defesa e da acusação.
- No final, o juiz decide com base na evidência.

As Hipóteses e os Erros

Quais são as hipóteses?

- Considerem um tribunal em que uma dada acusação contra um réu é credibilizada ou descredibilizada mediante a apresentação de provas por parte da defesa e da acusação.
- No final, o juiz decide com base na evidência.
- Um ensaio de hipóteses inspira-se na mesma envolvência com umas diferenças:

As Hipóteses e os Erros

Quais são as hipóteses?

- Considerem um tribunal em que uma dada acusação contra um réu é credibilizada ou descredibilizada mediante a apresentação de provas por parte da defesa e da acusação.
- No final, o juiz decide com base na evidência.
- Um ensaio de hipóteses inspira-se na mesma envolvência com umas diferenças:
  - A prova é a amostra aleatória.

As Hipóteses e os Erros

Quais são as hipóteses?

- Considerem um tribunal em que uma dada acusação contra um réu é credibilizada ou descredibilizada mediante a apresentação de provas por parte da defesa e da acusação.
- No final, o juiz decide com base na evidência.
- Um ensaio de hipóteses inspira-se na mesma envolvência com umas diferenças:
  - A prova é a amostra aleatória.
  - O juiz é o teste estatístico que foi escolhido (i.e. a estatística de teste do formulário).

As Hipóteses e os Erros

Quais são as hipóteses?

- Considerem um tribunal em que uma dada acusação contra um réu é credibilizada ou descredibilizada mediante a apresentação de provas por parte da defesa e da acusação.
- No final, o juiz decide com base na evidência.
- Um ensaio de hipóteses inspira-se na mesma envolvência com umas diferenças:
  - A prova é a amostra aleatória.
  - O juiz é o teste estatístico que foi escolhido (i.e. a estatística de teste do formulário).
- Assim sendo iremos ter duas hipóteses:

As Hipóteses e os Erros

Quais são as hipóteses?

- Considerem um tribunal em que uma dada acusação contra um réu é credibilizada ou descredibilizada mediante a apresentação de provas por parte da defesa e da acusação.
- No final, o juiz decide com base na evidência.
- Um ensaio de hipóteses inspira-se na mesma envolvência com umas diferenças:
  - A prova é a amostra aleatória.
  - O juiz é o teste estatístico que foi escolhido (i.e. a estatística de teste do formulário).
- Assim sendo iremos ter duas hipóteses:
  - H_0 - A hipótese nula é a presunção de inocência de um réu.

As Hipóteses e os Erros

Quais são as hipóteses?

- Considerem um tribunal em que uma dada acusação contra um réu é credibilizada ou descredibilizada mediante a apresentação de provas por parte da defesa e da acusação.
- No final, o juiz decide com base na evidência.
- Um ensaio de hipóteses inspira-se na mesma envolvência com umas diferenças:
  - A prova é a amostra aleatória.
  - O juiz é o teste estatístico que foi escolhido (i.e. a estatística de teste do formulário).
- Assim sendo iremos ter duas hipóteses:
  - H_0 - A hipótese nula é a presunção de inocência de um réu.
  - H_1 ou H_a - A hipótese alternativa é o contraditório da anterior.

As Hipóteses e os Erros

Quais são as hipóteses?

- Considerem um tribunal em que uma dada acusação contra um réu é credibilizada ou descredibilizada mediante a apresentação de provas por parte da defesa e da acusação.
- No final, o juiz decide com base na evidência.
- Um ensaio de hipóteses inspira-se na mesma envolvência com umas diferenças:
  - A prova é a amostra aleatória.
  - O juiz é o teste estatístico que foi escolhido (i.e. a estatística de teste do formulário).
- Assim sendo iremos ter duas hipóteses:
  - H_0 - A hipótese nula é a presunção de inocência de um réu.
  - H_1 ou H_a - A hipótese alternativa é o contraditório da anterior.
- Quando a realização da amostra (i.e. as provas) é usada numa determinada formulação de teste (i.e. o juiz) podemos ter os seguintes resultados:

As Hipóteses e os Erros

Os Erros Possíveis

- Os dados concretos fornecidos (i.e., as provas) à estatística de teste (i.e., o juiz) **não levam** a uma **rejeição** da hipótese nula (i.e., do status quo).

As Hipóteses e os Erros

Os Erros Possíveis

- Os dados concretos fornecidos (i.e., as provas) à estatística de teste (i.e., o juiz) **não levam** a uma **rejeição** da hipótese nula (i.e., do status quo).
- Os dados fornecidos (i.e., as provas) à estatística de teste (i.e., o juiz) **levam** a uma **rejeição** da hipótese nula (i.e., do status quo).

As Hipóteses e os Erros

Os Erros Possíveis

- Os dados concretos fornecidos (i.e., as provas) à estatística de teste (i.e., o juiz) **não levam** a uma **rejeição** da hipótese nula (i.e., do status quo).
- Os dados fornecidos (i.e., as provas) à estatística de teste (i.e., o juiz) **levam** a uma **rejeição** da hipótese nula (i.e., do status quo).
- Qual é a diferença entre o não rejeitar e o aceitar?

As Hipóteses e os Erros

Os Erros Possíveis

- Os dados concretos fornecidos (i.e., as provas) à estatística de teste (i.e., o juiz) **não levam** a uma **rejeição** da hipótese nula (i.e., do status quo).
- Os dados fornecidos (i.e., as provas) à estatística de teste (i.e., o juiz) **levam** a uma **rejeição** da hipótese nula (i.e., do status quo).
- Qual é a diferença entre o não rejeitar e o aceitar?
- Tal e qual como num julgamento podem existir erros...

As Hipóteses e os Erros

Os Erros Possíveis

- Os dados concretos fornecidos (i.e., as provas) à estatística de teste (i.e., o juiz) **não levam** a uma **rejeição** da hipótese nula (i.e., do status quo).
- Os dados fornecidos (i.e., as provas) à estatística de teste (i.e., o juiz) **levam** a uma **rejeição** da hipótese nula (i.e., do status quo).
- Qual é a diferença entre o não rejeitar e o aceitar?
- Tal e qual como num julgamento podem existir erros...
- Um instrumento de avaliação comum em vários domínios da estatística é a matriz de confusão.

As Hipóteses e os Erros

Erro tipo I e tipo II

Decisão Baseada Nas Provas	Situação Real (i.e. Observada) H_0 Verdadeira	H_0 Falsa
Não Rejeita H_0	Decisão Correta	Erro do Tipo II
Rejeita H_0	Erro do Tipo I	Decisão Correta

As Hipóteses e os Erros

Erro tipo I e tipo II

Decisão Baseada Nas Provas	Situação Real (i.e. Observada) H_0 Verdadeira	H_0 Falsa
Não Rejeita H_0	Decisão Correta	Erro do Tipo II
Rejeita H_0	Erro do Tipo I	Decisão Correta

- Erro Tipo I - A Hipótese nula **é rejeitada** quando, de facto, é verdadeira. - O réu é injustamente condenado.

As Hipóteses e os Erros

Erro tipo I e tipo II

Decisão Baseada Nas Provas	Situação Real (i.e. Observada)	
	H_0 Verdadeira	H_0 Falsa
Não Rejeita H_0	Decisão Correta	Erro do Tipo II
Rejeita H_0	Erro do Tipo I	Decisão Correta

- Erro Tipo I - A Hipótese nula **é rejeitada** quando, de facto, é verdadeira. - O réu é injustamente condenado.
- Erro Tipo II - A Hipótese nula **não é rejeitada** quando, de facto, é falsa. - O réu é injustamente ilibado.

As Hipóteses e os Erros

Erro tipo I e tipo II

Decisão Baseada Nas Provas	Situação Real (i.e. Observada) H_0 Verdadeira	H_0 Falsa
Não Rejeita H_0	Decisão Correta	Erro do Tipo II
Rejeita H_0	Erro do Tipo I	Decisão Correta

- Erro Tipo I - A Hipótese nula **é rejeitada** quando, de facto, é verdadeira. - O réu é injustamente condenado.
- Erro Tipo II - A Hipótese nula **não é rejeitada** quando, de facto, é falsa. - O réu é injustamente ilibado.
- Cada vez que um ensaio de hipóteses é executado os dois erros estão presentes e devem ser controlados.

Execução de Um Ensaio de Hipóteses

Passos Necessários - Problema Exemplo

- Vamos considerar um exemplo para delinear os passos necessários.

Execução de Um Ensaio de Hipóteses

Passos Necessários - Problema Exemplo

- Vamos considerar um exemplo para delinear os passos necessários.
- "Uma pizzaria recebe diariamente encomendas por telefone, que se têm comportado segundo uma lei normal. A empresa está dimensionada para uma procura média diária que não ultrapasse as 200 pizzas, admitindo um desvio-padrão de 15. Uma campanha promocional realizada nos últimos 9 dias levou a uma procura média de 210 pizzas. O problema consiste em avaliar a necessidade de reforçar a capacidade média de venda, estudando se houve de facto uma alteração significativa na procura média diária de pizzas"

Execução de Um Ensaio de Hipóteses

Passo I - Definir as Hipóteses

- Existiu uma quebra de estrutura na procura? Ou seja, tenho de redimensionar a minha loja?

Execução de Um Ensaio de Hipóteses

Passo I - Definir as Hipóteses

- Existiu uma quebra de estrutura na procura? Ou seja, tenho de redimensionar a minha loja?
  - Considere-se $X =$ Procura diária de pizzas no estabelecimento após a campanha promocional.

Execução de Um Ensaio de Hipóteses

Passo I - Definir as Hipóteses

- Existiu uma quebra de estrutura na procura? Ou seja, tenho de redimensionar a minha loja?
  - Considere-se $X =$ Procura diária de pizzas no estabelecimento após a campanha promocional.
  - Sabe-se que $X \sim \mathcal{N}(\mu, \sigma = 15)$. Queremos perceber se μ se alterou.

