

**Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Estatística**

**Teste de Significância via
Testes de Permutação na
Comparação de Médias em
Pequenas Amostras**

M. I. Zeghzeghi & F. R. B. Cruz

**Relatório Técnico
RTE-01/2010**

**Relatório Técnico
Série Ensino**

Teste de Significância via Testes de Permutação na Comparação de Médias em Pequenas Amostras

Miled Ibrahim Zeghzeghi & Frederico R. B. Cruz

Departamento de Estatística - ICEX - UFMG

zeghi1000@hotmail.com, fcruz@est.ufmg.br

Resumo

Neste trabalho apresentamos uma comparação entre duas formas de verificação de diferenças entre médias. Uma delas é a forma clássica, que faz uso de modelos matemáticos. A outra é via reamostragem, da família dos métodos computacionalmente intensivos, os quais incluem os métodos *bootstrap* e os métodos Monte Carlo. Lidamos inicialmente com dados simulados, de diferentes distribuições, discretas e contínuas, para analisarmos depois dados reais. Os resultados confirmam a superioridade dos métodos de reamostragem, quando as amostras são pequenas e assimétricas.

Palavras-chave: Testes de permutação, teste t , valor- p , Monte Carlo, teste de hipóteses, inferência com base em duas amostras.

1. Introdução

1.1 Preliminares

A estatística tem ocupado um lugar de crescente destaque no dia a dia, seja no setor acadêmico, seja no público ou ainda diretamente no mercado, em relevantes aplicações práticas. No campo acadêmico, as diversas áreas do conhecimento (ciências biológicas, ciências exatas e engenharias, lingüística, letras e artes e humanidades) são beneficiadas pelo uso da estatística, como ferramenta fundamental para análise dos dados obtidos em suas pesquisas. Nos órgãos do governo, os modelos estatísticos ajudam a planejar investimentos e a avaliar políticas sociais e econômicas. No mercado, lidar com incertezas é uma constante. Profissionais capazes de tomar decisões de forma rápida e eficaz são extremamente desejáveis. Assim, as técnicas estatísticas são uma

importante ferramenta de auxílio aos profissionais, na organização, disposição e análise das informações que compõem o cenário em estudo.

Conforme afirma Triola (2005), as famílias, o governo e as empresas utilizam largamente dados estatísticos, mesmo sem se darem conta disso. Taxas de desemprego, índices de preços ao consumidor, índices de inflação, entre outros, são calculados cuidadosamente por meio de sofisticadas ferramentas estatísticas e os resultados (adequadamente organizados e dispostos por conhecidas ferramentas estatísticas de descrição e apresentação de dados), são utilizados por empresários na tomada de importantes decisões que afetam a contratação de empregados, o nível de produção e a expansão da economia. Triola (2005) alerta ainda para o uso e o abuso da estatística, por descuido, ignorância ou má fé, pela apresentação de dados de forma enganosa, para alcançar objetivos pessoais. Enfim, todas estas observações reforçam ainda mais a importância do domínio das técnicas estatísticas.

Tabela 1.1: Médicos por sexo e região no Brasil em 1995

| Brasil | Homens | | Mulheres | | Total | |
|---------------|----------------|-------------|-----------------|-------------|----------------|--------------|
| | v. abs. | % | v. abs. | % | v. abs. | % |
| Norte | 3.763 | 66,0 | 1.935 | 34,0 | 5.698 | 100,0 |
| Nordeste | 17.224 | 58,9 | 12.025 | 41,1 | 29.249 | 100,0 |
| Sudeste | 72.656 | 67,5 | 34.909 | 32,5 | 107.565 | 100,0 |
| Sul | 20.559 | 73,1 | 7.549 | 26,9 | 28.108 | 100,0 |
| Centro-Oeste | 8.951 | 72,0 | 3.481 | 28,0 | 12.432 | 100,0 |
| Total | 123.153 | 67,3 | 59.899 | 32,7 | 183.052 | 100,0 |

Fonte: Pesquisa “Perfil dos Médicos no Brasil”, Fiocruz

A Tabela 1.1 e a Figura 1.1 apresentam exemplos interessantes do uso da estatística. É apresentado o resultado de um estudo sobre o perfil dos médicos no Brasil, no ano de 1995, com relação ao sexo. Na Tabela 1.1, podemos observar que a maioria dos médicos no Brasil é do sexo masculino (aproximadamente 67% de homens contra 33% de mulheres). Além disso, a região sul apresenta o *maior* desequilíbrio entre os sexos (aproximadamente 73% de homens contra 27% de mulheres). A região que apresenta a

menor diferença entre homens e mulheres é a nordeste (aproximadamente 59% contra 41%, respectivamente).

O gráfico representado na Figura 1.1 apresenta o perfil de idade dos médicos no Brasil, em 1995. Nota-se que a média de idade dos profissionais em 1995 era em torno de 42 anos. Este resultado era aparentemente independente do local de residência (capital, com média de 42.3 anos, e interior, com média de 41.7 anos). Uma questão de interesse é saber se essa pequena diferença entre médias seria significativa, ou apenas uma flutuação aleatória na amostra analisada.

