Pesquisa Quantitativa

Lupércio França Bessegato
Mestrado em Administração/UFJF

Testes de Significância

Roteiro Geral

1. Introdução
2. Coleta de dados
3. Modelos probabilísticos
4. Distribuições amostrais e estimação
5. **Testes de significância**
6. Comparações de médias
7. Tabelas de contagem
8. Análise de regressão
9. Referências

Pesquisa Quantitativa - 2016 3

Roteiro do Módulo

5. Testes de significância:
   a) Introdução
   b) Tipos de hipóteses
   c) Medição da evidência contra H_0
   d) Teste de hipótese como tomada de decisão
   e) Testes de hipóteses e intervalos de confiança
   f) Questões de uma amostra

Pesquisa Quantitativa - 2016 4

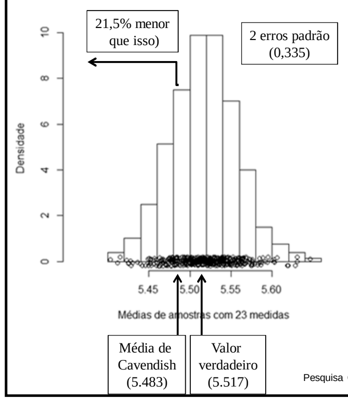
Introdução

Exemplo

- Medidas da densidade da terra.
  - √ Medidas da densidade da terra.
    - Quantidade de medidas: $n = 23$
    - Média da amostra: $\bar{x} = 5{,}4835$
    - Desvio-padrão da amostra: $s = 0{,}1904$
  - √ Valor teórico da densidade da terra:
    - $\mu = 5{,}517$
- O instrumento de Cavendish era viciado?

Pesquisa Quantitativa - 2016 6

- Simulação de 400 medidas da densidade
  - √ Instrumento não viciado de mesma precisão
    - $\mu = 5{,}517$ e $\sigma = 0{,}1904$



- Valor obtido não é incomum para as médias de amostras de experimento não viciado
- Valor obtido cai dentro dos 60% centrais da distribuição
- Dados não podem estabelecer que as medidas são viciadas
  - √ Aceitável que são não viciadas
- Conclusão:
  - √ Não há evidência de que procedimento fosse viciado

Pesquisa Quantitativa - 2016 7

Distância entre Estimativa e Valor Teórico

- Se valor hipotético está correto
  - √ (admitido por hipótese, valor teórico, etc.)
  - √ Estimativa do parâmetro, obtida a partir dos dados, não estará muito longe do valor suposto
- Distância entre a estimativa e o valor hipotético
  - √ Medida em termos de número de erros padrão da estimativa

$$T = \frac{\text{estimativa} - \text{valor hipotético do parâmetro}}{\text{erro padrão da estimativa}} .$$

Pesquisa Quantitativa - 2016 8

$$T = \frac{\text{estimativa} - \text{valor hipotético do parâmetro}}{\text{erro padrão da estimativa}} \ .$$

- Se distância é inaceitavelmente grande:
  - √ Rejeita-se o valor admitido por hipótese

O que é uma distância inaceitável?

- Se o problema estiver em conformidade com uma distribuição t ou Normal(0, 1)
  - √ Mesmo que aproximadamente
  - √ Distribuição de T seria conhecida!

1. Valor observado (t_0) é valor típico dessa distribuição
   - √ Verdadeiro valor do parâmetro é o valor suposto
2. Valor observado não é típico
   - √ Verdadeiro valor não é o valor suposto
   - √ Conclusão:
     - Não temos o verdadeiro valor do parâmetro

Exemplo

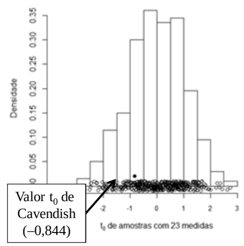
- Medidas da densidade da terra.
  - √ Medidas da densidade da terra.
    - Quantidade de medidas: $n = 23$
    - Média da amostra: $\bar{x} = 5{,}4835$
    - Desvio-padrão da amostra: $s = 0{,}1904$
  - √ Valor teórico da densidade da terra:
    - $\mu = 5{,}517$
- Distância entre estimativa e valor hipotético

$$\text{ep}(\bar{x}) = \frac{0,1904}{\sqrt{23}} = 0,0397 \ .$$

$$t_0 = \frac{\text{média amostral} - 5,517}{\text{erro padrão da média amostral}} = \frac{5,4835 - 5,517}{0,0397} = -0,844 \ .$$

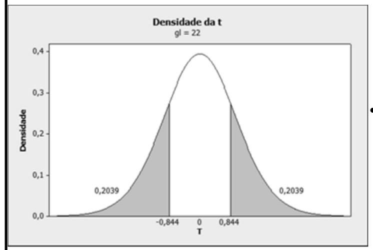
$$t_0 = \frac{5,4835 - 5,517}{0,0397} = -0,844 \ .$$

- √ Média das observações de Cavendish está 0,844 erros padrão abaixo do valor teórico
- Distribuição amostral da distância (T)
  - √ $T \sim t_{23-1}$



- √ Essa diferença pode ser explicada em termos da variação amostral?

- Caso o valor teórico seja o verdadeiro:
  - √ Probabilidade de erro amostral produzir distâncias maiores que a observada ($|t_0|$=0,844)



- 40,8% das médias amostrais estariam mais afastadas do verdadeiro valor
  - √ Medidas em erros padrão
- É plausível que a média populacional verdadeira seja o valor hipotético (teórico)
  - √ Diferença entre $\bar{x}$ e μ pode ser devido à variação amostral
- Não foi provada uma verdade
  - √ Ela é crível

13

Pesquisa Quantitativa - 2016

Teste de Significância e Intervalo de Confiança

- Intervalo de confiança:
  - √ Escolhe-se um nível de confiança
  - √ Determina-se um intervalo de valores possíveis para o parâmetro, consistente com os dados
  - √ Ex.:
    - [5,40; 5,57]: IC com 95% para a média verdadeira μ

14

Pesquisa Quantitativa - 2016

Teste de Significância e Intervalo de Confiança

- Teste de significância:
  - √ Testa-se apenas um valor possível para o parâmetro (valor hipotético)
  - √ Determina-se a a força da evidência fornecida pelos dados contra a proposição
    - – Proposto: valor hipotético é o verdadeiro valor
  - √ Ex.:
    - μ pode ser 5,517?

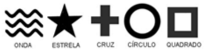
15

Pesquisa Quantitativa - 2016

Tipos de Hipóteses

Exemplo – Percepção Extra-sensorial

- Experimento para investigar PES
  - √ Pratt e Woodruf (1938)
  - √ Experimentador virava uma carta aleatória de baralho Zener (ou Rhine)



  - √ Sujeito experimental não podia ver a carta e indicava o formato que "achava" estava desenhado
  - √ Sujeitos experimentais: vários estudantes

17

Pesquisa Quantitativa - 2016

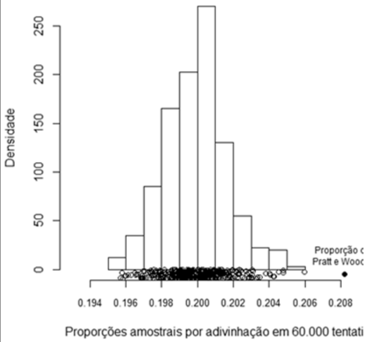
Exemplo – Percepção Extra-sensorial

- √ Dados amostrais:
  - 12.489 adivinhações corretas em 60.000 tentativas
- √ Se todos os sujeitos estivessem apenas adivinhando
  - Obter cerca de 12.000 observações corretas
  - 20% (1/5) das tentativas
- √ Taxa de sucesso do experimento: $\hat{p} = \frac{12.489}{60.000} = 20,8\%$.
- É possível que todos os estivessem apenas adivinhando?
  - √ Variação amostral pode explicar sozinha a taxa de sucesso obtida?