Execução de Um Ensaio de Hipóteses

Passo I - Definir as Hipóteses

- Existiu uma quebra de estrutura na procura? Ou seja, tenho de redimensionar a minha loja?
  - Considere-se $X =$ Procura diária de pizzas no estabelecimento após a campanha promocional.
  - Sabe-se que $X \sim \mathcal{N}(\mu, \sigma = 15)$. Queremos perceber se μ se alterou.
- Deste modo queremos que os dados respondam a isto:

Execução de Um Ensaio de Hipóteses

Passo I - Definir as Hipóteses

- Existiu uma quebra de estrutura na procura? Ou seja, tenho de redimensionar a minha loja?
  - Considere-se $X =$ Procura diária de pizzas no estabelecimento após a campanha promocional.
  - Sabe-se que $X \sim \mathcal{N}(\mu, \sigma = 15)$. Queremos perceber se μ se alterou.
- Deste modo queremos que os dados respondam a isto:
  - $H_0 : \mu \leq 200$ vs. $H_1 : \mu > 200$ - Teste Unilateral Direito

Execução de Um Ensaio de Hipóteses

Passo I - Definir as Hipóteses

- Existiu uma quebra de estrutura na procura? Ou seja, tenho de redimensionar a minha loja?
  - Considere-se $X =$ Procura diária de pizzas no estabelecimento após a campanha promocional.
  - Sabe-se que $X \sim \mathcal{N}(\mu, \sigma = 15)$. Queremos perceber se μ se alterou.
- Deste modo queremos que os dados respondam a isto:
  - $H_0 : \mu \leq 200$ vs. $H_1 : \mu > 200$ - Teste Unilateral Direito
  - Portanto queremos saber se a procura média se estabeleceu num novo patamar pelo que o COO tem de propor ao board um aumento da capacidade da pizzaria.
- Como vamos ver existem várias formulações de hipóteses (i.e. Testes)

Execução de Um Ensaio de Hipóteses

Passo I - Definir as Hipóteses

- Existiu uma quebra de estrutura na procura? Ou seja, tenho de redimensionar a minha loja?
  - Considere-se $X =$ Procura diária de pizzas no estabelecimento após a campanha promocional.
  - Sabe-se que $X \sim \mathcal{N}(\mu, \sigma = 15)$. Queremos perceber se μ se alterou.
- Deste modo queremos que os dados respondam a isto:
  - $H_0 : \mu \leq 200$ vs. $H_1 : \mu > 200$ - Teste Unilateral Direito
  - Portanto queremos saber se a procura média se estabeleceu num novo patamar pelo que o COO tem de propor ao board um aumento da capacidade da pizzaria.
- Como vamos ver existem várias formulações de hipóteses (i.e. Testes)
- Unilateral $\Rightarrow$ Sinal da alteração $\Rightarrow H_a(< \text{ ou } >)$, $H_0(=, \leq \text{ ou } \geq)$.
- Bilateral $\Rightarrow$ contra um valor concreto $K \Rightarrow H_0(=)$, $H_a(\neq)$.

Execução de Um Ensaio de Hipóteses

Passo II - Definir o nível de significância $\alpha = 1 - \lambda$

- Temos de definir um nível de significância do teste α que se liga diretamente com o nível de confiança que já foi mencionado nos ICs.

Execução de Um Ensaio de Hipóteses

Passo II - Definir o nível de significância $\alpha = 1 - \lambda$

- Temos de definir um nível de significância do teste α que se liga diretamente com o nível de confiança que já foi mencionado nos ICs.
- Este valor estabelece o limite para a região de aceitação (RA) e para a região crítica ou de rejeição (RC).

Execução de Um Ensaio de Hipóteses

Passo II - Definir o nível de significância $\alpha = 1 - \lambda$

- Temos de definir um nível de significância do teste α que se liga diretamente com o nível de confiança que já foi mencionado nos ICs.
- Este valor estabelece o limite para a região de aceitação (RA) e para a região crítica ou de rejeição (RC).
- O nível de significância impõe um trade-off à decisão? Sim.

Execução de Um Ensaio de Hipóteses

Passo II - Definir o nível de significância $\alpha = 1 - \lambda$

- Temos de definir um nível de significância do teste α que se liga diretamente com o nível de confiança que já foi mencionado nos ICs.
- Este valor estabelece o limite para a região de aceitação (RA) e para a região crítica ou de rejeição (RC).
- O nível de significância impõe um trade-off à decisão? Sim.
- O mais standard será $\alpha = 0.05$. Normalmente os papers que fazem uso de um teste de hipóteses mostram a decisão com 3 níveis de significância: 1% (***), 5%(**), 10%(*) e $> 10\%$ ().

Execução de Um Ensaio de Hipóteses

Passo III - Escolha da Estatística Adequada ao Ensaio

- Como é que se escolhe a estatística de teste adequada?

Execução de Um Ensaio de Hipóteses

Passo III - Escolha da Estatística Adequada ao Ensaio

- Como é que se escolhe a estatística de teste adequada?
  - Tal como nos intervalos de confiança, vai depender do que sabemos (i.e. do setup estatístico) - tipo de população, parâmetros conhecidos, dimensão da amostra, etc.

Execução de Um Ensaio de Hipóteses

Passo III - Escolha da Estatística Adequada ao Ensaio

- Como é que se escolhe a estatística de teste adequada?
  - Tal como nos intervalos de confiança, vai depender do que sabemos (i.e. do setup estatístico) - tipo de população, parâmetros conhecidos, dimensão da amostra, etc.
  - Com base no ponto anterior e na definição das hipóteses escolhe-se uma estatística de teste que terá uma distribuição amostral conhecida (Está no formulário).

Execução de Um Ensaio de Hipóteses

Passo III - Escolha da Estatística Adequada ao Ensaio

- Como é que se escolhe a estatística de teste adequada?
  - Tal como nos intervalos de confiança, vai depender do que sabemos (i.e. do setup estatístico) - tipo de população, parâmetros conhecidos, dimensão da amostra, etc.
  - Com base no ponto anterior e na definição das hipóteses escolhe-se uma estatística de teste que terá uma distribuição amostral conhecida (Está no formulário).
  - No nosso caso concreto trata-se de um ensaio unilateral à média da população com variância conhecida pelo que a estatística adequada a todo este setup é:

Execução de Um Ensaio de Hipóteses

Passo III - Escolha da Estatística Adequada ao Ensaio

- Como é que se escolhe a estatística de teste adequada?
  - Tal como nos intervalos de confiança, vai depender do que sabemos (i.e. do setup estatístico) - tipo de população, parâmetros conhecidos, dimensão da amostra, etc.
  - Com base no ponto anterior e na definição das hipóteses escolhe-se uma estatística de teste que terá uma distribuição amostral conhecida (Está no formulário).
  - No nosso caso concreto trata-se de um ensaio unilateral à média da população com variância conhecida pelo que a estatística adequada a todo este setup é:

$$T = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1) \tag{1}$$

Execução de Um Ensaio de Hipóteses

Passo IV - Tomada de Decisão

- Então e agora como é que se decide para uma realização da amostra aleatória (i.e., uma amostra observada)?

Execução de Um Ensaio de Hipóteses

Passo IV - Tomada de Decisão

- Então e agora como é que se decide para uma realização da amostra aleatória (i.e., uma amostra observada)?
- Para o nível de significância do Passo II (i.e $\alpha = 0.05$) calcula-se/calculam-se o/os valor/valores crítico/críticos da distribuição amostral:

Execução de Um Ensaio de Hipóteses

Passo IV - Tomada de Decisão

- Então e agora como é que se decide para uma realização da amostra aleatória (i.e., uma amostra observada)?
- Para o nível de significância do Passo II (i.e $\alpha = 0.05$) calcula-se/calculam-se o/os valor/valores crítico/críticos da distribuição amostral:
- O valor crítico em causa é $Z_{95\%} = 1.645$ - **Ensaio à direita**.

Execução de Um Ensaio de Hipóteses

Passo IV - Tomada de Decisão

- Então e agora como é que se decide para uma realização da amostra aleatória (i.e., uma amostra observada)?
- Para o nível de significância do Passo II (i.e $\alpha = 0.05$) calcula-se/calculam-se o/os valor/valores crítico/críticos da distribuição amostral:
- O valor crítico em causa é $Z_{95\%} = 1.645$ - **Ensaio à direita**.
- Substitui-se os valores do nosso setup estatístico na estatística de teste:

$$T^* = \frac{210 - 200}{\frac{15}{\sqrt{9}}} = \frac{10}{5} = 2. \tag{2}$$

Execução de Um Ensaio de Hipóteses

Passo IV - Tomada de Decisão

- Então e agora como é que se decide para uma realização da amostra aleatória (i.e., uma amostra observada)?
- Para o nível de significância do Passo II (i.e $\alpha = 0.05$) calcula-se/calculam-se o/os valor/valores crítico/críticos da distribuição amostral:
- O valor crítico em causa é $Z_{95\%} = 1.645$ - **Ensaio à direita**.
- Substitui-se os valores do nosso setup estatístico na estatística de teste:

$$T^* = \frac{210 - 200}{\frac{15}{\sqrt{9}}} = \frac{10}{5} = 2. \tag{2}$$

- O valor da estatística para a realização da amostra pertence à região crítica (RC): $\underbrace{2}_{T^*} \in \underbrace{[1.645, +\infty[}_{RC}$.

Execução de Um Ensaio de Hipóteses

Passo IV - Tomada de Decisão

- Então e agora como é que se decide para uma realização da amostra aleatória (i.e., uma amostra observada)?
- Para o nível de significância do Passo II (i.e $\alpha = 0.05$) calcula-se/calculam-se o/os valor/valores crítico/críticos da distribuição amostral:
- O valor crítico em causa é $Z_{95\%} = 1.645$ - **Ensaio à direita**.
- Substitui-se os valores do nosso setup estatístico na estatística de teste:

$$T^* = \frac{210 - 200}{\frac{15}{\sqrt{9}}} = \frac{10}{5} = 2. \tag{2}$$

- O valor da estatística para a realização da amostra pertence à região crítica (RC): $\underbrace{2}_{T^*} \in \underbrace{[1.645, +\infty[}_{RC}$.
- A pizzaria deverá de facto aumentar a sua capacidade.

Execução de Um Ensaio de Hipóteses

Passo IV - Tomada de Decisão

- Um outra maneira de avaliar o teste é calcular o valor crítico (i.e. o valor limite) para a estimativa associada ao parâmetro de interesse.

Execução de Um Ensaio de Hipóteses

Passo IV - Tomada de Decisão

- Um outra maneira de avaliar o teste é calcular o valor crítico (i.e. o valor limite) para a estimativa associada ao parâmetro de interesse.
- No nosso caso é μ e portanto trata-se de um valor limite à estimativa da média amostral $\bar{X}$.

Execução de Um Ensaio de Hipóteses

Passo IV - Tomada de Decisão

- Um outra maneira de avaliar o teste é calcular o valor crítico (i.e. o valor limite) para a estimativa associada ao parâmetro de interesse.
- No nosso caso é μ e portanto trata-se de um valor limite à estimativa da média amostral $\bar{X}$.

$$\frac{\bar{X}_C - 200}{\frac{15}{\sqrt{9}}} = 1.645 \Longrightarrow \bar{X}_C = 208.225 \tag{3}$$

Execução de Um Ensaio de Hipóteses

Passo IV - Tomada de Decisão

- Um outra maneira de avaliar o teste é calcular o valor crítico (i.e. o valor limite) para a estimativa associada ao parâmetro de interesse.
- No nosso caso é μ e portanto trata-se de um valor limite à estimativa da média amostral $\bar{X}$.

$$\frac{\bar{X}_{\mathrm{C}} - 200}{\frac{15}{\sqrt{9}}} = 1.645 \Longrightarrow \bar{X}_{\mathrm{C}} = 208.225 \qquad (3)$$

- Para não rejeitar a hipótese nula (H_0) a estimativa da média amostral teria de ter sido no máximo $\bar{X}_{\mathrm{C}} = 208.225$.