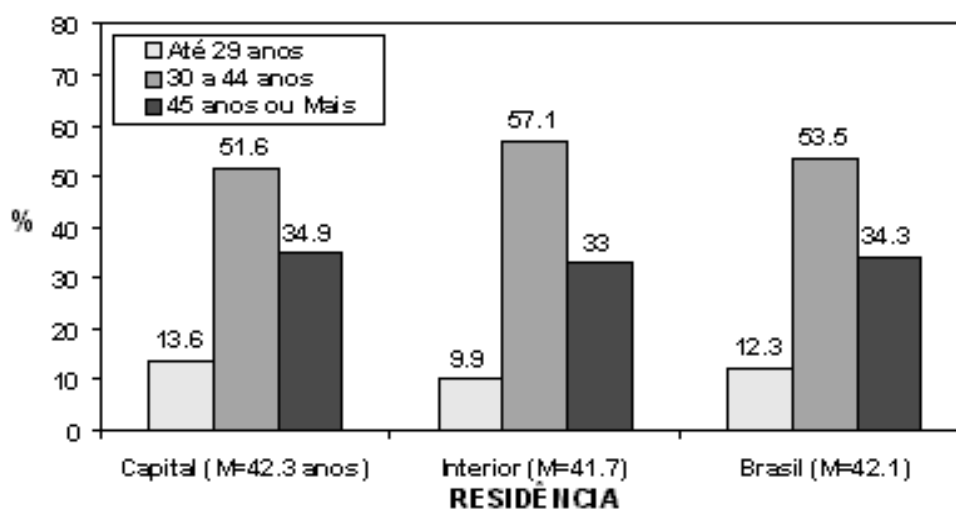


Figura 1.1: Perfil de idade de médicos conforme sua localização (capital ou interior)

Neste trabalho apresentaremos uma comparação entre duas formas de verificação de diferenças entre médias. Uma delas é a forma clássica, que faz uso de modelos matemáticos para a distribuição das populações. Dependendo de diversos fatores, tais como o tamanho da amostra (grande ou pequena), tipo de distribuição dos dados (normal, não-normal), conhecimento ou não da variância, há métodos específicos para o teste entre médias de duas populações. Maiores detalhes serão apresentados nas próximas seções.

Também temos a família dos métodos computacionalmente intensivos. Incluem-se nesta categoria os métodos *bootstrap* (detalhes em Montgomery & Runger, 2003), métodos de reamostragem (maiores informações em Duczmal et al., 2003) e métodos Monte

Carlo (ver Santos, 1998). Neste trabalho focaremos os métodos de reamostragem. Maiores detalhes serão apresentados a seguir.

1.2 Motivação

A motivação deste estudo baseia-se no fato de que em certas situações de interesse prático os pressupostos para aplicação dos métodos clássicos não se aplicam ou são de difícil verificação. Assim, tem-se como objetivo principal deste trabalho a análise de desempenho de testes de permutação, frente aos testes clássicos de diferenças entre médias.

1.3 Organização

Na Seção 2 são apresentados resumidamente os testes clássicos para comparação entre médias de duas populações. A Seção 3 detalha os testes de permutação, ressaltando sua importância para os tipos de dados aqui tratados. A Seção 4 descreve o desenvolvimento da rotina de programação (*macro*) que foi construída para realização do teste de permutação para diferenças entre médias. Finalmente, a Seção 5 conclui este trabalho com algumas observações finais.

2. Testes Clássicos para Comparação entre Médias

Sobre os testes clássicos para comparação entre médias, podemos dizer que existem situações específicas que direcionam sua aplicação (ver, por exemplo, Soares & Siqueira, 1999). A depender de as variâncias σ_1^2 e σ_2^2 serem conhecidas ou não, de os dados serem pareados ou não e de qual são os tamanhos das amostras n_1 e n_2 , teremos para cada situação abordagens específicas. As subseções a seguir descrevem os testes clássicos específicos a cada situação, a começar pelo teste da diferença das médias μ_1 e μ_2 com as variâncias σ_1^2 e σ_2^2 conhecidas.

2.1 Teste de Hipótese para Diferença de Médias com Variância Conhecida

Inicialmente consideraremos o teste de hipóteses relacionado com a comparação entre as médias μ_1 e μ_2 de duas populações, tomando-se como premissa que as variâncias σ_1^2 e σ_2^2 são conhecidas. Para o teste de hipóteses da diferença das médias, atribuímos ao resultado $(\mu_1 - \mu_2)$ um valor especificado Δ_0 , assim, estabeleceremos como hipótese nula:

$$H_0: \mu_1 - \mu_2 = \Delta_0.$$

Em muitos casos, faremos simplesmente $\Delta_0 = 0$. Assim estaremos testando a igualdade das duas médias μ_1 e μ_2 . A estatística de teste para este caso é

$$Z_0 = (\bar{X}_1 - \bar{X}_2 - \Delta_0) / \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_1}},$$

que assumirá (sob H_0) uma distribuição normal padrão $N(0,1)$. Podemos ainda afirmar que encontrar uma diferença considerável entre os valores das amostras, $\bar{x}_1 - \bar{x}_2$, em comparação com Δ_0 , é uma evidência de que a hipótese alternativa, H_1 , é verdadeira.