18

Pesquisa Quantitativa - 2016

- Simulação de 4.000 experimentos
  - √ Adivinhação apenas



- Nenhuma das proporções dos 4.000 experimentos de adivinhação foi tão grande quanto o de Pratt e Woodruff
- Adivinhação e variação amostral parecem não explicar a taxa de sucesso obtida
  - √ Estimativa não é consistente com o padrão de variação que espera-se que ocorra com adivinhação

19

Pesquisa Quantitativa - 2016

Hipóteses

- Não podemos nunca demonstrar que um valor hipotético para um parâmetro seja o valor verdadeiro
- Abordagem oposta:
  - √ Podemos determinar se há evidência para descartar o valor admitido por hipótese

20

Pesquisa Quantitativa - 2016

Hipótese Nula

- Em Estatística, testamos a hipótese que podemos descartar
  - √ Proposta para se verificar se podemos derrubá-la

Exemplo – Taxa de Retorno

- Empresas alvos para aquisição dão taxa de retorno menor do capital do investidor?
  - √ Considera-se que empresas com retornos menores do capital do acionista têm maior probabilidade de atrair ofertas de aquisição
- Suspeita:
  - √ Em média, os tipos de empresas visados para aquisição tendem a ser aquelas com retorno de capital pior do que a média
    - – (Hipótese de pesquisa)

Hipótese de Pesquisa

- Lança conjecturas que planejamos investigar através da pesquisa
- Se os ‘palpites’ dos pesquisadores se mostram corretos, são estabelecidos como verdadeiros

Hipóteses Mais Comumente Testadas

- Hipóteses do tipo ‘não faz nenhuma diferença’
  - √ Não há diferença na motivação dos trabalhadores entre companhias com novo tipo de gestão e aquelas que não o adotam
  - √ Não há diferenças nas taxas de promoções entre homens e mulheres
  - √ Uma campanha de propaganda não fez qualquer diferença no volume de vendas

Teste de Significância

- Uso principal do teste de significância
  - √ Verificar se as diferenças (ou efeitos) aparentes observados nos dados podem ser explicados simplesmente em termos de variação amostral
- Se elas podem, ninguém ficará impressionado

Pesquisa Quantitativa - 2016 25

Hipótese Nula

- Reação cética à hipótese de pesquisa (palpite)
  - √ São do tipo 'não faz nenhuma diferença'
- Diferenças observadas nos dados são explicáveis em termos de variação amostral
  - √ Não fornecem evidências convincentes de que existam diferenças reais

Pesquisa Quantitativa - 2016 26

Exemplo – Percepção Extra-sensorial

- Reação cética:
  - √ "Eles estavam apenas adivinhando"
- H_0:
  - √ $p = 0{,}20$

Pesquisa Quantitativa - 2016 27

Hipótese Alternativa (H_1)

- Corresponde à hipótese de pesquisa
- Especifica o tipo de afastamento da hipótese nula que esperamos detectar
  - √ Temos que pensar claramente a respeito do tipo de comportamento dos dados que seria evidência contra a hipótese nula, confirmando a hipótese alternativa

Pesquisa Quantitativa - 2016 28

Exemplo – PES

- Hipótese de pesquisa:
  - √ Alguns indivíduos poderiam receber informações sobre as cartas por meio de PES
- Hipótese nula (reação cética):
  - √ Taxa de sucesso deve-se apenas à adivinhação
- Hipóteses:
  - √ H_0: p = 0,2
  - √ H_1: p ≠ 0,2

Comportamento dos dados que fornece evidências contra H_0 (confirma H_1)

Pesquisa Quantitativa - 2016 29

Construção do Teste

- Hipóteses:
  - √ H_0: p = 0,2
  - √ H_1: p ≠ 0,2
- Comportamento dos dados que confirma H_1
  - √ Fornece evidências contra H_0
  - √ Taxa amostral de sucesso suficientemente maior que 0,2
    - – Não pode ser explicada simplesmente em termos da variação amostral

Pesquisa Quantitativa - 2016 30

Medição da Evidência contra a Hipótese Nula

Teste de Significância

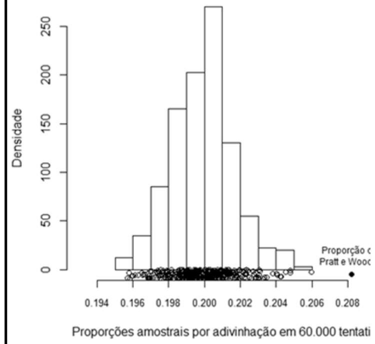
- Comparação de resultado dos dados com o padrão de variação amostral que seria gerado se H_0 fosse verdadeira
- Evidência contra a hipótese:
  - √ Valor dos dados que está fora da amplitude usual daquela variação

Pesquisa Quantitativa - 2016 32

Exemplo – Percepção Extra-sensorial

- Experimento para investigar PES
  - √ Pratt e Woodruf (1938)
  - √ Experimentador virava uma carta aleatória de baralho Zener (ou Rhine)
- Hipóteses:
  - √ H_0: sujeitos adivinharam vs. H_1: não adivinharam
  - √ H_0: p = 0,20 vs. H_1: p ≠ 0,20
- Estimativas (valores amostrais):
  - √ n = 60.000 $\hat{p} = \dfrac{12.489}{60.000} = 20,8\%$.

- Simulação de 4.000 experimentos
  - √ Adivinhação apenas

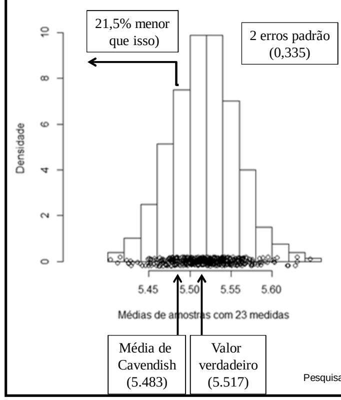


- Estimativa não é consistente com o padrão de variação que espera-se que ocorra com adivinhação (H_0)

Exemplo

- Medidas da densidade da terra.
  - √ Efetuadas por Cavendish (1797-1798)
  - √ Valor teórico (hoje): $\mu = 5{,}517$
- Hipóteses:
  - √ H_0: instrumento não viciado vs. H_1: viciado
  - √ H_0: $\mu = 5{,}517$ vs. H_1: $\mu \neq 5{,}517$
- Valores amostrais:
  - √ Quantidade de medidas: $n = 23$
  - √ Média da amostra: $\bar{x} = 5{,}4835$
  - √ Desvio-padrão da amostra: $s = 0{,}1904$

- Simulação de 400 medidas da densidade
  - √ Instrumento não viciado de mesma precisão
    - $\mu = 5{,}517$ e $\sigma = 0{,}1904$



- Valor obtido não é incomum para as médias de amostras de experimento não viciado
  - √ Valor obtido cai dentro dos 60% centrais da distribuição

Medição da Evidência contra H_0

Procedimento formal:

- Passo 1: Calcule a estatística de teste

$$T = \frac{\hat{\theta} - \theta_0}{\text{ep}(\hat{\theta})} = \frac{\text{estimativa} - \text{valor admitido por hipótese}}{\text{erro padrão}}.$$

√ com $T \sim t_{\text{gl}}$ (ou N(0,1) se gl $= \infty$).

√ Estatística de teste:

- Quantos erros padrão a estimativa obtida a partir dos dados está afastada do valor hipotético
- Média: T e t_0
- Proporção: Z e z_0

√ Significado de "mais afastada" depende de H_1

- Passo 2: Calcule o p-valor:

√ Medida da evidência contra H_0

	Hipótese alternativa	Evidência contra H_0 fornecida por	p-valor
Unilaterais	H_1: $\theta > \theta_0$	$\hat{\theta}$ muito maior do que θ_0 (ou seja $\hat{\theta} - \theta_0$ muito grande)	$p = \Pr\{T \geq t_0\}$
	H_1: $\theta < \theta_0$	$\hat{\theta}$ muito menor do que θ_0 (ou seja $\hat{\theta} - \theta_0$ muito negativa)	$p = \Pr\{T \leq t_0\}$[a]
Bilateral	H_1: $\theta \neq \theta_0$	$\hat{\theta}$ muito distante de θ_0 (ou seja $\lvert\hat{\theta} - \theta_0\rvert$ muito grande)	$p = 2 \times \Pr\{T \geq \lvert t_0\rvert\}$

√ [a]Se $\theta < \theta_0$ é positivo, o p-valor é maior do que 50% e não há necessidade de fazer o cálculo

√ p-valor:

- Probabilidade de, se H_0 for verdadeiro, a variação amostral produzir uma estimativa que esteja mais afastada do valor hipotético que a estimativa obtida a partir dos dados

√ Consequência:

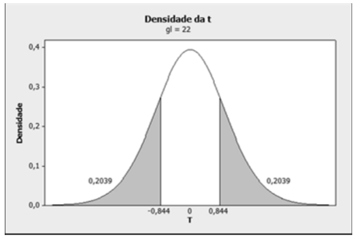
- Quanto menor o p-valor, mais forte a evidência contra H_o.