Execução de Um Ensaio de Hipóteses

Passo IV - Tomada de Decisão

- Um outra maneira de avaliar o teste é calcular o valor crítico (i.e. o valor limite) para a estimativa associada ao parâmetro de interesse.
- No nosso caso é μ e portanto trata-se de um valor limite à estimativa da média amostral $\bar{X}$.

$$\frac{\bar{X}_C - 200}{\frac{15}{\sqrt{9}}} = 1.645 \Longrightarrow \bar{X}_C = 208.225 \tag{3}$$

- Para não rejeitar a hipótese nula (H_0) a estimativa da média amostral teria de ter sido no máximo $\bar{X}_C = 208.225$.
- Contudo a amostra observada (i.e., realização da amostra) deu origem a uma estimativa $\bar{X} = 210$, pelo que a evidência estatística leva à rejeição do status quo (i.e. H_0).

Exercício Nº3 - (a)

O Ministério da Saúde afirma que, com os meios agora postos à disposição dos Hospitais Civis, o número médio de dias de internamento é no máximo 15.
Estas declarações foram postas em causa por alguns gestores hospitalares que decidiram proceder à recolha de uma amostra de 225 doentes onde se observou que o número médio de dias de internamento foi de 18.

(a) Terão os gestores hospitalares razão ? Justifique a sua resposta, utilizando um teste adequado a 1% de significância.

Exercício Nº4

Um fabricante de fitas magnéticas para computadores sabe que a resistência à ruptura destas fitas é uma variável aleatória normalmente distribuída com média 300kg e desvio-padrão de 20Kg. Para ajuizar se uma nova técnica/processo de fabrico produz fitas em média mais fracas que as do processo antigo, foi usado o seguinte teste estatístico com um nível de significância de 5% e um tamanho de amostra de $n = 100$:

- $H_0 : \mu_0 = 300kg$ vs $H_a : \mu_a = 295kg$, em que::
- Se $\bar{X} \leq \bar{X}_c \implies$ rejeita-se H_0
- Se $\bar{X} > \bar{X}_c \implies$ não se rejeita H_0

(a) Calcule $\bar{X}_c$

(b) Use este teste, para com base numa amostra de tamanho 100, onde se obteve uma média igual a 290kg, tomar a respectiva decisão.

Exercício Nº8 - (a)

No exame de Estatística efectuado na 2ª época no ano lectivo de 2000/2001, foram avaliados 31 alunos. Considerando este alunos como uma amostra representativa da população dos alunos matriculados na cadeira de Estatística e tendo em conta que, para essa amostra, se obtiveram os seguintes resultados:

$$\sum_{i=1}^{n=31} X_i = 299 \text{ e } \sum_{i=1}^{n=31} (X_i - \bar{X})^2 = 120 \tag{4}$$

Comente a afirmação: A média dos resultados não difere significativamente de 10. Utilize $\alpha = 0.05$

Exercício Nº11

A BUESPEED opera no mercado da União Europeia na área da distribuição de encomendas. A empresa garante que todas as encomendas chegam ao seu destinatário, em média, em menos de 48 horas com uma variabilidade máxima de 8 horas. Para avaliar este desempenho, foram recolhidos os tempos (em horas) relativos a uma amostra de 65 encomendas tendo-se obtido os seguintes resultados:

$$\sum_{i=1}^{65} X_i = 3250 \text{ e } \sum_{i=1}^{65} X_i^2 = 578500 \tag{5}$$

O que se deve concluir sobre o desempenho da BUESPEED ?

Exercício Nº12 - (a)

Um empresa farmacêutica está disposta a lançar no mercado um medicamento, se 90% dos pacientes tratados com esse novo medicamento ficarem curados. Caso verifique que apenas 70% dos pacientes ficam curados, então não lança o novo medicamento. Para tomar uma decisão, a empresa procedeu ao tratamento com o novo medicamente de 50 docentes, tendo-se registado que 45 deles ficaram curados.

(a) Qual deverá ser a decisão tomada pela farmacêutica? Utilize $\alpha = 0.05$.

Medidas dos Erros Num Ensaio de Hipóteses I

- Retomamos a nossa matriz de confusão, agora com as probabilidades associadas a cada uma das situações:

Decisão Baseada Nas Provas	Situação Real (i.e. Observada) H_0 Verdadeira	H_0 Falsa
Não Rejeita H_0	Decisão Correcta: $\mathbb{P} \geq 1 - \alpha$	Erro do Tipo II: $\mathbb{P} = \beta$
Rejeita H_0	Erro do Tipo I: $\mathbb{P} \leq \alpha$	Decisão Correcta: $\mathbb{P} = 1 - \beta$

Medidas dos Erros Num Ensaio de Hipóteses I

- Retomamos a nossa matriz de confusão, agora com as probabilidades associadas a cada uma das situações:

Decisão Baseada Nas Provas	Situação Real (i.e. Observada) H_0 Verdadeira	H_0 Falsa
Não Rejeita H_0	Decisão Correcta: $\mathbb{P} \geq 1 - \alpha$	Erro do Tipo II: $\mathbb{P} = \beta$
Rejeita H_0	Erro do Tipo I: $\mathbb{P} \leq \alpha$	Decisão Correcta: $\mathbb{P} = 1 - \beta$

- Erro Tipo I - A Hipótese nula **é rejeitada** quando, de facto, é verdadeira. - O réu é injustamente condenado.

Medidas dos Erros Num Ensaio de Hipóteses I

- Retomamos a nossa matriz de confusão, agora com as probabilidades associadas a cada uma das situações:

Decisão Baseada Nas Provas	Situação Real (i.e. Observada) H_0 Verdadeira	H_0 Falsa
Não Rejeita H_0	Decisão Correcta: $\mathbb{P} \geq 1 - \alpha$	Erro do Tipo II: $\mathbb{P} = \beta$
Rejeita H_0	Erro do Tipo I: $\mathbb{P} \leq \alpha$	Decisão Correcta: $\mathbb{P} = 1 - \beta$

- Erro Tipo I - A Hipótese nula **é rejeitada** quando, de facto, é verdadeira. - O réu é injustamente condenado.
- $\mathbb{P}[\text{rejeitar } H_0 | H_0 \text{ é verdadeira}] \leq \alpha$.
- Erro Tipo II - A Hipótese nula **não é rejeitada** quando, de facto, é falsa. - O réu é injustamente ilibado.

Medidas dos Erros Num Ensaio de Hipóteses I

- Retomamos a nossa matriz de confusão, agora com as probabilidades associadas a cada uma das situações:

Decisão Baseada Nas Provas	Situação Real (i.e. Observada) H_0 Verdadeira	H_0 Falsa
Não Rejeita H_0	Decisão Correcta: $\mathbb{P} \geq 1 - \alpha$	Erro do Tipo II: $\mathbb{P} = \beta$
Rejeita H_0	Erro do Tipo I: $\mathbb{P} \leq \alpha$	Decisão Correcta: $\mathbb{P} = 1 - \beta$

- Erro Tipo I - A Hipótese nula **é rejeitada** quando, de facto, é verdadeira. - O réu é injustamente condenado.
- $\mathbb{P}[\text{rejeitar } H_0 | H_0 \text{ é verdadeira}] \leq \alpha$.
- Erro Tipo II - A Hipótese nula **não é rejeitada** quando, de facto, é falsa. - O réu é injustamente ilibado.
- $\mathbb{P}[\text{Não rejeitar } H_0 | H_0 \text{ é Falsa}] = \beta$.

Medidas dos Erros Num Ensaio de Hipóteses II

- Cada vez que fazemos um teste destes estamos perante um destes erros, dependendo se rejeitamos ou não H_0.

Medidas dos Erros Num Ensaio de Hipóteses II

- Cada vez que fazemos um teste destes estamos perante um destes erros, dependendo se rejeitamos ou não H_0.
- Vamos calcular cada um deles usando o exemplo da pizzaria.

Medidas dos Erros Num Ensaio de Hipóteses II

- Cada vez que fazemos um teste destes estamos perante um destes erros, dependendo se rejeitamos ou não H_0.
- Vamos calcular cada um deles usando o exemplo da pizzaria.
- Testámos se o efeito da promoção levava a uma alteração da procura média de pizzas: $\mu \leq 200$ vs $\mu > 200$.

Medidas dos Erros Num Ensaio de Hipóteses II

- Cada vez que fazemos um teste destes estamos perante um destes erros, dependendo se rejeitamos ou não H_0.
- Vamos calcular cada um deles usando o exemplo da pizzaria.
- Testámos se o efeito da promoção levava a uma alteração da procura média de pizzas: $\mu \leq 200$ vs $\mu > 200$.
- Como formulámos o teste a $\alpha = 0.05$ e rejeitámos H_0 ficamos sujeitos à probabilidade (i.e. medida) do erro *tipo I*:

Medidas dos Erros Num Ensaio de Hipóteses II

- Cada vez que fazemos um teste destes estamos perante um destes erros, dependendo se rejeitamos ou não H_0.
- Vamos calcular cada um deles usando o exemplo da pizzaria.
- Testámos se o efeito da promoção levava a uma alteração da procura média de pizzas: $\mu \leq 200$ vs $\mu > 200$.
- Como formulámos o teste a $\alpha = 0.05$ e rejeitámos H_0 ficamos sujeitos à probabilidade (i.e. medida) do erro *tipo I*:

$$\mathbb{P}[\text{Rejeitar } H_0 | H_0 \text{ é Verdadeira}] \leq \alpha \tag{6}$$

Medidas dos Erros Num Ensaio de Hipóteses II

- Cada vez que fazemos um teste destes estamos perante um destes erros, dependendo se rejeitamos ou não H_0.
- Vamos calcular cada um deles usando o exemplo da pizzaria.
- Testámos se o efeito da promoção levava a uma alteração da procura média de pizzas: $\mu \leq 200$ vs $\mu > 200$.
- Como formulámos o teste a $\alpha = 0.05$ e rejeitámos H_0 ficamos sujeitos à probabilidade (i.e. medida) do erro *tipo I*:

$$\mathbb{P}[\text{Rejeitar } H_0 | H_0 \text{ é Verdadeira}] \leq \alpha \tag{6}$$

- Para o exemplo em causa, esta probabilidade para diferentes valores de μ, para o quais H_0 é verdadeira, é dada por:

Medidas dos Erros Num Ensaio de Hipóteses II

- Cada vez que fazemos um teste destes estamos perante um destes erros, dependendo se rejeitamos ou não H_0.
- Vamos calcular cada um deles usando o exemplo da pizzaria.
- Testámos se o efeito da promoção levava a uma alteração da procura média de pizzas: $\mu \leq 200$ vs $\mu > 200$.
- Como formulámos o teste a $\alpha = 0.05$ e rejeitámos H_0 ficamos sujeitos à probabilidade (i.e. medida) do erro *tipo I*:

$$\mathbb{P}[\text{Rejeitar } H_0 | H_0 \text{ é Verdadeira}] \leq \alpha \tag{6}$$

- Para o exemplo em causa, esta probabilidade para diferentes valores de μ, para o quais H_0 é verdadeira, é dada por:
- $\mathbb{P}[\bar{X} > 208.225 | \mu = 200] = \mathbb{P}\left[\frac{\bar{X}-200}{\frac{15}{\sqrt{9}}} > \frac{208.225-200}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z > 1.645] = 0.05$

Medidas dos Erros Num Ensaio de Hipóteses III

- $\mathbb{P}[\bar{X} > 208.225|\mu = 199] = \mathbb{P}\left[\frac{\bar{X}-199}{\frac{15}{\sqrt{9}}} > \frac{208.225-199}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z > 1.845] = 0.03255$

Medidas dos Erros Num Ensaio de Hipóteses III

- $\mathbb{P}[\bar{X} > 208.225|\mu = 199] = \mathbb{P}\left[\frac{\bar{X}-199}{\frac{15}{\sqrt{9}}} > \frac{208.225-199}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z > 1.845] = 0.03255$
- $\mathbb{P}[\bar{X} > 208.225|\mu = 195] = \mathbb{P}\left[\frac{\bar{X}-195}{\frac{15}{\sqrt{9}}} > \frac{208.225-195}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z > 1.845] = 0.00405$
- Quanto maior a diferença face ao valor da hipótese nula (i.e., 200) , menor probabilidade de erro. Porquê ?