Como (quando a hipótese nula H_0 for verdadeira) Z_0 tem uma distribuição normal padrão, $N(0,1)$, tomaremos $-z_{\alpha/2}$ e $z_{\alpha/2}$ como os limites da região crítica. Isto fará com que o teste tenha um nível de significância α . Regiões críticas para alternativas unilaterais terão tratamento semelhante.

O quadro a seguir sumariza o procedimento usado para a comparação de médias com variância conhecida.

| Teste de Hipóteses para $\mu_1 - \mu_2$ com Variâncias Conhecidas | |
|--|---|
| Hipótese nula: $H_0: \mu_1 - \mu_2 = \Delta_0$ | |
| Estatística de teste: $Z_0 = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_1}}}$ | |
| <u>Hipóteses Alternativas</u> | <u>Critério de Rejeição</u> |
| $H_1: \mu_1 - \mu_2 \neq \Delta_0$ | $Z_0 > z_{\alpha/2}$ ou $Z_0 < -z_{\alpha/2}$ |
| $H_1: \mu_1 - \mu_2 > \Delta_0$ | $Z_0 > z_{\alpha}$ |
| $H_1: \mu_1 - \mu_2 < \Delta_0$ | $Z_0 < -z_{\alpha}$ |

2.2 Teste de Hipóteses para Diferença de Médias com Variância Desconhecida

Para testar a diferença entre as médias, $(\mu_1 - \mu_2)$, de duas distribuições normais com as variâncias σ_1^2 e σ_2^2 desconhecidas, usaremos a estatística t para o teste de hipóteses. No caso anterior supomos que as distribuições são normais, entretanto pequenos desvios não afetam o procedimento.

Existem duas situações que requerem tratamentos diferenciados. No primeiro caso, as variâncias são desconhecidas, mas são iguais, isto é, $\sigma_1^2 = \sigma_2^2 = \sigma^2$. No segundo caso, σ_1^2 e σ_2^2 são desconhecidas, porém não necessariamente iguais, isto é, $\sigma_1^2 \neq \sigma_2^2$.

O quadro a seguir apresenta um sumário do teste de hipóteses de diferença de médias com as variâncias desconhecidas e *iguais*. Detalhes sobre o teste para variâncias desconhecidas e *diferentes* não serão apresentados aqui, mas podem ser encontrados em Montgomery & Runger (2003).

| <u>Teste t Combinado para Duas Amostras</u> | |
|--|---|
| Hipótese nula: $H_0: \mu_1 - \mu_2 = \Delta_0$ | |
| Estatística de teste: $T_0 = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ | |
| <u>Hipóteses Alternativas</u> | <u>Critério de Rejeição</u> |
| $H_1: \mu_1 - \mu_2 \neq \Delta_0$ | $t_0 > t_{\alpha/2; n_1+n_2-2}$ ou $t_0 < -t_{\alpha/2; n_1+n_2-2}$ |
| $H_1: \mu_1 - \mu_2 > \Delta_0$ | $t_0 > t_{\alpha/2; n_1+n_2-2}$ |
| $H_1: \mu_1 - \mu_2 < \Delta_0$ | $t_0 < -t_{\alpha/2; n_1+n_2-2}$ |

2.3 Teste t Emparelhado

Nas seções anteriores, para o teste t , foram coletados dados de amostras em pares provenientes de observações homogêneas, mas estas condições podem mudar dependendo da população pesquisada. Quando os pares são coletados de amostras dependentes chamamos o teste t de *emparelhado*. Este procedimento torna o teste mais poderoso.

Em um conjunto de n observações emparelhadas, $(X_{11}, X_{21}), (X_{12}, X_{22}), \dots, (X_{1n}, X_{2n})$, provenientes de duas populações representadas por X_1 e X_2 , com médias μ_1 e μ_2 , e variâncias σ_1^2 e σ_2^2 , respectivamente, definiram-se as diferenças entre cada par das observações por

$$D_j = X_{1j} - X_{2j}, j = 1, 2, \dots, n.$$

As diferenças D_j são consideradas como distribuídas normalmente, com média μ_D e variância σ_D^2 . Assim, o teste de hipóteses da diferença entre μ_1 e μ_2 pode ser feito através do teste t para μ_D :

$$\mu_D = E(X_1 - X_2) = E(X_1) - E(X_2) = \mu_1 - \mu_2.$$

Testar

$$H_0: \mu_1 - \mu_2 = \Delta_0, \text{ contra}$$

$$H_1: \mu_1 - \mu_2 \neq \Delta_0,$$

é equivalente a testar

$$H_0: \mu_D = \Delta_0, \text{ contra}$$

$$H_1: \mu_D \neq \Delta_0.$$

A seguir, um quadro com o teste.