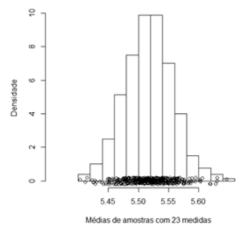
√ Usamos testes bilaterais, a menos que:

- exista uma hipótese de pesquisa a priori (antes de o dado ser coletado)
- razão forte a priori para acreditar que o resultado deveria ir numa direção em particular.

- Passo 3: Interprete o p-valor no contexto dos dados

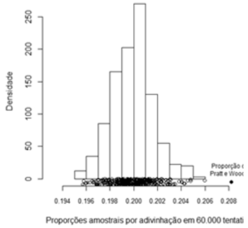
√ p-valor grande:





- Se H_0 fosse verdadeira, a variação amostral em torno do valor hipotético forneceria frequentemente valores que estariam ainda mais afastados do valor hipotético do que a estimativa (dados).

√ p-valor pequeno:



– Se H_0 fosse verdadeira, a variação amostral em torno do valor admitido por hipótese quase sempre daria valores que estariam mais próximos do valor hipotético

√ Quanto menor o p-valor, menos plausível a ideia de que o valor hipotético é o verdadeiro valor do parâmetro

– Mais forte a evidência de que o valor hipotético não é o verdadeiro valor do parâmtero

Tamanho aproximado do p-valor		**Tradução**[a]
> 0,12	(12%)	**Nenhuma** evidência contra H_0
0,10	(10%)	Evidência **fraca** contra H_0
0,05	(5%)	**Alguma** evidência contra H_0
0,01	(1%)	**Forte** evidência contra H_0
≤ 0,001	(0,1%)	Evidência **muito forte** contra H_0

[a]Essas traduções são dos autores e não são universalmente aceitas

Fonte: WILD, J. W; SEBER, G. A. F. *Encontros com o acaso: Um primeiro curso de análise de dados e inferência*. Rio de Janeiro: LTC, 2004

Exemplo – p-valor

- Medidas da densidade da Terra – Cavendish
  - √ Hipóteses:
    - – H_0: $\mu =$ 5,517 vs. H_1: $\mu \neq$ 5,517
- Estatísticas amostrais:

n	$\bar{x}$	s_X	**ep**($\bar{x}$)
23	5,4835	0,1904	0,0397

- Estatística de teste:

$$t_0 = \frac{\text{estimativa} - \text{valor admitido por hipótese}}{\text{erro padrão}} = \frac{5,4835 - 5,517}{0,0397} = -0,844\ .$$

- Determinação do p-valor
  - √ H_1: $\mu \neq$ 5,517 (bilateral)
  - √ Distribuição amostral: t_{22} (gl = 23 – 1)

$$p - valor = 2 \times \Pr\{T \geq |-0,84426|\} = 0,408.$$

  - √ P-valor é muito grande
- Conclusão:
  - √ Dados não fornecem nenhuma evidência contra o valor admitido por hipótese (instrumento viciado)
- IC com 95% de confiança para média μ:

$$\bar{x} \pm t_{gl} \times \text{ep}(\bar{x}) = 5,4835 \pm 2,073873 \times 0,0397 = [5,40; 5,57]\ .$$

Exemplo A

Empresas com retornos fracos são aquelas que atraem mais ofertas de aquisição?

√ *Retorno anormal*: retorno de capital da empresa menos retorno médio (todas as empresas)

– Valor negativo para retorno menor que a média

√ Hipótese de pesquisa:

– *retornos anormais* para empresas-alvo para aquisição seriam em média negativos

√ Dados amostrais

n	$\bar{x}$	s_X	$\mathbf{ep}(\bar{x})$
88	–0,0029	0,0169	0,0018015

$$\mathrm{ep}(\bar{x}) = \frac{s_X}{\sqrt{n}} = \frac{0,0169}{\sqrt{88}} = 0,0018015\ .$$

45

Pesquisa Quantitativa - 2016

- Hipótese de pesquisa (H_1):
  - √ *Retornos anormais* para os alvos para aquisição seriam negativos na média
- Reação cética a H_1 (H_0)
  - √ Aquelas empresas não são diferentes de outras
- Parâmetro de interesse:
  - √ μ: média verdadeira (desconhecida) para a variável retornos anormais para empresas-alvo.
- Hipóteses
  - √ H_0: μ = 0 vs. H_1: μ < 0

46

Pesquisa Quantitativa - 2016

- Estatística de teste:

$$t_0 = \frac{\text{estimativa} - \text{valor admitido por hipótese}}{\text{erro padrão}}$$
$$= \frac{-0,0029 - 0}{0,0018015} = -1,610\ .$$

  - √ Estimativa está 1,61 erro padrão abaixo do valor admitido por hipótese
- Determinação do p-valor:
  - √ Médias amostrais negativas grandes são evidência contra H_0.
  - √ Distribuição amostral: t_{87} (gl = 88 – 1)

$$p - valor = \Pr\{T \leq -1,610\} = 0,056.$$

47

Pesquisa Quantitativa - 2016

- Conclusão:
  - √ Como o p-valor é aproximadamente 5%, isto nos fornece alguma evidência contra H_0, em favor da hipótese de pesquisa de que alvos para aquisição têm menores valores de retornos anormais em média do que empresas em geral.

48

Pesquisa Quantitativa - 2016

- E se a hipótese alternativa fosse $\mu \neq 0$?
  - √ Determinação do p-valor

$$p - valor = 2 \times \Pr\{T \geq -1,610\} = 0,112.$$

  - √ Conclusão:
    - Os dados não fornecem evidência contra H_0, de que, em média, os valores de retornos anormais de empresas-alvo para aquisição são iguais aos das empresas em geral.
  - √ IC de 95% para o retorno anormal médio

$$\bar{x} \pm t_{gl} \times \text{ep}(\bar{x}) = -0,0029 \pm 1,98761 \times 0,0018015 = [-0,0065; 0,0007] \ .$$

Exemplo B

Reprodução mais veloz de propaganda pode aumentar a recordação?

- √ Estudo comparativo dos efeitos de comerciais de tempo reduzido, com amostras independentes
  - Grupo 1: versão acelerada (24s)
  - Grupo 2: versão normal (30 s)
- √ Hipótese de pesquisa:
  - Taxa de recordação dois dias depois para a versão acelerada é significativamente maior?

- √ Dados amostrais:
  - X: número de pessoas que recordaram a propaganda

Versão	n	X	$\hat{p}$
reduzida	74	32	0,4324
normal	57	15	0,2632

- √ Erro padrão:
  - Resposta dicotômica em duas amostras independentes

$$\begin{aligned} \text{ep}(\hat{p}_A - \hat{p}_N) &= \sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A} + \frac{\hat{p}_N(1-\hat{p}_N)}{n_N}} \\ &= \sqrt{\frac{0,4324 \times 0,5676}{74} + \frac{0,2632 \times 0,7368}{57}} \\ &= 0,0819667 \ . \end{aligned}$$

- Hipótese de pesquisa (H_1):
  - √ Acelerar a propaganda poderia mudar o nível de recordação
    - Sem crença a priori de que melhoraria recordação
- Reação cética a H_1 (H_0)
  - √ Não há diferença entre as versões quanto à taxa de recordação
- Parâmetro de interesse:
  - √ $p_A - p_N$: verdadeira diferença desconhecida.
- Hipóteses
  - √ H_0: $p_A - p_N = 0$ vs. H_1: $p_A - p_N \neq 0$

- Estatística de teste:

$$t_0 = \frac{\text{estimativa} - \text{valor admitido por hipótese}}{\text{erro padrão}}$$
$$= \frac{(0,4324 - 0,2632) - 0}{0,0819667} = 2,0652\ .$$

 √ Estimativa está 2,07 erro padrão acima da diferença admitida por hipótese

- Determinação do p-valor:
 √ Diferenças amostrais grandes são evidência contra H_0.
 √ Distribuição amostral: N(0, 1) (gl = ∞)
 - Amostras são grande o suficiente (Apêndice A3)

$$p - valor = 2 \times \Pr\{T \geq 2,0652\} = 0,039.$$

53

Pesquisa Quantitativa - 2016

- Conclusão:
 √ Como o p-valor é pequeno, temos evidência contra H_0. Assim, temos evidência de que existe uma diferença real.
 √ Estimativa mostra que é a versão acelerada deste comercial em particular que as pessoas mais provavelmente recordam
- Mas, quanto mais provável?