Medidas dos Erros Num Ensaio de Hipóteses III

- $\mathbb{P}[\bar{X} > 208.225|\mu = 199] = \mathbb{P}\left[\frac{\bar{X}-199}{\frac{15}{\sqrt{9}}} > \frac{208.225-199}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z > 1.845] = 0.03255$
- $\mathbb{P}[\bar{X} > 208.225|\mu = 195] = \mathbb{P}\left[\frac{\bar{X}-195}{\frac{15}{\sqrt{9}}} > \frac{208.225-195}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z > 1.845] = 0.00405$
- Quanto maior a diferença face ao valor da hipótese nula (i.e., 200) , menor probabilidade de erro. Porquê ?
- É mais difícil de errar quando o verdaediro valor de μ é tão díspar da hipótese considerada.

Medidas dos Erros Num Ensaio de Hipóteses III

- $\mathbb{P}[\bar{X} > 208.225|\mu = 199] = \mathbb{P}\left[\frac{\bar{X}-199}{\frac{15}{\sqrt{9}}} > \frac{208.225-199}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z > 1.845] = 0.03255$
- $\mathbb{P}[\bar{X} > 208.225|\mu = 195] = \mathbb{P}\left[\frac{\bar{X}-195}{\frac{15}{\sqrt{9}}} > \frac{208.225-195}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z > 1.845] = 0.00405$
- Quanto maior a diferença face ao valor da hipótese nula (i.e., 200) , menor probabilidade de erro. Porquê ?
- É mais difícil de errar quando o verdaediro valor de μ é tão díspar da hipótese considerada.
- Então a probabilidade do Erro Tipo I é dada por:

$$\mathbb{P}\left[\text{Erro Tipo I}\right] = \alpha(\mu), \forall \mu \in \Theta_0 \tag{7}$$

- Onde Θ_0 é o conjunto de valores que torna H_0 verdadeira.

Medidas dos Erros Num Ensaio de Hipóteses IV

- Vamos agora supor que $\alpha = 0.01$, pelo que H_0 não é rejeitada, dado que a amostra $\bar{X} = 210$ e $\bar{X}_c = 211.63$.
- Ao contrário da situação anterior ficamos sujeitos ao erro tipo II, cuja probabilidade é dada por:

$$\mathbb{P}\left[\text{Não Rejeitar } H_0 | H_0 \text{ é falsa}\right] = \beta \tag{8}$$

Medidas dos Erros Num Ensaio de Hipóteses IV

- Vamos agora supor que $\alpha = 0.01$, pelo que H_0 não é rejeitada, dado que a amostra $\bar{X} = 210$ e $\bar{X}_c = 211.63$.
- Ao contrário da situação anterior ficamos sujeitos ao erro tipo II, cuja probabilidade é dada por:

$$\mathbb{P}\left[\text{Não Rejeitar } H_0 | H_0 \text{ é falsa}\right] = \beta \tag{8}$$

- Para este exemplo concreto, esta probabilidade para vários valores de μ, para os quais H_1 é verdadeira, é dada por:
- $\mathbb{P}[\bar{X} < 211.63 | \mu = 200] = \mathbb{P}\left[\frac{\bar{X}-200}{\frac{15}{\sqrt{9}}} < \frac{211.63-200}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z < 2.233] = 0.99$

Medidas dos Erros Num Ensaio de Hipóteses V

- $\mathbb{P}[\bar{X} < 211.63|\mu = 205] = \mathbb{P}\left[\frac{\bar{X}-205}{\frac{15}{\sqrt{9}}} < \frac{211.63-205}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z < 1.326] = 0.90756$

Medidas dos Erros Num Ensaio de Hipóteses V

- $\mathbb{P}[\bar{X} < 211.63 | \mu = 205] = \mathbb{P}\left[\frac{\bar{X}-205}{\frac{15}{\sqrt{9}}} < \frac{211.63-205}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z < 1.326] = 0.90756$
- $\mathbb{P}[\bar{X} < 211.63 | \mu = 210] = \mathbb{P}\left[\frac{\bar{X}-210}{\frac{15}{\sqrt{9}}} < \frac{211.63-210}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z < 0.326] = 0.62778$

Medidas dos Erros Num Ensaio de Hipóteses V

- $\mathbb{P}[\bar{X} < 211.63|\mu = 205] = \mathbb{P}\left[\frac{\bar{X}-205}{\frac{15}{\sqrt{9}}} < \frac{211.63-205}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z < 1.326] = 0.90756$
- $\mathbb{P}[\bar{X} < 211.63|\mu = 210] = \mathbb{P}\left[\frac{\bar{X}-210}{\frac{15}{\sqrt{9}}} < \frac{211.63-210}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z < 0.326] = 0.62778$
- $\mathbb{P}[\bar{X} < 211.63|\mu = 215] = \mathbb{P}\left[\frac{\bar{X}-215}{\frac{15}{\sqrt{9}}} < \frac{211.63-215}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z < -0.674] = 0.250216$

Medidas dos Erros Num Ensaio de Hipóteses V

- $\mathbb{P}[\bar{X} < 211.63|\mu = 205] = \mathbb{P}\left[\frac{\bar{X}-205}{\frac{15}{\sqrt{9}}} < \frac{211.63-205}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z < 1.326] = 0.90756$
- $\mathbb{P}[\bar{X} < 211.63|\mu = 210] = \mathbb{P}\left[\frac{\bar{X}-210}{\frac{15}{\sqrt{9}}} < \frac{211.63-210}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z < 0.326] = 0.62778$
- $\mathbb{P}[\bar{X} < 211.63|\mu = 215] = \mathbb{P}\left[\frac{\bar{X}-215}{\frac{15}{\sqrt{9}}} < \frac{211.63-215}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z < -0.674] = 0.250216$
- $\mathbb{P}[\bar{X} < 211.63|\mu = 220] = \mathbb{P}\left[\frac{\bar{X}-220}{\frac{15}{\sqrt{9}}} < \frac{211.63-220}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z < -1.674] = 0.0471$
- Quanto maior a diferença face ao valor da hipótese nula (i.e., 200) , menor probabilidade de erro. Porquê ?

Medidas dos Erros Num Ensaio de Hipóteses V

- $\mathbb{P}[\bar{X} < 211.63|\mu = 205] = \mathbb{P}\left[\frac{\bar{X}-205}{\frac{15}{\sqrt{9}}} < \frac{211.63-205}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z < 1.326] = 0.90756$
- $\mathbb{P}[\bar{X} < 211.63|\mu = 210] = \mathbb{P}\left[\frac{\bar{X}-210}{\frac{15}{\sqrt{9}}} < \frac{211.63-210}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z < 0.326] = 0.62778$
- $\mathbb{P}[\bar{X} < 211.63|\mu = 215] = \mathbb{P}\left[\frac{\bar{X}-215}{\frac{15}{\sqrt{9}}} < \frac{211.63-215}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z < -0.674] = 0.250216$
- $\mathbb{P}[\bar{X} < 211.63|\mu = 220] = \mathbb{P}\left[\frac{\bar{X}-220}{\frac{15}{\sqrt{9}}} < \frac{211.63-220}{\frac{15}{\sqrt{9}}}\right] = \mathbb{P}[Z < -1.674] = 0.0471$
- Quanto maior a diferença face ao valor da hipótese nula (i.e., 200) , menor probabilidade de erro. Porquê ?
- É mais difícil de errar quando o verdadeiro valor de μ é tão díspar da hipótese considerada.

Medidas dos Erros Num Ensaio de Hipóteses VI

- Então a probabilidade do erro tipo II é dada por:

Medidas dos Erros Num Ensaio de Hipóteses VI

- Então a probabilidade do erro tipo II é dada por:

$$\mathbb{P}\left[\text{Não Rejeitar } H_0 | H_0 \text{ é falsa}\right] = \beta(\mu_1), \forall \mu_1 \in \Theta_1, \tag{9}$$

Medidas dos Erros Num Ensaio de Hipóteses VI

- Então a probabilidade do erro tipo II é dada por:

$$\mathbb{P}\left[\text{Não Rejeitar } H_0 | H_0 \text{ é falsa}\right] = \beta(\mu_1), \forall \mu_1 \in \Theta_1, \qquad (9)$$

- Onde Θ_1 é o conjunto de valores que torna H_1 verdadeira.

Medidas dos Erros Num Ensaio de Hipóteses VI

- Então a probabilidade do erro tipo II é dada por:

$$\mathbb{P}\left[\text{Não Rejeitar } H_0 | H_0 \text{ é falsa}\right] = \beta(\mu_1), \forall \mu_1 \in \Theta_1, \tag{9}$$

- Onde Θ_1 é o conjunto de valores que torna H_1 verdadeira.
- Existe alguma maneira de reduzir estes erros?

Medidas dos Erros Num Ensaio de Hipóteses VI

- Então a probabilidade do erro tipo II é dada por:
$$\mathbb{P}\left[\text{Não Rejeitar } H_0 | H_0 \text{ é falsa}\right] = \beta(\mu_1), \forall \mu_1 \in \Theta_1, \qquad (9)$$

- Onde Θ_1 é o conjunto de valores que torna H_1 verdadeira.
- Existe alguma maneira de reduzir estes erros?
- Só é possível diminuir os dois erros simultaneamente se pudermos aumentar a dimensão da amostra que implica mais custos.