Teste t Emparelhado

Hipótese nula:

$$H_0: \mu_D = \Delta_0$$

Estatística de teste:

$$T_0 = \frac{\bar{D} - \Delta_0}{S_D / \sqrt{n}}$$

Hipóteses Alternativas

$$H_1: \mu_D \neq \Delta_0$$

$$H_1: \mu_D > \Delta_0$$

$$H_1: \mu_1 - \mu_2 < \Delta_0.$$

Critério de Rejeição

$$t_0 > t_{\alpha/2; n_1+n_2-2} \text{ ou } t_0 < -t_{\alpha/2; n_1+n_2-2}$$

$$t_0 > t_{\alpha/2; n_1+n_2-2}$$

$$t_0 < -t_{\alpha/2; n_1+n_2-2}$$

3. Testes de Permutação

No teste de permutação, comparamos o valor da estatística observada com as saídas apresentadas através do algoritmo computacional, que consiste em criar reamostras por simples permutações dos dados originais (maiores detalhes podem ser vistos, por exemplo, em Duczmal et al., 2003). Se estas saídas produzem estatísticas com valores semelhantes à estatística observada nos dados, então a estatística observada terá sido um mero acaso. O teste terá validade, desde que se mantenha a estrutura dos dados.

Os testes de permutação apresentam uma vantagem sobre os testes clássicos, que é não necessitar da suposição de aleatoriedade na coleta dos dados, e nem mesmo de uma suposição inicial da função de densidade da estatística em estudo (Edgington, 1995)

4. Desenvolvimento de uma Macro para o Minitab®

Apresentaremos agora os detalhes do desenvolvimento de uma macro para o Minitab® 15. Inicialmente apresentamos um algoritmo que descreve um teste de permutação para o teste de diferenças entre médias de duas populações, como pode ser visto na Figura 4.1, a seguir.

```
algoritmo  
  leia amostra1, amostra2  
  calcule a diferença entre as médias da amostra1 e da amostra2 (diferença observada)  
  combine amostra1 e amostra2 e coloque em amostra3 (simulação da hipótese nula)  
  repetir  
    reparta elementos de amostra3 entre amostra4 e amostra5  
    calcule a diferença entre as médias de amostra4 e amostra5  
    armazene a diferença calculada anteriormente  
    atualize o valor-p, por comparação com a diferença observada  
  até número de replicações ser alcançado  
  escreva resultados  
fim algoritmo
```

Figura 4.1: Algoritmo para teste de permutação

Note-se que o algoritmo é iniciado com a leitura das amostras das duas populações de interesse e com o cálculo da diferença observada nas duas amostras originais. As duas amostras são então combinadas, como forma de simular a hipótese nula, H_0 , isto é, como meio de simular que as amostras vêm de uma mesma população. Por um número de vezes pré-estabelecido, os elementos da amostra combinada são repartidos entre duas reamostras e a diferença entre as médias destas reamostras é calculada e acumulada, para futura análise. O valor- p é então estimado. Tal estimativa é feita pela comparação entre a diferença observada entre as médias das amostras originais e a diferença entre as médias das reamostras.

É importante ressaltar o papel desempenhado pelo número de simulações Monte Carlo, que deve ser criteriosamente escolhido para garantir um compromisso entre uma aproximação de boa qualidade e um baixo tempo computacional. Em outras palavras, se o número de simulações for muito grande, o tempo de processamento pode ser proibitivo e, por outro lado, se for muito baixo, uma estimativa muito grosseira do valor- p será obtida. Depois de alguma experimentação, verificamos que 1.000 era um valor adequado para o número de simulações Monte Carlo, para os casos testados.

Note que se os tamanhos das duas amostras forem *pequenos*, podemos partir para a enumeração completa de todas as reamostras possíveis, em lugar de usar um número fixo de simulações, igual a 1.000 (p.e., para duas amostras de tamanho 9, teremos exatamente 48.620 reamostras possíveis, que é o número de combinação de 18 elementos 9 a 9).

Outra observação importante é que o número de simulações pode ser mantido em aberto. No caso dos testes Monte Carlo sequenciais, o número de simulações não é fixado a priori. Os testes prosseguem até que uma determinada regra de parada seja alcançada. Essa estratégia pode reduzir consideravelmente o tempo computacional para se chegar a uma aproximação aceitável do valor- p . Maiores detalhes sobre os testes Monte Carlo sequenciais podem ser obtidos no recente artigo de Silva et al. (2009).

5. Resultados Experimentais

Apresentaremos e discutiremos agora os resultados obtidos. Inicialmente, lidamos com dados simulados. Testamos amostras *únicas* de várias distribuições (simétricas, assimétricas, contínuas e discretas). Também utilizamos vários tamanhos de amostras. Inicialmente, apresentamos os resultados para dados simulados.