54

Pesquisa Quantitativa - 2016

√ IC de 95% para a verdadeira proporção (p_A–p_N)

$$(\hat{p}_A - \hat{p}_N) \pm t_\infty \times \text{ep}(\hat{p}_A - \hat{p}_N) =$$
$$(0,4324 - 0,2632) \pm 1,98761 \times 0,0819667 =$$
$$[0,01; 0,33]\ .$$

 - Intervalo não é muito preciso

√ A verdadeira diferença pode ser muito pequena
 - 1% a mais de telespectadores conseguindo recordar seu comercial provavelmente não causaria um impacto real na efetividade da campanha publicitária

√ A verdadeira diferença pode ser muito grande
 - Ter 33% a mais de público sonseguindo recordar seu comercial (com redução dos custos de exibição) é um bônus interessante.

55

Pesquisa Quantitativa - 2016

Exemplo C

Quem responde ‘sim’ está mais seguro?

√ Pessoas com opinião forte a respeito de uma questão tendem a responder rapidamente
 - Se não têm opinião a priori forte, tendem a levar um tempo maior pensando

√ Pesquisa de opinião por telefone:
 - “Supondo que seja conveniente para você, você espera obter uma injeção contra gripe suína?”
 - Medidas do tempo de resposta (*valores de certeza)*

√ Questão de interesse:
 - Pessoas que responderam sim tinham mais ou menos confiança acerca de sua resposta?

56

Pesquisa Quantitativa - 2016

√ Dados amostrais:

– Dados dos valores de certeza (s^{-1})

	Responderam sim	Responderam não
Tamanho da amostra	100	43
Média amostral	0,76	1,41
Desvio padrão	0,50	0,72
Fonte: LaBarbera e MacLachlan (1979)		

- Hipótese de pesquisa (H_1):
  - √ Quem responde 'sim' tem mais segurança
    - – Tempo médio de resposta é menor (μ_S)
- Reação cética a H_1 (H_0)
  - √ Não há diferença entre os tempos médios de resposta

- Parâmetro de interesse:
  - √ $\mu_N - \mu_S$: diferença desconhecida entre os valores de certeza
- Hipóteses
  - √ H_0: $\mu_N - \mu_S = 0$ vs. H_1: $\mu_N - \mu_S \neq 0$
    - – Não temos nenhuma razão *a priori* que, como grupo, os 'nãos' deveriam tender a ser respondidos com mais ou menos convicção do que os 'sins'.

√ Erro padrão:

– Grupos independentes (grupos inteiramente distintos de pessoas)

$$\text{ep}(\bar{x}_N - \bar{x}_S) = \sqrt{\frac{s_N^2}{n_N} + \frac{s_S^2}{n_S}} = \sqrt{\frac{0,72^2}{43} + \frac{0,50^2}{100}} = 0,1206475\ .$$

- Estatística de teste:

$$t_0 = \frac{\text{estimativa} - \text{valor admitido por hipótese}}{\text{erro padrão}} = \frac{(1,41 - 0,76) - 0}{0,1206475} = 5,387597\ .$$

√ Estimativa está 5,38 erros padrão acima da diferença admitida por hipótese (parece grande!)

- Determinação do p-valor:
  - √ Distância amostral grande é evidência contra H_0.
  - √ Distribuição amostral: t_{42} [gl=Min(100– 1, 43– 1)]

$$p - valor = 2 \times \Pr\{T \geq 5,387597\} = 0,000003.$$

- Conclusão:
  - √ Temos evidência muito forte contra H_0, ou seja, evidência muito forte de uma diferença real.
- Quão maior é μ_N do que μ_S?
  - √ Média amostral de certeza do grupo 'não' é maior

√ IC de 95% para a verdadeira diferença ($\mu_N - \mu_S$)

$$(\bar{x}_N - \bar{x}_S) \pm t_{42} \times \text{ep}(\bar{x}_N - \bar{x}_S) =$$
$$(1,41 - 0,76) \pm 2,018082 \times 0,1206475 =$$
$$[0,41; 0,89] \ .$$

- Com 95% de confiança, o valor de certeza média no grupo 'não' está entre 0,41 e 0,89 s^{-1} maior do que o valor médio do grupo 'sim'

√ IC de 95% para os tempos

- certeza = 1/tempo

$$\left[\frac{1}{0,89}; \frac{1}{0,41}\right]$$
$$[1,1; 2,4] \ .$$

- 'sins' são respondidos mais vagarosamente entre 1,1 e 2,4 segundos

Exemplo D

Nascimentos subsequentes tendem a ser do mesmo sexo?

√ Pesquisa secundária para confirmar evidências

- Dados de 20 anos de registro de hospital

√ Hipótese de pesquisa:

- Se primeiro filho foi uma menina, é mais provável de segundo filho ser menina
- Idem, para meninos

√ Dados amostrais:

- Primeiro e segundo nascimentos por sexo

		Segundo filho		
		Menino	**Menina**	**Total**
Primeiro filho	**Menino**	3.202	2.776	5.978
	Menina	2.620	2.792	5.412
	Total	5.822	5.568	11.390

√ Tabela reformatada

- Análise do sexo dos segundos filhos

grupo	**Nascimentos (n)**	**Número de meninas (X)**	$\hat{p}$
1 (filho anterior foi uma menina)	5.412	2.792	0,516
2 (filho anterior foi um menino)	5.978	2.776	0,464

- Hipótese de pesquisa (H_1):
  √ Nascimentos do primeiro grupo (1º filho = menina) são mais prováveis de ser menina do que do segundo grupo (1º filho = menino)
- Reação cética a H_1 (H_0)
  √ Não há diferença nas verdadeiras proporções entre os dois grupos

- Parâmetro de interesse:
  - √ $p_1 – p_2$: diferença entre as verdadeiras proporções
    - p_1: verdadeira proporção de meninas de mães cujo 1º filho foi uma menina
    - p_2: verdadeira proporção de meninas de mães cujo 1º filho foi uma menino
- Hipóteses
  - √ H_0: $p_1 – p_2 = 0$ vs. H_1: $p_1 – p_2 > 0$
    - Há evidência *a* priori de que nascimentos subsequentes tendem a ser do mesmo sexo.

√ Erro padrão:

- Grupos independentes (grupos inteiramente distintos de pessoas)

$$\mathrm{ep}(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \sqrt{\frac{0,5159 \times 0,4841}{5412} + \frac{0,4644 \times 0,5356}{5978}} = 0,0093677 \; .$$

$$\hat{p}_1 = \frac{2792}{5412} = 0,5159 \qquad \hat{p}_2 = \frac{2776}{5978} = 0,4644 \; .$$

- Estatística de teste:

$$t_0 = \frac{\text{estimativa} - \text{valor admitido por hipótese}}{\text{erro padrão}} = \frac{(0,5159 - 0,4644) - 0}{0,0093677} = 5,49986 \; .$$

√ Estimativa está 5,50 erros padrão acima da diferença admitida por hipótese (parece grande!)

- Determinação do p-valor:
  - √ Distribuição amostral: N(0, 1) [gl = ∞]
    - Amostras são claramente grandes o suficiente

$$p - valor = \Pr\{T \geq 5,49986\} = 1,9 \times 10^{-8}.$$

- Conclusão:
  - √ Temos evidência muito forte contra H_0, que confirma a alternativa.
  - √ Temos evidência extremamente forte de que mães cujo 1º. filho foi uma menina *têm* maior chance de ter uma menina na segunda vez, do que mães cujo 1º. filho foi um menino.
- Quão mais provável?

√ IC de 95% para a verdadeira proporção (p_1–p_2)

$$(\hat{p}_1 - \hat{p}_2) \pm t_\infty \times \mathrm{ep}(\hat{p}_1 - \hat{p}_2) = (0,5159 - 0,4644) \pm 1,960 \times 0,0093677 = [0,033; 0,070] \; .$$

- Com 95% de confiança, a % populacional de meninas nascidas no 1º. grupo é maior do que no último grupo por algo entre 3 a 7 pontos percentuais.