Medidas dos Erros Num Ensaio de Hipóteses VI

- Então a probabilidade do erro tipo II é dada por:

$$\mathbb{P}\left[\text{Não Rejeitar } H_0 | H_0 \text{ é falsa}\right] = \beta(\mu_1), \forall \mu_1 \in \Theta_1, \qquad (9)$$

- Onde Θ_1 é o conjunto de valores que torna H_1 verdadeira.
- Existe alguma maneira de reduzir estes erros?
- Só é possível diminuir os dois erros simultaneamente se pudermos aumentar a dimensão da amostra que implica mais custos.
- Diminuir a probabilidade de um tipo de erro, implica aumentar do outro.

Medidas dos Erros Num Ensaio de Hipóteses VI

- Então a probabilidade do erro tipo II é dada por:
$$\mathbb{P}\left[\text{Não Rejeitar } H_0 | H_0 \text{ é falsa}\right] = \beta(\mu_1), \forall \mu_1 \in \Theta_1, \qquad (9)$$
- Onde Θ_1 é o conjunto de valores que torna H_1 verdadeira.
- Existe alguma maneira de reduzir estes erros?
- Só é possível diminuir os dois erros simultaneamente se pudermos aumentar a dimensão da amostra que implica mais custos.
- Diminuir a probabilidade de um tipo de erro, implica aumentar do outro.
- Se α diminui/aumenta $\Longrightarrow$ região crítica diminui/aumenta $\Longrightarrow$ região de aceitação aumenta/diminui $\Longrightarrow \beta$ aumenta/diminui.

Função Potência do Ensaio

- Precisamos então de uma medida que nos indique que o nosso teste tem qualidade, isto é, se tem capacidade de avaliar quais das hipóteses está certa através de uma amostra.
- Para isto calcula-se o que se chama da função potência do ensaio $\pi(\mu)$.
- É a probabilidade complementar do erro tipo II e como tal é obtida da seguinte forma:

$$\pi(\mu) = \mathbb{P}\left[\text{Rejeitar } H_0 | H_0 \text{ é falsa}\right] = 1 - \mathbb{P}\left[\text{Não Rejeitar } H_0 | H_0 \text{ é falsa}\right] = 1 - \beta(\mu_1) \quad (10)$$

- Com $\mu_1 \in \Theta_1$, tal como anteriormente Θ_1 é o conjunto de todos os valores que tornam H_0 falsa.

Exercício Nº3 - (b), (c) e (d)

O Ministério da Saúde afirma que, com os meios agora postos à disposição dos Hospitais Civis, o número médio de dias de internamento é no máximo 15.
Estas declarações foram postas em causa por alguns gestores hospitalares que decidiram proceder à recolha de uma amostra de 225 doentes onde se observou que o número médio de dias de internamento foi de 18.

(b) Na decisão que tomou, qual a probabilidade de estar a cometer um erro ?

(c) Com que probabilidade é dada razão aos gestores hospitalares, se o verdadeiro valor médio de dias de internamento for 117 ?

(d) Como variaria aquela probabilidade se a hipótese alternativa fosse superior ao valor especificado na alínea (3) ? E se o tamanho da amostra aumentasse ?

Exercício Nº9

A **Frigo** é uma conceituada marca de camarão congelado. Quando a linha de empacotamento está bem ajustada, as embalagens pesam 250 gramas com uma variabilidade de 15 gramas. Periodicamente, o técnico responsável pela linha de empacotamento recolhe amostras de 20 embalagens e, caso o peso médio encontrado não se situe entre os 243 e 257 gramas, decide que a máquina terá de ser ajustada. Admita que o peso das embalagens é normalmente distribuído.

(a) Se pretender fazer um ensaio de hipóteses sobre o peso médio das embalagens, que hipóteses irá testar?

(b) Qual o nível de significância que está implícito neste problema?

(c) Esboce o gráfico da função potência deste ensaio, calculando $\Pi(260)$.

Exercício Nº14

Uma agência de publicidade desenvolveu determinado tema para um anúncio baseado no pressuposto de que 50% dos espectadores que o vissem tinham mais de 30 anos. A Agência está interessada em saber se, naquele grupo etário, houve alteração do índice de audiência. Para testar essa possibilidade, a agência efectuou uma sondagem sobre 400 espectadores, escolhidos aleatoriamente, dos quais 210 tinham idade superior a 30 anos. O teste foi efectuado ao nível da significância de 50%.

(a) A que conclusão chegou a agência de publicidade? Justifique.

(b) Represente Graficamente a função potência do ensaio para os telespectadores do referido anúncio, sabendo que: para $p = 0.2$ vem $\beta = 0$; para $p = 0.3$ vem $\beta = 0$; para $p = 0.55$ vem $\beta = 0.48$ e para $p = 0.425$ vem $\beta =?$. Determine β.

(c) A agência de publicidade considerou que se a verdadeira proporção de telespectadores que vêem o referido anúncio fosse de 42.5%, o risco associado à não rejeição da hipótese nula ainda seria elevado. A agência pretende diminuir esse risco. Indique, justificando, que tomadas de decisão são possíveis para atingir esse objectivo.

Exercícios de Outputs de SPSS

- Exercícios com outputs sobre ensaios de hipóteses - Parte A.

Exercício Nº17

Exercícios de Outputs de SPSS

- Exercícios com outputs sobre ensaios de hipóteses - Parte B.

Exercício Nº19

Um nutricionista está convencido que a nova dieta que prescreve aos seus doentes é eficaz no tratamento da obsidade provocando perda de peso ao fim de 4 semanas e, contrariamente a outras dietas reduz o estado de ansiedade dos doentes.

Doente obeso	Peso inicial (Kg)	Peso após 4 semanas (Kg)
1	90	86
2	85	85
3	95	92
4	95	90
5	105	100
6	102	95
7	83	80
8	85	81
9	93	90
10	94	88

Com base na amostra recolhida e cosntante no quadro anterior, considera que o nutricionista tem razão ? Utilize 0.05 como nível de significância.

Exercício Nº19 - Resolução I

- Neste caso temos uma amostra antes da dieta e outra depois da dieta

Exercício Nº19 - Resolução I

- Neste caso temos uma amostra antes da dieta e outra depois da dieta
- X_1 - Peso, em kg, dos doentes antes da dieta.

Exercício Nº19 - Resolução I

- Neste caso temos uma amostra antes da dieta e outra depois da dieta
- X_1 - Peso, em kg, dos doentes antes da dieta.
- X_2 - Peso, em kg, dos doentes após 4 semanas de dieta.

Exercício Nº19 - Resolução I

- Neste caso temos uma amostra antes da dieta e outra depois da dieta
- X_1 - Peso, em kg, dos doentes antes da dieta.
- X_2 - Peso, em kg, dos doentes após 4 semanas de dieta.
- Obviamente que nesta situação não podemos considerar que amostras provenientes da popuação X_1 e amostras provenientes da população X_2 sejam independentes porque X_1 e X_2 não são populações independentes...são o mesmo doente!

Exercício Nº19 - Resolução I

- Neste caso temos uma amostra antes da dieta e outra depois da dieta
- X_1 - Peso, em kg, dos doentes antes da dieta.
- X_2 - Peso, em kg, dos doentes após 4 semanas de dieta.
- Obviamente que nesta situação não podemos considerar que amostras provenientes da popuação X_1 e amostras provenientes da população X_2 sejam independentes porque X_1 e X_2 não são populações independentes...são o mesmo doente!
- Como nada nos é dito relativamente à distribuição da população e o tamanho de cada amostra não permite o uso do teorema do limite central $n \geq 30$, então temos se assumir-mos que:
- $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ e $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$

Exercício Nº19 - Resolução II

- Podemos executar um teste de diferença de médias para amostras emparelhadas:

Exercício Nº19 - Resolução II

- Podemos executar um teste de diferença de médias para amostras emparelhadas:

$$T = \frac{\bar{D} - \mu_D}{\frac{S'_D}{\sqrt{n}}} \sim t_{n-1}, \qquad D = X_1 - X_2 \tag{11}$$

Exercício Nº19 - Resolução II

- Podemos executar um teste de diferença de médias para amostras emparelhadas:

$$T = \frac{\bar{D} - \mu_D}{\frac{S'_D}{\sqrt{n}}} \sim t_{n-1}, \qquad D = X_1 - X_2 \tag{11}$$

- Como assumimos anteriormente que X_1 e X_2 são variáveis normais então $D \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 - 2COV(X_1, X_2))$ também segue uma normal e podemos executar o teste.

Exercício Nº19 - Resolução II

- Podemos executar um teste de diferença de médias para amostras emparelhadas:

$$T = \frac{\bar{D} - \mu_D}{\frac{S'_D}{\sqrt{n}}} \sim t_{n-1}, \qquad D = X_1 - X_2 \tag{11}$$

- Como assumimos anteriormente que X_1 e X_2 são variáveis normais então $D \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 - 2COV(X_1, X_2))$ também segue uma normal e podemos executar o teste.
- Então o nosso teste de hipóteses vai ser formalizado da seguinte forma:

Exercício Nº19 - Resolução II

- Podemos executar um teste de diferença de médias para amostras emparelhadas:

$$T = \frac{\bar{D} - \mu_D}{\frac{S'_D}{\sqrt{n}}} \sim t_{n-1}, \qquad D = X_1 - X_2 \tag{11}$$

- Como assumimos anteriormente que X_1 e X_2 são variáveis normais então $D \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 - 2COV(X_1, X_2))$ também segue uma normal e podemos executar o teste.
- Então o nosso teste de hipóteses vai ser formalizado da seguinte forma:
- $H_0 : \mu_1 - \mu_2 \leq 0$ vs $H_1 : \mu_1 - \mu_2 > 0$, ou seja:

Exercício Nº19 - Resolução II

- Podemos executar um teste de diferença de médias para amostras emparelhadas:

$$T = \frac{\bar{D} - \mu_D}{\frac{S'_D}{\sqrt{n}}} \sim t_{n-1}, \qquad D = X_1 - X_2 \tag{11}$$

- Como assumimos anteriormente que X_1 e X_2 são variáveis normais então $D \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 - 2COV(X_1, X_2))$ também segue uma normal e podemos executar o teste.
- Então o nosso teste de hipóteses vai ser formalizado da seguinte forma:
- $H_0 : \mu_1 - \mu_2 \leq 0$ vs $H_1 : \mu_1 - \mu_2 > 0$, ou seja:
- $H_0 : \mu_D \leq 0$ vs $H_1 : \mu_D > 0$

Exercício Nº19 - Resolução III

- Para uma probabilidade do erro tipo I fixa em 5%, $\alpha = 0.05$, temos que $t(95\%; 10-1) = t(95\%, 9) = 1.833$.