5.1 Dados Simulados

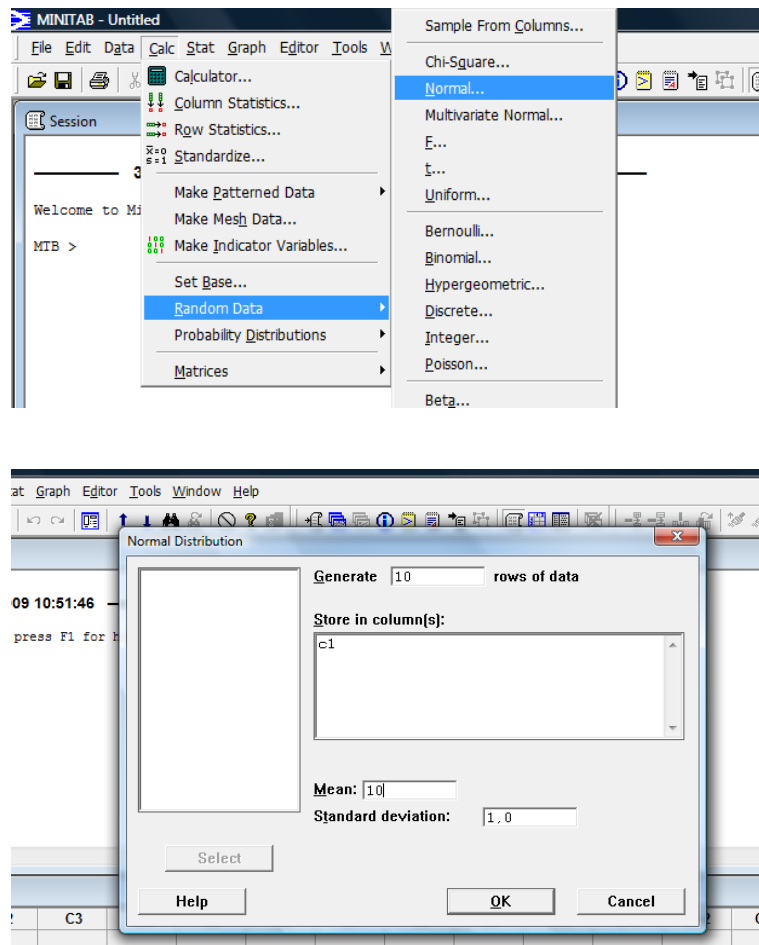


Figura 5.1: Geração de uma amostra $N(10; 1)$ de tamanho 10

A geração de dados com uso do Minitab® é bastante fácil e intuitiva. Por exemplo, a geração de dados de uma distribuição normal, com média 10, desvio-padrão 1, de tamanho 10, pode ser conseguida com simples cliques de mouse, conforme ilustrado na Figura 5.1.

A Figura 5.1 mostra também a caixa de diálogo resultante do comando para geração da amostra aleatória. Note que a geração de outras distribuições (por exemplo, exponencial, Poisson etc.) é igualmente simples e facilmente acessível via menus.

A chamada à macro desenvolvida não é também de grande complexidade. Na janela de seção do Minitab®, podemos executar o teste de permutação por meio do comando ilustrado na Figura 5.2. Note que a macro *precisa* estar no diretório padrão do Minitab® (por exemplo, “C:\Program Files\MINITAB 15\English\Macros”), para que o comando funcione.

```
#primeira opção de chamada
MTB> %permuta C3 C4
...
#segunda opção de chamada
MTB> %permuta C3 C4;
SUBC> hist.
...
```

Figura 5.2: Opções de utilização da macro desenvolvida

Na Figura 5.2 vemos as duas formas pelas quais a macro *permuta* pode ser chamada. Na primeira forma, queremos verificar a diferença entre médias de duas amostras disponíveis nas colunas C3 e C4 do Minitab®. Note que nenhuma restrição é feita quanto ao tamanho (se são iguais ou não) destas duas amostras. A segunda forma de chamada tem a mesma função, mas adiciona-se ao resultado a geração de um histograma das diferenças das reamostras, para uma melhor comparação com a diferença observada.

Para validação dos resultados da macro *permuta*, realizaremos comparações com os testes clássicos de comparação entre médias. A realização de tais testes no Mintab® é também bastante simples, conforme pode ser visto na Figura 5.3.

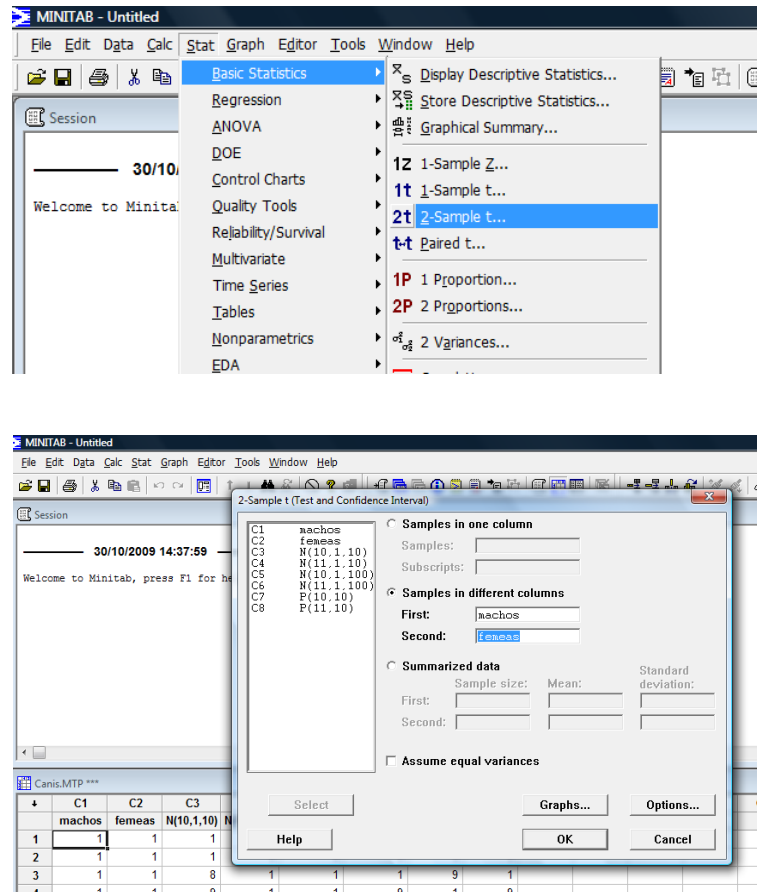
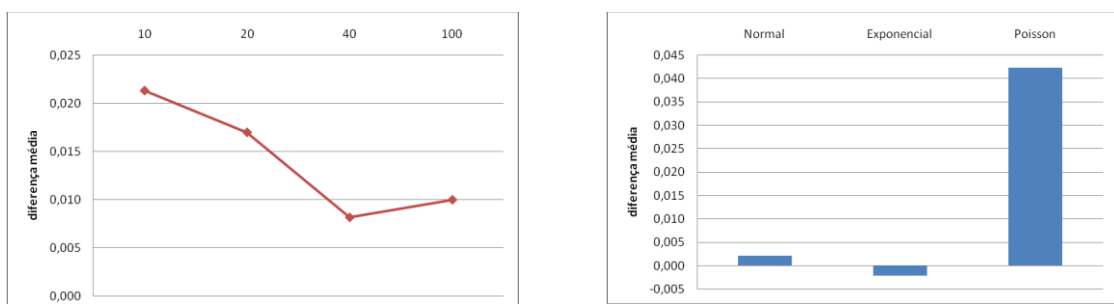