Exemplo D – Outra Pergunta

- Ter um menino, seguido por uma menina é mais ou menos comum do que ter uma menina, seguida por um menino?

		Segundo filho		
		Menino	Menina	Total
Primeiro filho	Menino	3.202	2.776	5.978
	Menina	2.620	2.792	5.412
	Total	5.822	5.568	11.390

√ Contagens maiores para menino seguido por menina do que o contrário

69

Pesquisa Quantitativa - 2016

- Ordens de nascimento:
  √ {MM, MF, FM, FF}
- Situação de amostragem:
  √ Única amostra, resposta em qualquer uma das quatro ordens de nascimento (categorias)

	Nascimentos	Proporção amostral
MM	3.202	0,28
MF	2.776	0,24
FM	2.792	0,25
MM	2.620	0,23
Total	11.390	1,00

70

Pesquisa Quantitativa - 2016

- Hipótese de pesquisa (H_1):
  √ Ter um menino, seguido por uma menina é mais ou menos comum do que ter uma menina seguida por um menino
- Reação cética a H_1 (H_0)
  √ Não há diferença nas verdadeiras proporções das duas situação

71

Pesquisa Quantitativa - 2016

- Parâmetro de interesse:
  √ p_{MF}–p_{FM}: diferença entre as verdadeiras proporções
  - p_{MF}: verdadeira proporção de meninas de mães cujo 1º filho foi uma menino
  - p_{FM}: verdadeira proporção de meninos de mães cujo 1º filho foi uma menina
- Hipóteses
  √ H_0: $p_{MF} – p_{FM} = 0$ vs. H_1: $p_{MF} – p_{FM} \neq 0$
  - Não temos nenhuma ideia a respeito da direção na qual a diferença deveria cair, antes de observarmos os dados.

72

Pesquisa Quantitativa - 2016

√ Erro padrão:

– Uma amostra, várias categorias de resposta (situação 2)

$$\text{ep}(\hat{p}_{MF}-\hat{p}_{FM}) = \sqrt{\frac{\hat{p}_{MF}+\hat{p}_{FM}-(\hat{p}_{MF}-\hat{p}_{FM})^2}{n}} \qquad \hat{p}_{MF} = \frac{2.776}{11.390} = 0,24372$$

$$= \sqrt{\frac{0,24372+0,23003-(0,24372-0,23003)^2}{11.390}} \qquad \hat{p}_{FM} = \frac{2.620}{11.390} = 0,23003 \, .$$

$$= 0,006448 \, .$$

- Estatística de teste:

$$t_0 = \frac{\text{estimativa} - \text{valor admitido por hipótese}}{\text{erro padrão}}$$

$$= \frac{(0,24372-0,23003)-0}{0,006448} = 2,124098 \, .$$

√ Estimativa está 2,12 erros padrão acima da diferença admitida por hipótese (parece grande?)

73

Pesquisa Quantitativa - 2016

- Determinação do p-valor:
  - √ Distribuição amostral: N(0, 1) [gl = ∞]
    - – Amostras são claramente grandes o suficiente

$$p - valor = 2 \times \Pr\{T \geq 2,124098\} = 0,034.$$

- Conclusão:
  - √ Temos evidência contra H_0 e, portanto, evidência de que um menino seguido por uma menina é mais comum do que uma menina seguida por um menino.
    - – A diferença é positiva ($t_0 > 0$)!
- Quão mais comum?

74

Pesquisa Quantitativa - 2016

√ IC de 95% para a verdadeira proporção (p_{MF}–p_{FM})

$$(\hat{p}_{MF}-\hat{p}_{FM}) \pm t_{\infty} \times \text{ep}(\hat{p}_{MF}-\hat{p}_{FM}) =$$

$$(0,24372-0,23003) \pm 1,960 \times 0,006448 =$$

$$[0,001; 0,026] \, .$$

– Com 95% de confiança, o excesso da % populacional de meninos seguidos de meninas (MF) em relação a meninas seguidas de meninos (FM) está entre 0,1 (essencialmente 0) e 2,6 pontos percentuais.

75

Pesquisa Quantitativa - 2016

Exemplo E

- Adiamento da morte
  - √ Pessoas parecem adiar sua morte, até a ocorrência de algum evento de importância
  - √ Pesquisa secundária para verificar teoria
    - – Datas de nascimento e morte de 348 pessoas encontradas no livro *400 Americanos Notáveis*.
    - – 16 morreram no mês que antecedia seus aniversários
  - √ Há evidência de que as pessoas têm menos chance de morrer no mês anterior ao aniversário?

76

Pesquisa Quantitativa - 2016

- Parâmetro de interesse:
  - √ p: probabilidade de uma pessoa morrer no mês anterior ao de seu aniversário
- Hipóteses
  - √ H_0: p = $\frac{1}{12}$ vs. H_1: p < $\frac{1}{12}$
    - A teoria é de que é menos provável que a pessoa morra no mês que antecede seu aniversário.

77

Pesquisa Quantitativa - 2016

√ Erro padrão:

- Uma amostra com resposta binária

$$\mathrm{ep}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0,04598 \times 0,95402}{348}} = 0,01123\ .$$

$$\hat{p} = \frac{16}{348} = 0,04598\ .$$

- Estatística de teste:

$$t_0 = \frac{\text{estimativa} - \text{valor admitido por hipótese}}{\text{erro padrão}} = \frac{0,04598 - \frac{1}{12}}{0,01123} = -3,327\ .$$

√ Estimativa está 3,33 erros padrão abaixo da proporção admitida por hipótese (parece grande?)

78

Pesquisa Quantitativa - 2016

- Determinação do p-valor:
  - √ Distribuição amostral: N(0, 1) [gl = ∞]
    - A regra de 10% (Apêndice A3) dá *n* como sendo pelo menos 960, o que não é o caso

$$p-valor = \Pr\{T \leq -3,327\} = 0,0004.$$

- Conclusão:
  - √ Há evidência muito forte contra H_0 em favor de H_1, portanto, evidência de que é menos provável que a pessoa morra no mês que antecede seu aniversário.

79

Pesquisa Quantitativa - 2016

- Teste exato – saída R:

```
> prop.test(16, 348, p = 1/12, alternative = "less")

        1-sample proportions test with continuity correction

data:  16 out of 348, null probability 1/12
X-squared = 5.8777, df = 1, p-value = 0.007667
alternative hypothesis: true p is less than 0.08333333
95 percent confidence interval:
 0.00000000 0.06989543
sample estimates:
         p
0.04597701
```

- p-valor é bastante diferente daquele obtido com o teste aproximado (aproximação normal)

√ Há forte evidência contra H_0 em favor da hipótese de que é menos provável que a pessoa morra no mês que antecede seu aniversário.

80

Pesquisa Quantitativa - 2016

- Caso optássemos pela hipótese alternativa bilateral

$$\frac{1}{12} = 0,0833 \, .$$

√ H_1: $p \neq \frac{1}{12}$

```
> prop.test(16, 348, p = 1/12, alternative = "two.sided")

        1-sample proportions test with continuity correction

data:  16 out of 348, null probability 1/12
X-squared = 5.8777, df = 1, p-value = 0.01533
alternative hypothesis: true p is not equal to 0.08333333
95 percent confidence interval:
 0.02739140 0.07510479
sample estimates:
         p
0.04597701
```

√ Há forte evidência contra H_0 em favor de H_1

– No caso, levaria à mesma conclusão

- Cuidados com a extrapolação dos resultados:

√ Supomos que as pessoas 'comuns' têm o mesmo comportamento que as pessoas 'notáveis'?

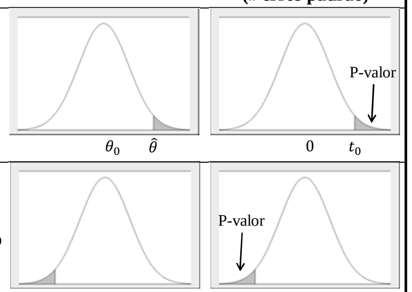
√ Há uniformidade ao longo do ano nas taxas de nascimento e de morte?