Exercício Nº19 - Resolução III

- Para uma probabilidade do erro tipo I fixa em 5%, $\alpha = 0.05$, temos que $t(95\%; 10-1) = t(95\%, 9) = 1.833$.
- Trata-se de um teste unilateral à direita: $R.C = [1.833, +\infty]$ e $R.C =]-\infty, 1.833[$

Exercício Nº19 - Resolução III

- Para uma probabilidade do erro tipo I fixa em 5%, $\alpha = 0.05$, temos que $t(95\%; 10-1) = t(95\%, 9) = 1.833$.
- Trata-se de um teste unilateral à direita: $R.C = [1.833, +\infty]$ e $R.C =]-\infty, 1.833[$

Doentes	1	2	3	4	5	6	7	8	9	10	Total
$D_i = X_1 - X_2$	4	0	3	5	5	7	3	4	3	6	40
$(D_i - \bar{D})$	0	-4	-1	1	1	3	-1	0	-1	2	0
$(D_i - \bar{D})^2$	0	16	1	1	1	9	1	0	1	4	34

- $\bar{D} = \frac{1}{n}\sum_{i=1}^{n=10} D_i = \frac{40}{10} = 4$

Exercício Nº19 - Resolução III

- Para uma probabilidade do erro tipo I fixa em 5%, $\alpha = 0.05$, temos que $t(95\%; 10-1) = t(95\%, 9) = 1.833$.
- Trata-se de um teste unilateral à direita: $R.C = [1.833, +\infty]$ e $R.C =]-\infty, 1.833[$

Doentes	1	2	3	4	5	6	7	8	9	10	Total
$D_i = X_1 - X_2$	4	0	3	5	5	7	3	4	3	6	40
$(D_i - \bar{D})$	0	-4	-1	1	1	3	-1	0	-1	2	0
$(D_i - \bar{D})^2$	0	16	1	1	1	9	1	0	1	4	34

- $\bar{D} = \frac{1}{n}\sum_{i=1}^{n=10} D_i = \frac{40}{10} = 4$
- $(S'_D)^2 = \frac{1}{n-1}\sum_{i=1}^{n=10}(D_i - \bar{D})^2 = \frac{34}{9} \approx 3,78$

Exercício Nº19 - Resolução III

- Para uma probabilidade do erro tipo I fixa em 5%, $\alpha = 0.05$, temos que $t(95\%; 10-1) = t(95\%, 9) = 1.833$.
- Trata-se de um teste unilateral à direita: $R.C = [1.833, +\infty]$ e $R.C =]-\infty, 1.833[$

Doentes	1	2	3	4	5	6	7	8	9	10	Total
$D_i = X_1 - X_2$	4	0	3	5	5	7	3	4	3	6	40
$(D_i - \bar{D})$	0	-4	-1	1	1	3	-1	0	-1	2	0
$(D_i - \bar{D})^2$	0	16	1	1	1	9	1	0	1	4	34

- $\bar{D} = \frac{1}{n}\sum_{i=1}^{n=10} D_i = \frac{40}{10} = 4$
- $(S'_D)^2 = \frac{1}{n-1}\sum_{i=1}^{n=10}(D_i - \bar{D})^2 = \frac{34}{9} \approx 3,78$
- $S'_D = \sqrt{(S'_D)^2} = \sqrt{3,78} \approx 1,94$

Exercício Nº19 - Resolução IV

- Então a estatística de teste realizada, ou a estatística de teste observada é dada por:

$$T^* = \frac{4-0}{\frac{1{,}94}{\sqrt{10}}} = 6.52 \tag{12}$$

Exercício Nº19 - Resolução IV

- Então a estatística de teste realizada, ou a estatística de teste observada é dada por:

$$T^* = \frac{4-0}{\frac{1{,}94}{\sqrt{10}}} = 6.52 \qquad (12)$$

- Como $T^* \in$ R.C então rejeito H_0 para este nível de significância e para esta realização da amostra aleatória ou para esta amostra observada. Logo o nutricionista deve ter razão!

Exercícios de Outputs de SPSS

- Exercícios com outputs sobre ensaios de hipóteses - Parte B.
- Exercício Nº19

Teste ANOVA - Análise de Variância I

- O teste ANOVA, generaliza o teste à diferença de duas médias para k médias de k populações independentes, portanto vamos ter k amostras aleatórias para executar o ensaio:

Teste ANOVA - Análise de Variância I

- O teste ANOVA, generaliza o teste à diferença de duas médias para k médias de k populações independentes, portanto vamos ter k amostras aleatórias para executar o ensaio:
- Amostra 1: $(X_{11}, X_{21}, ..., X_{n_1 1})$
- Amostra 2: $(X_{12}, X_{22}, ..., X_{n_2 2})$
- ...
- Amostra k: $(X_{1k}, X_{2k}, ..., X_{n_k k})$

Teste ANOVA - Análise de Variância I

- O teste ANOVA, generaliza o teste à diferença de duas médias para k médias de k populações independentes, portanto vamos ter k amostras aleatórias para executar o ensaio:
- Amostra 1: $(X_{11}, X_{21}, ..., X_{n_1 1})$
- Amostra 2: $(X_{12}, X_{22}, ..., X_{n_2 2})$
- ...
- Amostra k: $(X_{1k}, X_{2k}, ..., X_{n_k k})$
- $X_{i,j}$ é o indivíduo $i(i = 1, 2, ..., n_j)$ pertencente à amostra $j(j = 1, 2, ..., k)$ e $n_1, n_2, ..., n_k$ a dimensão da amostra de cada uma das k populações.

Teste ANOVA - Análise de Variância I

- O teste ANOVA, generaliza o teste à diferença de duas médias para k médias de k populações independentes, portanto vamos ter k amostras aleatórias para executar o ensaio:
- Amostra 1: $(X_{11}, X_{21}, ..., X_{n_1 1})$
- Amostra 2: $(X_{12}, X_{22}, ..., X_{n_2 2})$
- ...
- Amostra k: $(X_{1k}, X_{2k}, ..., X_{n_k k})$
- $X_{i,j}$ é o indivíduo $i(i = 1, 2, ..., n_j)$ pertencente à amostra $j(j = 1, 2, ..., k)$ e $n_1, n_2, ..., n_k$ a dimensão da amostra de cada uma das k populações.
- **Pressupostos:** Além de k amostras aleatórias, precisamos de garantir populações normais com variâncias desconhecidas mas iguais!

Teste ANOVA - Análise de Variância I

- O teste ANOVA, generaliza o teste à diferença de duas médias para k médias de k populações independentes, portanto vamos ter k amostras aleatórias para executar o ensaio:
- Amostra 1: $(X_{11}, X_{21}, ..., X_{n_1 1})$
- Amostra 2: $(X_{12}, X_{22}, ..., X_{n_2 2})$
- ...
- Amostra k: $(X_{1k}, X_{2k}, ..., X_{n_k k})$
- $X_{i,j}$ é o indivíduo $i(i = 1, 2, ..., n_j)$ pertencente à amostra $j(j = 1, 2, ..., k)$ e $n_1, n_2, ..., n_k$ a dimensão da amostra de cada uma das k populações.
- **Pressupostos:** Além de k amostras aleatórias, precisamos de garantir populações normais com variâncias desconhecidas mas iguais!
- $X_j \sim \mathcal{N}(\mu_j, \sigma)$, $\forall j = 1, 2, ..., k$

Teste ANOVA - Análise de Variância II

- Portanto formalmente o teste é dado por:

Teste ANOVA - Análise de Variância II

- Portanto formalmente o teste é dado por:
- $H_0 : \mu_1 = ... = \mu_k$ VS $H_1 : \exists r, j : \mu_r \neq \mu_j, r \neq j$

Teste ANOVA - Análise de Variância II

- Portanto formalmente o teste é dado por:
- $H_0 : \mu_1 = ... = \mu_k$ VS $H_1 : \exists r, j : \mu_r \neq \mu_j, r \neq j$
- H_0 : A médias das k populações são todas iguais VS H_1 : Existe pelo menos um par (r, j) em que as médias diferem.

Teste ANOVA - Análise de Variância II

- Portanto formalmente o teste é dado por:
- $H_0 : \mu_1 = ... = \mu_k$ VS $H_1 : \exists r, j : \mu_r \neq \mu_j, r \neq j$
- H_0 : A médias das k populações são todas iguais VS H_1 : Existe pelo menos um par (r, j) em que as médias diferem.
- **Atenção!!** O erro mais comum é achar que a inferência deste teste é sobre a variância devido ao facto do método se chamar análise de variância. Trata-se de um teste à igualdade de k médias.

Teste ANOVA - Análise de Variância II

- Portanto formalmente o teste é dado por:
- $H_0 : \mu_1 = ... = \mu_k$ VS $H_1 : \exists r, j : \mu_r \neq \mu_j, r \neq j$
- H_0 : A médias das k populações são todas iguais VS H_1 : Existe pelo menos um par (r, j) em que as médias diferem.
- **Atenção!!** O erro mais comum é achar que a inferência deste teste é sobre a variância devido ao facto do método se chamar análise de variância. Trata-se de um teste à igualdade de k médias.
- A origem do nome do método vai ficar clara já de seguida, quando a estatística do teste for apresentada.

Teste ANOVA - Análise de Variância III

A estatística do teste ANOVA é dada por:

$$T = \frac{\frac{SSB}{(k-1)}}{\frac{SSW}{(n-k)}} = \frac{MSSB}{MSSW} \sim F_{(k-1,n-k)}, \tag{13}$$

onde SSW é a soma de quadrados dentro (within) dos grupos, ou a soma de quadrados devida aos erros intra-grupo

$$SSW = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (\bar{X}_{ij} - \bar{X}_j)^2, \tag{14}$$

e SSB é a soma de quadrados entre (between) os grupos, ou a soma de quadrados devido ao factor independente

$$SSB = \sum_{j=1}^{k} n_j (\bar{X}_j - \bar{X})^2. \tag{15}$$

Teste ANOVA - Análise de Variância IV

Estas duas parcelas somadas constituem a soma total dos quadrados dos desvios dos valores observados em torno da média global (SST):

$$SST = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (\bar{X}_{ij} - \bar{X})^2, \tag{16}$$

pelo que se obtem a seguitne igualdade

$$SST = SSW + SSB \tag{17}$$

No fundo o nome análise de variância surge porque de facto, através do rácio da estatística de teste, estamos a comparar duas fontes de variação: intra-grupo (SSW) e entre grupos (SSB).

Teste ANOVA - Análise de Variância V

- De notar que:

Teste ANOVA - Análise de Variância V

- De notar que:
- k - número de grupos/populações.

Teste ANOVA - Análise de Variância V

- De notar que:
- k - número de grupos/populações.
- n_j - dimensão da amostra $j (j = 1, 2, ..., k)$.

Teste ANOVA - Análise de Variância V

- De notar que:
- k - número de grupos/populações.
- n_j - dimensão da amostra $j(j = 1, 2, ..., k)$.
- X_{ij} - observação para o indivíduo i do grupo/população j.