Figura 5.3: Realização do teste clássico de diferença entre médias de duas populações



a) diferenças em função do tamanho da amostra

b) diferenças em função da distribuição

Figura 5.4: Diferenças médias entre o teste de permutação e o método clássico

A Tabela 5.1 apresenta os resultados obtidos para amostras *únicas* provenientes de diversas distribuições. Em geral, notamos que as diferenças entre os valores-*p*, pelo teste de permutação e pelo método clássico, são desprezíveis, para amostras grandes ou

para grandes diferenças entre as médias, δ , das duas populações, independente da distribuição considerada. Esse comportamento é bastante intuitivo e esperado. Por outro lado, essas diferenças crescem, à medida que os δ diminuem, principalmente em amostras pequenas e distribuições assimétricas.

Tabela 5.1: Resultados para amostras *únicas* de diversas populações e tamanhos

| Distribuição | | Tamanho da amostra | Valor- p | | |
|--|------------------|--------------------|------------|----------|-----------|
| População 1 | População 2 | | Permutação | Clássico | Diferença |
| Normal(10; 1.) $\delta = 11-10 = 1$ | Normal(11; 1) | 10 | 0,061 | 0,065 | -0,004 |
| | | 20 | 0,228 | 0,208 | 0,020 |
| | | 40 | 0,001 | 0,000 | 0,001 |
| | | 100 | 0,000 | 0,000 | 0,000 |
| Normal(10; 1.) $\delta = 20-10 = 10$ | Normal(20; 1) | 10 | 0,000 | 0,000 | 0,000 |
| | | 20 | 0,000 | 0,000 | 0,000 |
| | | 40 | 0,000 | 0,000 | 0,000 |
| | | 100 | 0,000 | 0,000 | 0,000 |
| Exponencial(1) $\delta = 1,1-1 = 0,1$ | Exponencial(1,1) | 10 | 0,744 | 0,758 | -0,014 |
| | | 20 | 0,253 | 0,261 | -0,008 |
| | | 40 | 0,504 | 0,494 | 0,010 |
| | | 100 | 0,059 | 0,063 | -0,004 |
| Exponencial(1) $\delta = 20-1 = 19$ | Exponencial(20) | 10 | 0,000 | 0,001 | -0,001 |
| | | 20 | 0,000 | 0,000 | 0,000 |
| | | 40 | 0,000 | 0,000 | 0,000 |
| | | 100 | 0,000 | 0,000 | 0,000 |
| Poisson(1) $\delta = 1,1-1 = 0,1$ | Poisson(1,1) | 10 | 0,539 | 0,392 | 0,147 |
| | | 20 | 0,643 | 0,553 | 0,090 |
| | | 40 | 0,204 | 0,166 | 0,038 |
| | | 100 | 0,922 | 0,858 | 0,064 |
| Poisson(1) $\delta = 20-1 = 19$ | Poisson(20) | 10 | 0,000 | 0,000 | 0,000 |
| | | 20 | 0,000 | 0,000 | 0,000 |
| | | 40 | 0,000 | 0,000 | 0,000 |
| | | 100 | 0,000 | 0,000 | 0,000 |

Obs: O δ é a diferença entre as esperanças das duas populações;
 Por simplicidade, os valores- p são obtidos de uma *única* amostra da respectiva população;
 A diferença da última coluna é a diferença entre os valores- p ., obtidos pelos dois métodos.

Finalmente, a Figura 5.4 mostra o comportamento da diferença média entre as estimativas dos valores- p , em função do tamanho da amostra, e em função da distribuição das populações. Conforme esperado, a diferença entre os valores- p estimados tende a diminuir, com o aumento do tamanho da amostra, e é maior para amostras da distribuição Poisson (que é discreta e assimétrica). Além disso, observamos que as estimativas pelo teste de permutação tendem a ser menores que os valores fornecidos pelo método clássico (o que resulta em diferenças positivas), nas amostras da distribuição normal. Ao contrário, no caso de amostras da distribuição exponencial, os

valores- p fornecidos pelo teste de permutação têm tendência a serem superiores aos valores fornecidos pelo método clássico.