– Ex.: Se a maioria dos nascimentos fosse no verão e a maioria das mortes no inverno, por outras razões que não o adiamento da morte, a estimativa de *p* seria pequena

√ Teste da H_0: $\theta = \theta_0$

$$t_0 = \frac{\hat{\theta} - \theta_0}{\text{ep}(\hat{\theta})} \, .$$

Hipótese Alternativa	Evidência contra H_0 fornecida por	Escala $\hat{\theta}$ →	Escala t (# erros padrão)
H_1: $\theta > \theta_0$	$\hat{\theta}$ muito maior do que θ_0	θ_0 $\hat{\theta}$	P-valor; 0 t_0
H_1: $\theta < \theta_0$	$\hat{\theta}$ muito menor do que θ_0	$\hat{\theta}$ θ_0	P-valor; t_0 0

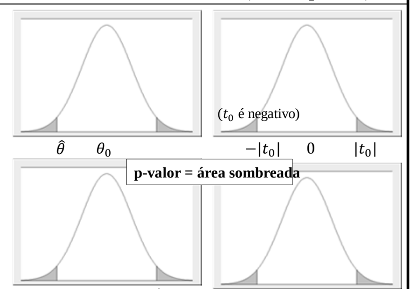


√ Teste da H_0: $\theta = \theta_0$

$$t_0 = \frac{\hat{\theta} - \theta_0}{\text{ep}(\hat{\theta})} \, .$$

Hipótese Alternativa	Evidência contra H_0 fornecida por	Escala $\hat{\theta}$	Escala t (# erros padrão)
H_1: $\theta \neq \theta_0$ (bilateral)	$\hat{\theta}$ muito longe de θ_0 (ambas as direções)	menor: $\hat{\theta}$ θ_0; maior: θ_0 $\hat{\theta}$	(t_0 é negativo) $-\lvert t_0\rvert$ 0 $\lvert t_0\rvert$; $-t_0$ 0 t_0

p-valor = área sombreada



Teste de Hipótese como Tomada de Decisão

Teste de Hipóteses

- Teste ao nível de significância de 5%:
  - √ Rejeitar H_0 se p-valor $\leq$ 5%
  - √ Se H_0 é verdadeira, p-valor $\leq$ 0,05 ocorrem apenas 5% das vezes
  - √ O que se busca?
    - Na maioria das vezes nossa decisão de rejeitar seria correta

86

Pesquisa Quantitativa - 2016

- H_0 tende a representar um ponto de vista cético
- Fixar um nível de significância muito baixo para evitar rejeitar H_0 verdadeira
  - √ Tentativa de tornar improvável a rejeição de H_0 falsas importantes

87

Pesquisa Quantitativa - 2016

- Efeitos significantes (em geral)
  - √ Efeitos ou diferenças para os quais o teste de hipóteses é significante ao nível de 5%, ou seja p-valor $\leq$ 0,05. (ou qualquer outro nível)
- Efeito não significante
  - √ Efeitos ou diferenças para os quais o p-valor é maior que o nível de significância do teste.
- Importante
  - √ p-valores de 0,052 e 0,048 devem ser tratados de maneira diferente?

88

Pesquisa Quantitativa - 2016

Teste de Significância de Nível Fixo

- Teste de hipóteses como processo de tomada de decisão

Decisão tomada	Situação real: H_0 é verdadeira	Situação real: H_0 é falsa
Não rejeita H_0 como verdadeira	Okay	Erro do tipo II
Rejeita H_0 como falsa	Erro do tipo I	Ok

√ Há muitas situações científicas nas quais continuaremos a aceitar uma teoria ou hipótese, a menos que tenha sido encontrada evidência contra ela.

Abordagens

- Fisher:
  - √ P-valores para medir o grau de evidência contra H_0
    - (testes de significância)
- Neyman-Pearson:
  - √ Teste de hipóteses de nível fixo
- Inferência Bayesiana
  - √ Ideias subjetivas do analista sobre o valor do parâmetro expressas na forma de distribuição de probabilidade
  - √ Atualização com as informações dos dados

Teste de Hipótese e Intervalos de Confiança

Intervalos de Confiança e Testes de Significância

- Intervalo de confiança:
  - √ Amplitude de valores que são possibilidades plausíveis para o verdadeiro valor do parâmetro (θ), num dado nível de confiança.
- Teste de significância:
  - √ Fornece uma medida da força da evidência contra um único valor admitido por hipótese (θ_0)

Relação entre Resultado de um Teste e dos Valores num IC

- Teste bilateral de H_0: $\theta = \theta_0$ é significante ao nível de 5% se, e somente se, θ_0 cair fora de um intervalo com 95% de confiança para θ.
  - √ No nível de 95%, dados fornecem evidência contrária a qualquer valor θ_0 que caia fora do IC
    - (Rejeitaríamos H_0: $\theta=\theta_0$ para qualquer valor que caia fora do IC)
  - √ Não fornece evidência, no nível de 5%, contrária a qualquer valor θ_0 que caia dentro do IC

93

Pesquisa Quantitativa - 2016

Exemplo

- Medidas da densidade da Terra – Cavendish
  - √ Hipóteses:
    - H_0: $\mu = 5{,}517$ vs. H_1: $\mu \neq 5{,}517$
- Estatísticas amostrais:

n	$\bar{x}$	s_X	ep($\bar{x}$)
23	5,4835	0,1904	0,0397

- IC com 95% de confiança para média μ:

$$\bar{x} \pm t_{gl} \times \text{ep}(\bar{x}) = 5,4835 \pm 2,073873 \times 0,0397 = [5,40; 5,57] .$$

  - √ Os testes de hipótese, ao nível de 5%, rejeitariam qualquer valor m_0 maior que 5,57 ou menor que 5,40

94

Pesquisa Quantitativa - 2016

Significância

- Significância estatística:
  - √ Se relaciona à existência de um efeito
  - √ p-valor pequeno:
    - Fornece evidência da existência do efeito, mas não informa sobre sua intensidade
- Significância prática:
  - √ Depende da intensidade do efeito
- Para estimar intensidade de um efeito construa um intervalo de confiança

95

Pesquisa Quantitativa - 2016

Exemplo

Determinação de se um efeito tem significância prática

- √ Estudo de doenças coronariana, descobriu-se que se o paciente fumava (ou não) era significativamente relacionado ao risco de morte
  - p-valor = 0,048
  - IC de 95% = [1,004; 2,5]
- √ Teste indica haver um aumento no risco (p-valor)
- √ Aumento real poderia ter sido minúsculo (de pouca importância prática ou poderia ter sido muito grande (mais que dobrado)

96

Pesquisa Quantitativa - 2016

Resultados Não Significantes – Interpretação

- Um teste não significante não é evidência de que H_0 é verdadeira
  - √ Podemos afirmar apenas que:
    - Dados não descartam a possibilidade de que H_0 seja verdadeira
- Usamos um IC, para verificar outros valores do parâmetro (diferente do especificado por H_0) são consistentes com os dados

Exemplo

- Teste de significância de correlação entre duas variáveis
  - √ $\alpha = 0{,}01$, com 1.000 observações, $r = 0{,}08$
  - √ Forte evidência de correlação não nula
  - √ Verdadeira correlação populacional situa-se provavelmente bastante perto do valor observado
  - √ Para efeitos práticos, é possível ignorar a associação entre essas variáveis, apesar de estarmos bastante confiantes de que a correlação é positiva

Importante

- Para evitar enganos:
  - √ Nunca cite um p-valor sobre a existência de um efeito, sem fornecer também um IC que estime seu tamanho

Como Mostrar que uma Hipótese é Verdadeira?