Teste ANOVA - Análise de Variância V

- De notar que:
- k - número de grupos/populações.
- n_j - dimensão da amostra $j (j = 1, 2, ..., k)$.
- X_{ij} - observação para o indivíduo i do grupo/população j.
- $\bar{X}_j$ - média amostral do grupo/população j.

Teste ANOVA - Análise de Variância V

- De notar que:
- k - número de grupos/populações.
- n_j - dimensão da amostra $j (j = 1, 2, ..., k)$.
- X_{ij} - observação para o indivíduo i do grupo/população j.
- $\bar{X}_j$ - média amostral do grupo/população j.
- $\bar{X}$ - média amostral de todas as observações.

Teste ANOVA - Análise de Variância V

- De notar que:
- k - número de grupos/populações.
- n_j - dimensão da amostra $j(j = 1, 2, ..., k)$.
- X_{ij} - observação para o indivíduo i do grupo/população j.
- $\bar{X}_j$ - média amostral do grupo/população j.
- $\bar{X}$ - média amostral de todas as observações.
- A dimensão total da amostra é o acumulado da dimensão de cada uma das amostras n_j referentes a cada grupo/população j:

$$n = \sum_{j=1}^{k} n_j \tag{18}$$

Teste ANOVA - Análise de Variância VI

- O teste ANOVA simples, que é o nosso caso, é sempre um teste **unilateral à direita**... Porquê ?

Teste ANOVA - Análise de Variância VI

- O teste ANOVA simples, que é o nosso caso, é sempre um teste **unilateral à direita**... Porquê ?
- Vamos ver a intuição fornecida pela estatística de teste:

Teste ANOVA - Análise de Variância VI

- O teste ANOVA simples, que é o nosso caso, é sempre um teste **unilateral à direita**... Porquê ?
- Vamos ver a intuição fornecida pela estatística de teste:

$$T = \frac{\frac{SSB}{(k-1)}}{\frac{SSW}{(n-k)}} = \frac{MSSB}{MSSW} \sim F_{(k-1,n-k)}, \qquad (19)$$

Teste ANOVA - Análise de Variância VI

- O teste ANOVA simples, que é o nosso caso, é sempre um teste **unilateral à direita**... Porquê ?
- Vamos ver a intuição fornecida pela estatística de teste:

$$T = \frac{\frac{SSB}{(k-1)}}{\frac{SSW}{(n-k)}} = \frac{MSSB}{MSSW} \sim F_{(k-1,n-k)}, \tag{19}$$

- Se as variações entre grupos, medidas por SSB, forem significativamente maiores que as variações intra grupo, medidas por SSW, então o rácio vai ser significativamente maior que 1..., ou seja

Teste ANOVA - Análise de Variância VI

- O teste ANOVA simples, que é o nosso caso, é sempre um teste **unilateral à direita**... Porquê ?
- Vamos ver a intuição fornecida pela estatística de teste:

$$T = \frac{\frac{SSB}{(k-1)}}{\frac{SSW}{(n-k)}} = \frac{MSSB}{MSSW} \sim F_{(k-1,n-k)}, \tag{19}$$

- Se as variações entre grupos, medidas por SSB, forem significativamente maiores que as variações intra grupo, medidas por SSW, então o rácio vai ser significativamente maior que 1..., ou seja
- Só faz sentido rejeitar a hipótese nula, igualdade de k médias populacionais, para valores elevados da estatística de teste, valores esses que ocorrem quando a variação entre os grupos, devido ao factor independente, for relativamente elevada quando comparada com a variação dentro dos grupos (ou devida a erros).

Teste ANOVA - Análise de Variância VII

- Um exemplo intuitivo:

Teste ANOVA - Análise de Variância VII

- Um exemplo intuitivo:
- Se 3 grupos têm médias próximas, tipo 12.1, 12.3 e 12.2, e grandes desvios internos então não existe um efeito de grupo (i.e., o factor independente não se faz notar) $\Longrightarrow$ T vai ser pequeno $\Longrightarrow$ Não rejeito H_0.

Teste ANOVA - Análise de Variância VII

- Um exemplo intuitivo:
- Se 3 grupos têm médias próximas, tipo 12.1, 12.3 e 12.2, e grandes desvios internos então não existe um efeito de grupo (i.e., o factor independente não se faz notar) $\Longrightarrow$ T vai ser pequeno $\Longrightarrow$ Não rejeito H_0.
- Se 3 grupos têm médias susbtancialmente diferentes, tipo 10, 12 e 18, e as diferenças intra-grupo são pequenas então existe um efeito de grupo (i.e., o factor independente faz-se notar) $\Longrightarrow$ T vai ser grande $\Longrightarrow$ Rejeito H_0.

Teste ANOVA - Análise de Variância VII

- Um exemplo intuitivo:
- Se 3 grupos têm médias próximas, tipo 12.1, 12.3 e 12.2, e grandes desvios internos então não existe um efeito de grupo (i.e., o factor independente não se faz notar) $\Longrightarrow$ T vai ser pequeno $\Longrightarrow$ Não rejeito H_0.
- Se 3 grupos têm médias susbtancialmente diferentes, tipo 10, 12 e 18, e as diferenças intra-grupo são pequenas então existe um efeito de grupo (i.e., o factor independente faz-se notar) $\Longrightarrow$ T vai ser grande $\Longrightarrow$ Rejeito H_0.
- Nesta sequência de raciocionio a pergunta mais óbvia será:

Teste ANOVA - Análise de Variância VII

- Um exemplo intuitivo:
- Se 3 grupos têm médias próximas, tipo 12.1, 12.3 e 12.2, e grandes desvios internos então não existe um efeito de grupo (i.e., o factor independente não se faz notar) $\Longrightarrow$ T vai ser pequeno $\Longrightarrow$ Não rejeito H_0.
- Se 3 grupos têm médias susbtancialmente diferentes, tipo 10, 12 e 18, e as diferenças intra-grupo são pequenas então existe um efeito de grupo (i.e., o factor independente faz-se notar) $\Longrightarrow$ T vai ser grande $\Longrightarrow$ Rejeito H_0.
- Nesta sequência de raciocionio a pergunta mais óbvia será:
- Então porque não consideramos apenas o numerador SSB?

Teste ANOVA - Análise de Variância VII

- Um exemplo intuitivo:
- Se 3 grupos têm médias próximas, tipo 12.1, 12.3 e 12.2, e grandes desvios internos então não existe um efeito de grupo (i.e., o factor independente não se faz notar) $\Longrightarrow$ T vai ser pequeno $\Longrightarrow$ Não rejeito H_0.
- Se 3 grupos têm médias susbtancialmente diferentes, tipo 10, 12 e 18, e as diferenças intra-grupo são pequenas então existe um efeito de grupo (i.e., o factor independente faz-se notar) $\Longrightarrow$ T vai ser grande $\Longrightarrow$ Rejeito H_0.
- Nesta sequência de raciocionio a pergunta mais óbvia será:
- Então porque não consideramos apenas o numerador SSB?
- Porque sem o respectivo contexto, uma variação não nos diz grande coisa...

Teste ANOVA - Análise de Variância VII

- Um exemplo intuitivo:
- Se 3 grupos têm médias próximas, tipo 12.1, 12.3 e 12.2, e grandes desvios internos então não existe um efeito de grupo (i.e., o factor independente não se faz notar) $\Longrightarrow$ T vai ser pequeno $\Longrightarrow$ Não rejeito H_0.
- Se 3 grupos têm médias susbtancialmente diferentes, tipo 10, 12 e 18, e as diferenças intra-grupo são pequenas então existe um efeito de grupo (i.e., o factor independente faz-se notar) $\Longrightarrow$ T vai ser grande $\Longrightarrow$ Rejeito H_0.
- Nesta sequência de raciocionio a pergunta mais óbvia será:
- Então porque não consideramos apenas o numerador SSB?
- Porque sem o respectivo contexto, uma variação não nos diz grande coisa...
- Ou seja, não haveria maneira de se dizer uma determinada variação entre grupos era grande ou pequena.

Teste ANOVA - Análise de Variância VII

- Um exemplo intuitivo:

Teste ANOVA - Análise de Variância VII

- Um exemplo intuitivo:
- Se 3 grupos têm médias próximas, tipo 12.1, 12.3 e 12.2, e grandes desvios internos então não existe um efeito de grupo (i.e., o factor independente não se faz notar) $\Longrightarrow$ T vai ser pequeno $\Longrightarrow$ Não rejeito H_0.

Teste ANOVA - Análise de Variância VII

- Um exemplo intuitivo:
- Se 3 grupos têm médias próximas, tipo 12.1, 12.3 e 12.2, e grandes desvios internos então não existe um efeito de grupo (i.e., o factor independente não se faz notar) $\Longrightarrow$ T vai ser pequeno $\Longrightarrow$ Não rejeito H_0.
- Se 3 grupos têm médias susbtancialmente diferentes, tipo 10, 12 e 18, e as diferenças intra-grupo são pequenas então existe um efeito de grupo (i.e., o factor independente faz-se notar) $\Longrightarrow$ T vai ser grande $\Longrightarrow$ Rejeito H_0.

Teste ANOVA - Análise de Variância VII

- Um exemplo intuitivo:
- Se 3 grupos têm médias próximas, tipo 12.1, 12.3 e 12.2, e grandes desvios internos então não existe um efeito de grupo (i.e., o factor independente não se faz notar) $\Longrightarrow$ T vai ser pequeno $\Longrightarrow$ Não rejeito H_0.
- Se 3 grupos têm médias susbtancialmente diferentes, tipo 10, 12 e 18, e as diferenças intra-grupo são pequenas então existe um efeito de grupo (i.e., o factor independente faz-se notar) $\Longrightarrow$ T vai ser grande $\Longrightarrow$ Rejeito H_0.
- Nesta sequência de raciocionio a pergunta mais óbvia será:

Teste ANOVA - Análise de Variância VII

- Um exemplo intuitivo:
- Se 3 grupos têm médias próximas, tipo 12.1, 12.3 e 12.2, e grandes desvios internos então não existe um efeito de grupo (i.e., o factor independente não se faz notar) $\Longrightarrow$ T vai ser pequeno $\Longrightarrow$ Não rejeito H_0.
- Se 3 grupos têm médias susbtancialmente diferentes, tipo 10, 12 e 18, e as diferenças intra-grupo são pequenas então existe um efeito de grupo (i.e., o factor independente faz-se notar) $\Longrightarrow$ T vai ser grande $\Longrightarrow$ Rejeito H_0.
- Nesta sequência de raciocionio a pergunta mais óbvia será:
- Então porque não consideramos apenas o numerador SSB?