5.2 Dados Reais

Para melhor ilustrar a utilidade dos testes de permutação para comparação entre médias de duas populações, usaremos dados reais. Utilizaremos os dados reais tais como apresentados em Santos (1998). Através de 18 amostras retiradas de camadas superficiais dos solos de dois locais diferentes (*Central Soil Salinity Research Institute*), foram coletados dados sobre seus devidos valores de pH. O processo utilizado foi a retirada de 9 amostras de cada uma das duas regiões pesquisadas. O objetivo é tentar provar se realmente as amostras podem ser classificadas em dois grupos distintos ou não.

Localidade A

8,53 8,52 8,01 7,99 7,93 7,89 7,85 7,82 7,80

Localidade B

7,85 7,73 7,58 7,40 7,35 7,30 7,27 7,27 7,23

Realizamos no Minitab® uma análise preliminar dos dados. Os resultados estão disponíveis na Tabela 5.2 e Figura 5.4.

Tabela 5.2: Estatística descritiva dos dados reais

| | N | Mean | StDev | SE Mean |
|-------|---|-------|-------|---------|
| loc A | 9 | 8,038 | 0,285 | 0,095 |
| loc B | 9 | 7,442 | 0,224 | 0,075 |

| | | | | |
|---|--|--|--|--|
| Difference = mu (loc A) - mu (loc B) | | | | |
| Estimate for difference: 0,596 | | | | |
| 95% CI for difference: (0,338; 0,853) | | | | |
| T-Test of difference = 0 (vs not =): T-Value = 4,92 P-Value = 0,000 DF = 15 | | | | |

Pela Tabela 5.2, observamos uma estimativa para a diferença entre as médias das duas localidades de 0,596, com um valor- p muito próximo de zero. Isso indica uma evidência bastante forte de que as duas localidades possuem, de fato, médias diferentes.

Pela Figura 5.4, entretanto, notamos uma distribuição bastante assimétrica dos dados. Esse fato, reforçado pelo tamanho reduzido das amostras, pode sugerir um baixo desempenho do método clássico e ser, conseqüente, uma indicação de uso do método de permutação.

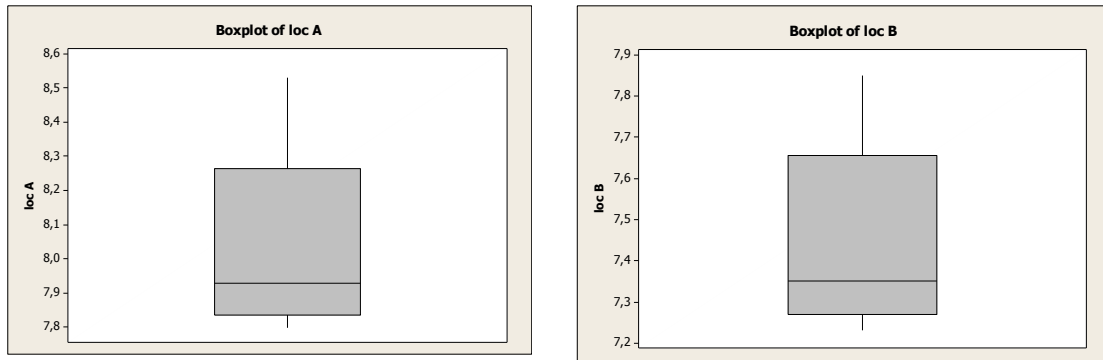


Figura 5.5: Boxplot dos dados

Utilizando, portanto, a macro que foi desenvolvida, obtivemos os resultados apresentados a seguir. Na Tabela 5.3, temos a saída, que igualmente indica uma evidência forte contra a hipótese nula de que as duas amostras vêm de populações com mesma média.

Tabela 5.3: Chamada à macro e saída do teste de permutação

```

MTB > %permuta c1 c2;
SUBC> hist.

Executing from file: permuta.MAC

Data Display
      Estatística original
k6      -0,595556

Data Display
      Valor-p
k7      0
  
```

De fato, a estatística original (obtida das amostras reais) se situa em uma posição bastante extrema no histograma de diferenças entre as médias das reamostras, conforme apresentado na Figura 5.6. Neste caso, temos uma concordância entre as conclusões

obtidas pelos dois métodos. O analista pode estar mais seguro da sua conclusão, pois o teste de permutação não assume nenhuma hipótese em relação aos dados examinados.

