

Análise de resultados (prática)

Jorge Alberto Wagner Filho

INF01043 Interação Humano-Computador

Instituto de Informática
Universidade Federal do Rio Grande do Sul

July 3, 2017

Sumário

- 1 Objetivos
- 2 Requisitos
- 3 Linguagem R
- 4 Testes estatísticos
- 5 Testes estatísticos em R
- 6 Dados do experimento
- 7 Exercícios

Objetivos

- Pôr em prática o que foi visto sobre **análise de resultados** de testes de interação baseada no método científico.
- Aprofundar experiência em testes com usuários (essencial em IHC).
- Aprender um pouco sobre a linguagem **R** para computação estatística.
- Ajudar nas análises dos trabalhos de vocês.
- Explorar os dados gerados por vocês mesmos no experimento com o Victor.

- ## 7 Exercícios

Requisitos

- ① Boot no Ubuntu
- ② Baixar em **<http://inf.ufrgs.br/~jawfilho/ihc/aula28.zip>**:
 - slides.pdf
 - exemplos/
 - typing.csv
 - pointing.csv
 - experimento/
 - demographics.csv
 - dv.csv
 - merged.csv
 - data.xlsx

- ### 3 Linguagem R

Sobre o R

- O que é?
 - Linguagem *open source* para computação estatística.
 - Largamente usada para análise e mineração de dados.
 - Tem se tornado cada vez mais popular nos últimos anos.

Sobre o R

- O que é?
 - Linguagem *open source* para computação estatística.
 - Largamente usada para análise e mineração de dados.
 - Tem se tornado cada vez mais popular nos últimos anos.
- Como usar?
 - **Linha de comando:** R, Rscript
 - IDE: RStudio
 - Web: R-Fiddle

Sobre o R

- O que é?
 - Linguagem *open source* para computação estatística.
 - Largamente usada para análise e mineração de dados.
 - Tem se tornado cada vez mais popular nos últimos anos.
- Como usar?
 - **Linha de comando:** R, Rscript
 - IDE: RStudio
 - Web: R-Fiddle
- Como aprender mais?
 - Apostila *R for Beginners*
 - Curso online *Try R* da Code School
 - Site *Quick-R - accessing the power of R*

Operações básicas em R

Carregar dados de um arquivo para um *data frame*

```
> mydata <- read.csv("data.csv")  
> str(mydata) # verificar formato  
> mydata$User <- as.factor(mydata$User) # categórico
```

Acessar linhas, colunas e posições

```
> mydata[1,] # primeira linha  
> mydata[,1] # primeira coluna  
> mydata[1,1] # posição  
> mydata$Participante # coluna nomeada  
> mydata[mydata$Mapa == "1", "Tempo"] # condicional
```

Operações básicas em R

Calcular médias, desvios, correlação

```
> mean(x)
> sd(x)
> summary(x)
> cor(x,y)
```

Instalar pacotes

```
> install.packages("ez")
> library(ez)
```

Gerando gráficos em R

Boxplot

```
> plot(Tempo ~ Técnica, mydata)
```

Scatterplot

```
> plot(mydata$Tempo, mydata$Técnica)
```

Histograma

```
> hist(demo$Idade)
```

Pizza

```
> mytable <- table(demo$Idade)
> lbls <- paste(names(mytable), "\n", mytable, sep="")
> pie(mytable, labels=lbls)
```

- 1 Objetivos
- 2 Requisitos
- 3 Linguagem R
- 4 Testes estatísticos**
- 5 Testes estatísticos em R
- 6 Dados do experimento
- 7 Exercícios

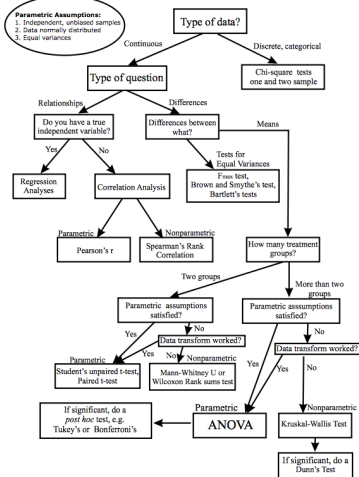
Recapitulando

Avaliação somativa

- Mede resultado final
- Compara diferentes técnicas
- Muitos usuários, protocolo estrito
- Variáveis independentes e dependentes
- Resultados quantitativos
- **Significância estatística**

Qual teste estatístico usar?¹

Flow Chart for Selecting Commonly Used Statistical Tests



¹Fonte: Robert Gerwien

Testes muito usados em IHC

Teste t de Student

Comparar 1 variável dependente obtida de 2 tratamentos de 1 variável independente

ANOVA - *Analysis of Variance*

Comparar 1 variável dependente obtida de n tratamentos com m variáveis independentes

Atenção: ambos são testes paramétricos → assumem que amostras são de distribuições normais com variâncias iguais.²

²Teste t de Welch não assume variâncias iguais.