Teste ANOVA - Análise de Variância VII

- Um exemplo intuitivo:
- Se 3 grupos têm médias próximas, tipo 12.1, 12.3 e 12.2, e grandes desvios internos então não existe um efeito de grupo (i.e., o factor independente não se faz notar) $\Longrightarrow$ T vai ser pequeno $\Longrightarrow$ Não rejeito H_0.
- Se 3 grupos têm médias susbtancialmente diferentes, tipo 10, 12 e 18, e as diferenças intra-grupo são pequenas então existe um efeito de grupo (i.e., o factor independente faz-se notar) $\Longrightarrow$ T vai ser grande $\Longrightarrow$ Rejeito H_0.
- Nesta sequência de raciocionio a pergunta mais óbvia será:
- Então porque não consideramos apenas o numerador SSB?
- Porque sem o respectivo contexto, uma variação não nos diz grande coisa...

Teste ANOVA - Análise de Variância VII

- Um exemplo intuitivo:
- Se 3 grupos têm médias próximas, tipo 12.1, 12.3 e 12.2, e grandes desvios internos então não existe um efeito de grupo (i.e., o factor independente não se faz notar) $\Longrightarrow$ T vai ser pequeno $\Longrightarrow$ Não rejeito H_0.
- Se 3 grupos têm médias susbtancialmente diferentes, tipo 10, 12 e 18, e as diferenças intra-grupo são pequenas então existe um efeito de grupo (i.e., o factor independente faz-se notar) $\Longrightarrow$ T vai ser grande $\Longrightarrow$ Rejeito H_0.
- Nesta sequência de raciocionio a pergunta mais óbvia será:
- Então porque não consideramos apenas o numerador SSB?
- Porque sem o respectivo contexto, uma variação não nos diz grande coisa...
- Ou seja, não haveria maneira de se dizer que uma determinada variação entre grupos era grande ou pequena.

Exercício Nº22

Uma empresa pretende testar se existem diferenças significativas nos tempos médios de vida (em milhares de km) de quatro marcas de pneus: A, B, C e D.

A	31	25	28	30	32	27.5	
B	24	26	27	25	30	32	28
C	30	30.5	29.5	28	31		
D	24.5	27	26	23	21	22	

(a) Utilize um nível de significância de 0.05 para testar se existem diferenças significativas nos tempos médios de vida das quatro marcas de pneus.
(b) Quais as marcas significativamente diferentes entre si ?
(c) O que conclui acerca do pressuposto da igualdade de variâncias entre os grupos ?

Exercício Nº22 - Resolução (a)

Pretende-se testar se existem diferenças significativas na duração (em 10^3 de km) de 4 marcas de pneus: A, B, C e D.

- Começamos com a definição das populações:
- X_1 — duração (em 10^3 Km) da marca de pneus A
- X_2 — duração (em 10^3 Km) da marca de pneus B
- X_3 — duração (em 10^3 Km) da marca de pneus C
- X_4 — duração (em 10^3 Km) da marca de pneus D

Pressupostos:

Admite-se que para além das amostras serem independentes, o tempo de vida (duração) se distribui normalmente e com igual variância (em cada um dos quatro tipos de pneus).

i) $X_i \sim \mathcal{N}(\mu_i, \sigma^2) \quad i = 1, 2, 3, 4$

ii) $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$ (variâncias desconhecidas mas idênticas)

Exercício Nº22 - Resolução (a)

Formulação das Hipóteses:

- $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ VS $H_1 : \exists_{i,j}, i \neq j : \mu_i \neq \mu_j$

A estatística do teste é:

$$T = \frac{\frac{SSB}{(k-1)}}{\frac{SSW}{(n-k)}} = \frac{MSSB}{MSSW} \sim F_{(k-1;n-k)},$$

Onde:

$$SSB = \sum_{j=1}^{k} n_j(\bar{X}_j - \bar{X})^2 \quad \text{e} \quad SSW = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij} - \bar{X}_j)^2$$

A região crítica é sempre unilateral direita. Para um nível de significância $\alpha = 0{,}05$, tem-se:

- $R.C. = [3{,}098; +\infty[$ e $R.A. = [0; 3{,}098[$, pois $k = 4$ e $n = 24$.

Exercício Nº22 - Resolução (a)

- Começamos com alguns cálculos auxiliares:

$\bar{X}_A = 28{,}92$, $n_A = 6$, $\bar{X}_B = 27{,}43$, $n_B = 7$, $\bar{X}_C = 29{,}8$, $n_C = 5$, $\bar{X}_D = 23{,}92$, $n_D = 6$ e $\bar{X} = 27{,}417$.

- A soma dos quadrados da variação entre grupos, SSB, é dada por:

$$SSB = \underbrace{\sum_{j=1}^{k}}_{\text{grupos}} \underbrace{n_j(\bar{X}_j - \bar{X})^2}_{\substack{\text{desvio da média} \\ \text{do grupo em relação à média global}}}$$

$$= \underbrace{6(\bar{X}_1 - \bar{X})^2}_{\text{grupo A}} + \underbrace{7(\bar{X}_2 - \bar{X})^2}_{\text{grupo B}} + \underbrace{5(\bar{X}_3 - \bar{X})^2}_{\text{grupo C}} + \underbrace{6(\bar{X}_4 - \bar{X})^2}_{\text{grupo D}}$$

$$= 6(28{,}92 - 27{,}417)^2 + 7(27{,}43 - 27{,}417)^2$$

$$+ 5(29{,}8 - 27{,}417)^2 + 6(23{,}92 - 27{,}417)^2 = \boxed{115{,}40}$$

Exercício Nº22 - Resolução (a)

- Para a soma dos quadrados intra grupo, SSW, seguimos também a fórmula:

$$SSW = \underbrace{\sum_{j=1}^{k}}_{\text{grupos}} \underbrace{\sum_{i=1}^{n_j}}_{\substack{\text{observações} \\ \text{no grupo } j}} (X_{ij} - \bar{X}_j)^2$$

$$= \underbrace{(31 - 28{,}92)^2 + (25 - 28{,}92)^2 + \ldots + (27{,}5 - 28{,}92)^2}_{\text{grupo A, } j=1;\, i=1\ldots6}$$

$$+ \underbrace{(24 - 27{,}43)^2 + \ldots + (28 - 27{,}43)^2}_{\text{grupo B, } j=2;\, i=1\ldots6}$$

$$+ \underbrace{(30 - 29{,}8)^2 + \ldots + (31 - 29{,}8)^2}_{\text{grupo C, } j=3;\, i=1\ldots5}$$

$$+ \underbrace{(24{,}5 - 23{,}92)^2 + \ldots + (22 - 23{,}92)^2}_{\text{grupo D, } j=4;\, i=1\ldots6}$$

$$= \boxed{113{,}43}$$

Exercício Nº22 - Resolução (a)

- A soma total dos quadrados, SST, também pode ser cálculada através da fórmula:

$$SST = \underbrace{\sum_{j=1}^{k}}_{\text{grupos}} \left(\underbrace{\sum_{i=1}^{n_j}}_{\substack{\text{observações} \\ \text{no grupo } j}} (x_{ij} - \bar{X})^2 \right) =$$

$$= \underbrace{(31 - 27{,}417)^2 + (25 - 27{,}417)^2 + \ldots + (27{,}5 - 27{,}417)^2}_{\text{grupo } A,\, j=1;\, i=1 \ldots 6}$$

$$+ \underbrace{(24 - 27{,}417)^2 + \ldots + (28 - 27{,}417)^2}_{\text{grupo } B,\, j=2;\, i=1 \ldots 6} + \underbrace{(30 - 27{,}417)^2 + \ldots + (31 - 27{,}417)^2}_{\text{grupo } C,\, j=3;\, i=1 \ldots 5}$$

$$+ \underbrace{(24{,}5 - 27{,}417)^2 + \ldots + (22 - 27{,}417)^2}_{\text{grupo } D,\, j=4;\, i=1 \ldots 6} = \boxed{228{,}83}$$

Exercício Nº22 - Resolução (a)

- A soma dos quadrados da variação intra grupos, SSW, também pode ser cálculada à custa de $SST = SSW + SSB$ e SST também pode ser cálculada à custa do conceito de variância global $\neq$ variância amostral:

$$S^2 = \frac{1}{n} \underbrace{\sum_{j=1}^{k} \sum_{i=1}^{n_j} (\bar{X}_{ij} - \bar{X})^2}_{SST} = \frac{1}{n} SST \Longrightarrow SST = nS^2.$$

- Atentem que esta variância global, ainda que se faça uso do mesmo S^2, não é a mesma coisa que a variância amostral...

Exercício Nº22 - Resolução (a)

- A soma dos quadrados da variação intra grupos, SSW, também pode ser cálculada à custa de $SST = SSW + SSB$ e SST também pode ser cálculada à custa do conceito de variância global $\neq$ variância amostral:

$$S^2 = \frac{1}{n} \underbrace{\sum_{j=1}^{k} \sum_{i=1}^{n_j} (\bar{X}_{ij} - \bar{X})^2}_{SST} = \frac{1}{n} SST \implies SST = nS^2.$$

- Atentem que esta variância global, ainda que se faça uso do mesmo S^2, não é a mesma coisa que a variância amostral...
- Esta fórmula mistura as amostras de várias populações, coisa que a variância amostral não faz.

Exercício Nº22 - Resolução (a)

- Usando a amostra fornecida para proceder ao cálculo:

$$S^2 = \frac{1}{n}\sum_{j=1}^{k}\sum_{i=1}^{n_j} x_{ij}^2 - \bar{X}^2$$

$$= \frac{1}{24}\left(\underbrace{31^2 + 25^2 + \ldots + 27{,}5^2}_{\text{Grupo A}} + \underbrace{24^2 + 26^2 + \ldots + 28^2}_{\text{Grupo B}} + \right.$$

$$\left. \underbrace{30^2 + 30{,}5^2 + \ldots + 31^2}_{\text{Grupo C}} + \underbrace{24{,}5^2 + 27^2 + \ldots + 22^2}_{\text{Grupo D}}\right) - (27{,}417)^2$$

$$= \frac{1}{24} \cdot 18269{,}00 - 751{,}69 = 761,208 - 751{,}674 = \boxed{9{,}535}$$

- Pelo que $SST = 24 \times 9.535 = 228,83$ e $SSW = 228.83 - 115.40 = 113,43.$

Exercício Nº22 - Resolução (a)

- Normalmente faz-se uma tabela resumo com os valores das medidas necessárias à realização do teste:

Fonte de variação	Graus de liberdade	Soma de quadrados (SS)	Soma média de quadrados (MSS)	Estatística de teste
Entre os grupos	3	SSB=115,40	MSSB=38,467	$T^* = 6{,}783$
Dentro dos grupos	20	SSW=113,43	MSSW=5,672	
Total	23	SST=228,83		

- Dado que a estatística do teste pertence à região crítica, devemos rejeitar H_0, ou seja, a um nível de significância de 5% e para esta realização da amostra aleatória, existem diferenças entre pelo menos uma par de médias. As marcas não apresentam a mesma duração média de vida.

Comparações Múltiplas

- Quando se rejeita H_0 no teste ANOVA então pelo menos um par de médias μ_r, μ_i são diferentes...
- Então temos de descubrir quais são... Surgem os testes de comparação múltipla.