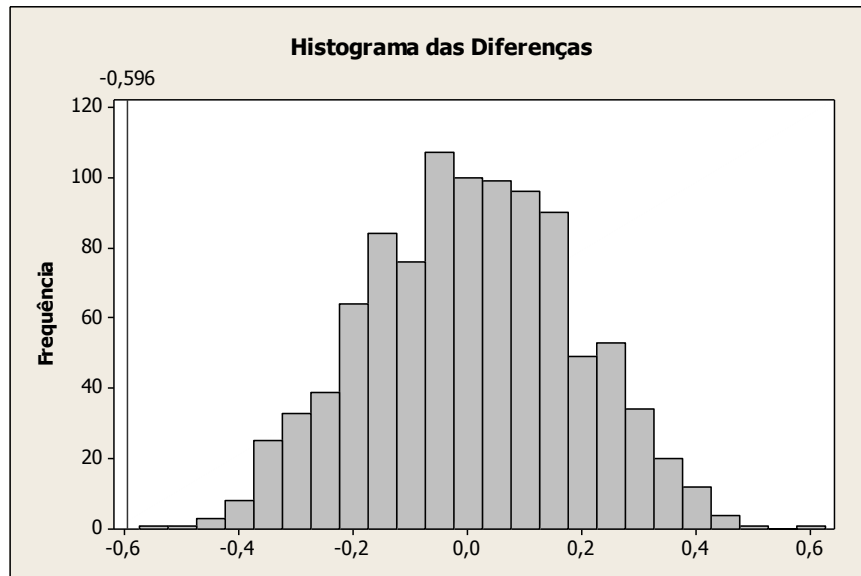


Figura 5.6: Histograma das diferenças das médias das reamostras

6. Conclusões e Observações Finais

Descrevemos neste trabalho o desenvolvimento e teste de uma macro em Minitab® para comparação entre médias de duas populações arbitrárias, por meio de uma técnica computacionalmente intensiva conhecida como teste de permutação. A grande vantagem desta técnica é que nenhuma suposição precisa ser feita quanto a distribuição dos dados sob teste.

7. Bibliografia

- Duczmal, L. H.; Bessegato, L. F.; Santos, M. A. C. & Ferreira, S. J. (2003). *Introdução às Pilhas e Filas e Teste de Permutação*. RTE-03/2003, Departamento de Estatística – UFMG, Belo Horizonte, MG.
- Edgington, E. S. (1995). *Randomization Tests*. Marcel Dekker, Inc., 3ª ed., New York, NY.
- Montgomery, D. C. & Runger, G. C. (2003). *Estatística Aplicada e Probabilidade para Engenheiros*. 2ª ed., LTC Livros Técnicos e Científicos Editora S.A., Rio de Janeiro, RJ.
- Santos, M. A. C. (1998). *Noções de Estatística Computacional*. RTE-03/1998, Departamento de Estatística – UFMG, Belo Horizonte, MG.

- Silva, I.; Assunção, R. & Costa, M. (2009). Power of the Sequential Monte Carlo Test. *Sequential Analysis* 28, 163-174.
- Soares, J. F. & Siqueira, A. L. (1999). *Introdução à Estatística Médica*. Departamento de Estatística – UFMG, Belo Horizonte, MG.
- Triola, M. F. (2005). *Introdução à Estatística*. 9ª ed., LTC Livros Técnicos e Científicos Editora S.A., Rio de Janeiro, RJ.

Apêndice I – Macro Permuta

Apresentamos abaixo, a macro, desenvolvida para o Minitab®

```
MACRO
# template
permuta c1 c2;
hist.
# Autor:
# Miled Ibrahim Zeghezghi & F. R. B. Cruz
# Departamento de Estatística - UFMG
# E-mail: fcruz@est.ufmg.br - (c) 2009
# Funcao:
# ...
# Utilizacao:
# MTB> %DRIVE:\CAMINHO\permuta c1 c2
# Declaração das variáveis
MCOLUMN c1 c2 c3 c4 c5 c6 c7
MCONSTANT k1 k2 k3 k4 k5 k6 k7 k8 k9 k10
k11 k12
# Corpo:
BRIEF 0
# Calcular tamanho da amostra 1
Count c1 k1.
# Calcular tamanho da amostra 2
Count c2 k2.
# Calcular tamanho total das amostras 1 e 2
LET k3 = k1+k2
# Calcular as médias das amostras 1 e 2
Mean C1 k4.
Mean C2 k5.
# Calcular a diferença entre as médias das
amostras originais
LET k6 = k5 - k4
# Juntar amostras 1 e 2 em amostra 3
Stack C1 C2 C3.
# inicializar o valor-p
LET K7 = 0
# inicializar o número de permutações
LET K8 = 1000
```

```
# Fazer permutações
DO k9 = 1:k8
# permutar amostra 3 e armazenar em
amostra 4
Sample k3 C3 c4.
# colocar k1 primeiros elementos da
amostra 4 em 5
DO k10 = 1:k1
LET c5[k10] = c4[k10]
ENDDO
# colocar k2 elementos restantes da
amostra 4 em 6
DO k10 = 1:k2
LET c6[k10] = c4[k1+k10]
ENDDO
# calcular as médias das amostras 5 e 6
Mean C5 k11.
Mean C6 k12.
# acumular a diferença entre as médias
LET c7[k9] = k12 - k11
# contribuir para o valor-p
IF ABSO(C7[k9]) >= ABSO(K6)
LET k7 = k7 + 1
ENDIF
ENDDO
# calcular o valor-p
LET k7 = k7/k8
# Escrever resultados
BRIEF 1
Print "Estatística original" k6
print "Valor-p" k7
IF hist <> 0
Histogram C7;
Bar.
ENDIF
ENDMACRO
```