- Como mostrar que o verdadeiro valor de um parâmetro θ é θ_0?
  - √ O mais próximo que podemos chegar é quando todos os valores contidos em um intervalo de confiança apropriado de θ são essencialmente os mesmo que θ_0 para todos os propósitos práticos
- Para todos os propósitos práticos
  - √ Decisão da área de estudo, não estatística

Detecção de Diferenças Quando São Importantes

- Planejamento de uma investigação:
  - √ Determinação do tamanho da amostra
- Teste com nível de significância fixo
  - √ Evidência contra H_0: $\theta = \theta_0$ se o p-valor for menor que algum patamar (p-valor $\leq$ 0,05)

θ_{verd}: verdadeiro valor do parâmetro

- Quando $\theta_{verd} \neq \theta_0$:
  - √ Espera-se obter um resultado significante sempre que fosse coletada amostra de dados (impossível)
  - √ Desejável:
    - Probabilidade de obter resultado significante fosse a mais alta possível sempre que $\theta_{verd} \neq \theta_0$.
    - P{erro tipo II} fosse a mais baixa possível

Ideias

- Tamanho da amostra relativamente pequeno
  - √ Variabilidade amostral de uma estimativa é tipicamente grande
  - √ Pode não ser possível detectar diferença entre θ_{verd} e θ_0 mesmo quando essa é grande
    - Amostras pequenas estão sujeitas a erros amostrais grandes
  - √ Pode ser bem pequena a probabilidade de se obter resultado significante

- Tamanho da amostra é muito grande
  - √ Variabilidade amostral de uma estimativa é tipicamente pequena
  - √ É possível detectar com confiabilidade diferenças moderadas e grandes entre θ_{verd} e θ_0.
  - √ Pode ser possível detectar com bastante confiabilidade diferenças razoavelmente pequenas entre θ_{verd} e θ_0.
  - √ Probabilidade de obtenção de resultado significante pode ser muito alta

Processo para Determinação do Tamanho da Amostra

1. Determinação do menor valor de diferença (δ) a qual tenha importância prática ser detectada
   - √ $\delta = \theta_{\text{verd}} - \theta_0$
2. Suponha que a verdadeira diferença seja δ, calcula-se o menor tamanho de amostra necessário de modo que a probabilidade de se obter um resultado significante seja alta.

- Se um estudo não pode nem mesmo ser confiável para demonstrar uma diferença importante entre θ_{verd} e θ_0, o estudo provavelmente não vale a pena ser realizado.

Cálculo de Tamanho de Amostras para Testes de Significância

- São semelhantes, mas mais complicados que aqueles para intervalos de confiança
- Dependem de usar informação disponível sobre a variabilidade dos dados
  - √ Informação que tende a não ser confiável
- É difícil com frequência produzir um δ razoável e de aceitação geral

Estatística de Teste e o p-valor – Ideias Gerais

- p-valor:
  - √ Usado como medida da evidência contra H_0
- Classe de problemas (até agora)
  - √ H_0: $\theta = \theta_0$
  - √ Distribuição amostral:

$$t_0 = \frac{\text{estimativa} - \text{valor admitido por hipótese}}{\text{erro padrão}} \sim t_{\text{gl}}\ . \quad \boxed{\text{ou N(0, 1)}}$$

Classe de Problemas Mais Complexos

- Teste F
  - √ H_0: médias de 3 ou mais grupos são iguais
- Teste χ^2:
  - √ H_0: Probabilidades que governam como indivíduos se distribuem em tabela de contagem

Conceitos Ampliados

- Estatística de teste:
  - √ Medida da discrepância entre o que observamos nos dados e aquilo que esperaríamos observar se H_0 fosse verdadeira
  - √ Nos exemplos anteriores
    - Medida de discrepância = distância entre a estimativa e o valor hipotético

- p-valor:
  - √ Probabilidade de que a variação amostral sozinha produza dados que sejam mais discrepantes que o conjunto dos dados observados, calculada supondo-se que a hipótese nula seja verdadeira
  - √ Se a variação amostral, sob H_0, por si só produz dados que frequentemente estão ainda mais afastados do que os dados reais
    - Dados não fornecem evidência contra H_0.
  - √ Se a variação amostral dificilmente produz dados ainda mais afastados do que os dados reais
    - Dados evidenciam contra H_0

Dados de Variável Contínua – Questões de uma Amostra

Testes e Intervalos de Confiança para uma Média

- Teste t de uma hipótese sobre a média
  - √ Mostrará qualquer afastamento estatisticamente significante de H_0
- Intervalo de confiança da média
  - √ Informará o tamanho do afastamento
- Suposição para uso
  - √ Amostra proveniente de população normal
  - √ Em geral, para amostras grande essa suposição não é necessária

Exemplo

Concentração de íon nitrato (μg/ml) em amostra d'água

- √ 10 medidas de concentração do composto
  - – Pequeno conjunto de dados
- √ Preocupação de que a concentração tivesse sido alterada com relação ao valor original
  - – Valor desejado: 0,492

Procedimento

- Parâmetro de interesse:
  - √ μ: concentração média verdadeira
- Hipóteses:
  - √ H_0: $\mu = 0{,}492$
  - √ H_1: $\mu \neq 0{,}492$

- Carregamento dos dados:

```
            # Importação do conjunto de dados

concentracao <- scan(file = "composto.txt",sep=":", dec=",")
concentracao
 [1] 0.513 0.524 0.529 0.481 0.492 0.499 0.518 0.490 0.494 0.501
```

- Estatísticas descritivas:

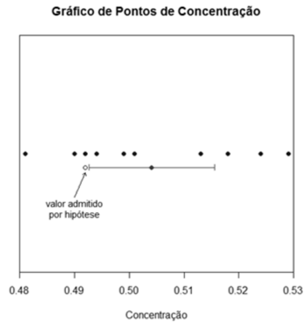
```
            # Estatísticas descritivas

library(pastecs)
stat.desc(concentracao, basic = FALSE)
```

- Saída:

Margem de erro

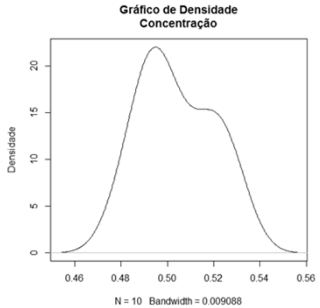
```
> stat.desc(concentracao, basic = FALSE)
      median         mean      SE.mean CI.mean.0.95          var      std.dev
 0.500000000  0.504100000  0.005060632  0.011447946  0.000256100  0.016003125
    coef.var
 0.031745933
```

- Gráfico de pontos:



√ Não há nada no gráfico de pontos sugerindo que os dados não pudessem se originar de uma distribuição normal

- Gráfico de densidade



√ Pequenas amostras podem produzir flutuações na estimativa da densidade

√ Não há indícios de não normalidade dos dados

- Execução do teste – comando:

```
# Criação do objeto com os resultados do teste t

h0 <- 0.492
concentracao.t <- t.test(concentracao, y = NULL, alternative = "two.sided",
                         mu = h0, paired = FALSE, conf.level = 0.95)
```

- Resultado - saída:

```
> concentracao.t

        One Sample t-test

data:  concentracao
t = 2.391, df = 9, p-value = 0.04049
alternative hypothesis: true mean is not equal to 0.492
95 percent confidence interval:
 0.4926521 0.5155479
sample estimates:
mean of x
   0.5041
```

p-valor

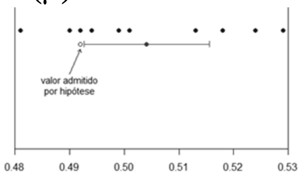

```
> concentracao.t

        One Sample t-test

data:  concentracao
t = 2.391, df = 9, p-value = 0.04049
alternative hypothesis: true mean is not equal to 0.492
95 percent confidence interval:
 0.4926521 0.5155479
sample estimates:
mean of x
   0.5041
```

- Conclusão

√ Evidência, a um nível de 5%, de que a concentração sofreu alteração com relação a 0,492

  - p-valor pequeno

- Intervalo de confiança para a média verdadeira (μ):



√ Com 95% de confiança, a verdadeira concentração cai entre 0,493 e 0,516.
  – Valor verdadeiro está muito próximo do limite inferior do intervalo

√ Alteração pode ter sido pequena

121

Pesquisa Quantitativa - 2016

Amostra Pequena – Consequências

- Nunca saberemos se os dados são normais
- Testes e intervalos t são robustos a desvios da hipótese de normalidade
- Robusto:
  √ Procedimento ainda "funciona" quando amostramos de distribuição diferente da normal

122

Pesquisa Quantitativa - 2016

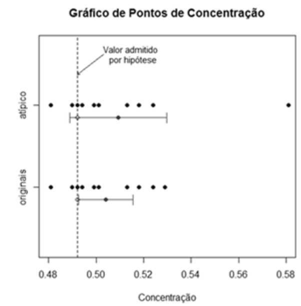
- Funcionar significa:
  √ Os intervalos de confiança contêm o verdadeiro valor (μ) em aproximadamente 95% das amostras extraídas
  √ Obtemos um resultado "significante ao nível de 5%" em aproximadamente 5% das vezes
- Testes e intervalos t não são robustos a não normalidade severa
  √ Ex.: presença de outliers
  √ $\bar{x}$ e s_X são muito sensíveis a valores atípicos