- 

Testes estatísticos básicos em R ³

Teste de normalidade: e.g. Shapiro-Wilk

H0: distribuição normal (confirmada se $p > 0.05$)

```
> shapiro.test(data)
```

Teste t (de Welch)

H0: médias são iguais (rejeitada se $p < 0.05 \rightarrow$ diferença encontrada)

```
> t.test(x,y) # (between-subjects)
```

```
> t.test(x,y, paired=T) # (within-subjects)
```

³Muito útil para testes em IHC:
páginas de Koji Yatani e Anthony Tang.

Exemplo⁴

pointing.csv

Hipótese: mouse é mais rápido do que stylus para apontamento.

H0: mouse = stylus

```
> pointing <- read.csv("exemplos/pointing.csv")  
> summary(pointing)  
  
> shapiro.test(pointing$mousedata)           # normalidade  
> shapiro.test(pointing$stylusdata)  
  
> t.test(pointing$mousedata, pointing$stylusdata,  
paired=T)                                     # teste t pareado/within-subjects
```

⁴Fonte: hctang.org

Testes estatísticos básicos em R

One-way ANOVA

H0: médias são iguais (rejeitada se $p < 0.05 \rightarrow$ diferença encontrada)

```
> aov <- aov(Value ~ Group, data)
> summary(aov)
```

Teste par a par *post-hoc*: e.g. TukeyHSD

```
> TukeyHSD(aov)
```

Exemplo⁵

typing.csv

Hipótese: Digitar com a predição do Android é mais rápido do que com a do iOS ou do que sem predição.

H0: Todas as formas são igualmente rápidas.

```
> typing <- read.csv("exemplos/typing.csv")
> boxplot(perf ~ type, typing)

> aov = aov(perf ~ type, data=typing)    # one-way anova
> summary(aov)

> TukeyHSD(aov)
```

⁵Fonte: hcritang.org

Testes estatísticos básicos em R

One-way ANOVA com medidas repetidas

```
> aov <- aov(Value ~ factor(Group) +  
Error(factor(Participant)/factor(Group)), data)
```

Two-way ANOVA com medidas repetidas

```
> library(ez)  
> options(contrasts=c("contr.sum", "contr.poly"))  
> ezANOVA(data=data, dv=.(Time), wid=.(Participant),  
within=.(Technique, Device), type=3)
```

- ## 6 Dados do experimento

Variáveis do experimento

Independentes:

- Condição
 - ① TactileMap (alocêntrico)
 - ② ImmersiveMap (egocêntrico)
 - ③ TactileMap + ImmersiveMap (híbrido)
- Mapa
 - ① Oblíquo
 - ② Reto
 - ③ Agudo

Dependentes: Tempo, Workload, Easiness.

- 1 Objetivos
- 2 Requisitos
- 3 Linguagem R
- 4 Testes estatísticos
- 5 Testes estatísticos em R
- 6 Dados do experimento
- 7 Exercícios

Exercício 1

1) Como você caracterizaria sua população se precisasse descrevê-la?

Exercício 2

2) Analise os dados e elabore algumas hipóteses que possam ser provadas ou refutadas a partir da análise dos dados qualitativos fornecidos pelos usuários. Por exemplo: *“Homens com menos de 30 anos e algum problema de visão acharam que foi difícil chegar ao destino.”*

Em seguida, tente provar a sua hipótese usando os dados disponíveis. Por exemplo: *“Esta afirmação é embasada pelas respostas à questão n, onde vê-se que 75% dos homens com problema de visão disseram que foi difícil chegar ao destino, dando notas iguais ou inferiores a 3 para este quesito”*. Gráficos ilustrativos são bem vindos. Tente ser o mais genérico possível em suas conclusões. Caso não seja possível, tente identificar correlações entre as variáveis.

Exercício 3

3) Agora, analise os dados quantitativos disponíveis. Você consegue perceber alguma diferença ou semelhança de comportamento entre usuários de condições e mapas diferentes? Justifique sua resposta usando gráficos ou análise estatística.