123

Pesquisa Quantitativa - 2016

Exemplo – Valor Atípico

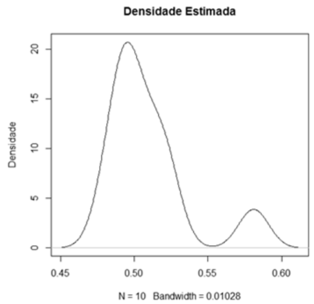
- Maior observação deslocada para a direita
  √ Trocado 0,529 por 0,581



124

Pesquisa Quantitativa - 2016

- Gráfico de densidade



√ Há indício de não normalidade dos dados

√ Outlier produz aumento da densidade na cauda superior (atipicidade)

- Execução do teste – comando:

```
              # Perturbação - introdução do valor atípico
atipico <- sort(composto)[-10]    # insere outlier
atipico[10] <- 0.581
              # Criação do objeto com os resultados do teste t
atipico.t <- t.test(atipico, y = NULL, alternative = "two.sided",
                       mu = h0, paired = FALSE, conf.level = 0.95)
```

- Resultado - saída:

p-valor

```
> atipico.t

        One Sample t-test

data:  atipico
t = 1.9172, df = 9, p-value = 0.08745
alternative hypothesis: true mean is not equal to 0.492
95 percent confidence interval:
 0.4888875 0.5297125
sample estimates:
mean of x
   0.5093
```


Comparação entre os Resultados

Estatísticas amostrais	Original	Atípico
$\bar{x}$	0,5041	0,5093
s_X	0,0160	0,02853
ep($\bar{x}$)	0,00506	0,00902
t_0	2,39	1,92
p-valor	0,040	0,087
IC com 95%	[0,49265; 0,51555]	[0,48889; 0,52971]

- Valores amostrais
  - √ Média se deslocou para a direita
    - Se afastou de μ_0
  - √ Desvio padrão e erro padrão aumentaram
  - √ Distância t_0 diminuiu
  - √ P-valor aumentou
  - √ IC mais impreciso
- A conclusão é diferente!

Resultado de Testes e Intervalos t – Desvio de Normalidade

- Assimetria
  - √ Testes t bilaterais e intervalos t
    - Pode-se tolerar muita assimetria na distribuição, mesmo em amostras pequenas
  - √ Testes t unilaterais
    - Podem ser muito afetados por assimetria pronunciada (amostras menores que 40)
    - Quanto menor a amostra, menor o grau de assimetria que pode ser tolerado

- Curtose (caudas pesadas)
  - √ Processos aleatórios com a característica de produzirem valores extremos
  - √ Afeta bastante os procedimentos
  - √ Caudas pesadas são difíceis de serem detectadas em pequenas amostras
- Outliers
  - √ Em amostras pequenas, levam a não confiar nos resultado do teste t e nos resultados dos intervalos t

129

Pesquisa Quantitativa - 2016

Exploração Gráfica

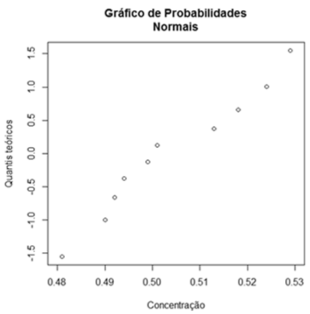
- Sempre represente graficamente os dados antes de usar ferramentas formais de análise
- Análise gráfica univariada:
  - √ Gráfico de pontos
  - √ Histograma
  - √ Gráfico de probabilidade normal

130

Pesquisa Quantitativa - 2016

Gráfico de Probabilidade Normal

- Ferramenta gráfica para detectar a falta de normalidade dos dados



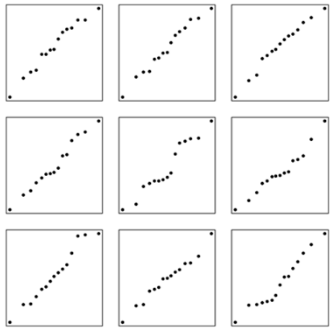
- Se dados provêm de normal
  - √ QQ plot será aproximadamente uma reta

131

Pesquisa Quantitativa - 2016

- Dados amostrados de normal podem ser bem diferentes de uma reta

Simulação amostras normais
(n = 15)



- √ Muita flutuação, com valores deslocados
  - principalmente nos extremos
- √ Pode ocorrer:
  - Conjuntos de dados pequenos (n<15) e mesmo moderados (15≤n≤40)
- √ Aconselha-se comparar o gráfico de seus dados com gráficos de vários “conjuntos de dados” simulados de mesmo tamanho.

132

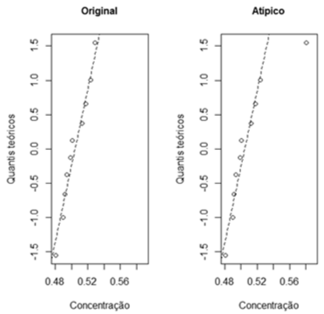
Pesquisa Quantitativa - 2016

Teste de Normalidade

- Hipóteses:
  - √ H_0: dados provêm de distribuição normal
  - √ H_1: dados não provêm de distribuição normal
- P-valor:
  - √ Valor pequeno é evidência contra H_0, em favor da hipótese de não normalidade dos dados
- Procedimentos mais comuns
  - √ Shapiro-Wilk
  - √ Anderson-Darling

Exemplo

- Gráfico de normalidade concentração
  - √ Dados originais e com perturbação (outlier)



- Dados originais:
  - √ Gráfico sugere que dados provêm de população normal
- Dados com perturbação
  - √ Gráfico não indica amostra de distribuição
  - √ Gráfico destaca outlier
    - Fora da tendência dos outros pontos

- Testes de normalidade:
  - √ Dados originais

```
> shapiro.test(concentracao)

 Shapiro-Wilk normality test

data:  concentracao
W = 0.94658, p-value = 0.6283
```

  - √ Teste sugere que a hipótese de normalidade não é inaceitável
    - p-valor não forneceu qualquer evidência contrária à hipótese de normalidade (p > 10%)

√ Dados com valor atípico:

```
> shapiro.test(atipico)

 Shapiro-Wilk normality test

data:  atipico
W = 0.7995, p-value = 0.0143
```

√ Teste sinaliza forte evidência contrária à hipótese de normalidade
- p-valor de aproximadamente 0,01

Procedimentos t – Recomendações de Uso

- Procedimentos t são amplamente utilizados devido a sua robustez
- Não são aplicáveis quando:
  - √ Dados mostrarem estarem separados em clusters
  - √ Situações com mais de uma moda

Procedimentos t – Cuidados

- Pequenas amostras ($n < 15$)
  - √ Não use procedimentos t na presença de valores atípicos
  - √ Não use procedimentos t unilaterais se dados forem mais que ligeiramente assimétricos

- Amostras moderadas ($15 \leq n \leq 40$):
  - √ Não use procedimentos t se houver outliers
  - √ Não use os procedimentos t se os dados forem fortemente assimétricos
- Amostras grandes ($n > 40$)
  - √ Procedimentos t funcionam bem mesmo para dados nitidamente assimétricos
  - √ Podem ser afetados adversamente por valores atípicos grosseiros

Lidando com Valores Atípicos

- Se parece haver valores atípicos nos dados
  - √ Verificar valores a partir das fontes originais
    - Observações podem ser erros
  - √ Se forem erros:
    - Se possível, devem ser corrigidos
    - Caso contrário, devem ser removidos do conjunto de dados

- Se houver dúvidas sobre a legitimidade das observações:
  - √ Verifique se a presença dos mesmos altera substancialmente as conclusões alcançadas a partir dos dados

Enfrentando a Não normalidade

- Usar métodos não paramétricos
  - √ Não fazem nenhuma hipótese a respeito da distribuição subjacente dos dados
- Usar métodos robustos
  - √ Procedimentos insensíveis à presença de valores atípicos
- Transformar os dados:
  - √ Pode dificultar a interpretação dos resultados

Referências

Bibliografia Recomendada

- AGRESTI, A.; FINLAY, B. *Métodos estatísticos para as ciências sociais.* Penso, 2012.
- MOORE, D. S.; MCCABE, G. P.; DUCKWORTH, W. M.; SLOVE, S. L. *A prática da estatística empresarial: como usar dados para tomar decisões*. LTC, 2006.
- WILD, J. W.; SEBER, G. A. F. *Encontros com o acaso: um primeiro curso de análise de dados e inferência*. LTC, 2004.