

Estudo de Caso 01: estudantes da disciplina *Design and Analysis of Experiments* são bons estimadores para quantidade e valor de moedas colocadas em um copo?

Team 04

abril de 2017

Coordenador: Gustavo Vieira

Relator: Danny Tonidandel

Verificador: Alessandro Dias

Monitor: Bernardo Marques

1- O experimento

1.1 Introdução

Até que ponto a opinião de pessoas comuns, reunidas em grandes quantidades, podem revelar “*verdades*” acerca da natureza de determinado objeto ou fenômeno? Segundo Steiner [1], que realizou uma série de testes baseados no *best seller* *The Wisdom of Crowds* [2], o mais famoso experimento desta natureza foi realizado pelo Cientista Vitoriano *Francis Galton*, em uma carta enviada à revista *Nature* [3], na qual analisa uma competição realizada em *Plymouth* (Inglaterra), em que diversas pessoas deveriam estimar a massa de um boi. Obviamente ninguém acertou exatamente o valor, mas a média das tentativas das quase 800 pessoas que participaram do concurso refletiu, com bastante proximidade, o real valor da medida procurada. E o que Steiner realizou foi testar a ideia utilizando-se de uma garrafa cheia de moedas, convidando pessoas que acessavam a internet a fazerem o mesmo, a partir de uma foto que mostrava a garrafa com as moedas. Mas seria isto verdade?

1.2. Descrição do Problema

Da mesma forma podemos conjecturar que o experimento proposto pelo professor da disciplina *Design and Analysis of Experiments* foi inspirado nos mesmos experimentos. Com a diferença de que o material utilizado foram dois recipientes *A* e *B*, cheios de moedas, conforme descrito na referência [4]. O vigente estudo busca, portanto, investigar se as opiniões de 29 estudantes, isto é, o quanto a média das opiniões dos estudantes pode refletir o número e o valor real das moedas depositadas nos recipientes *A* e *B*?

1.3. Design Experimental

Como a média real não foi dada a conhecer pelo proponente do estudo, o time decidiu realizar uma montagem experimental semelhante (replicação do experimento), utilizando um recipiente de mesma natureza (copo plástico de 200ml) para uma estimativa inicial do número de moedas no recipiente *A*, sabendo que era composto por moedas de natureza diferente (25 e 50 centavos e 1 real) e, no recipiente *B*, moedas de mesma natureza (5 centavos), utilizando contagem manual das moedas. O resultado seria utilizado como estimativa inicial para as médias:

Recipiente A : 182 moedas; *Recipiente B* : 9 reais e 10 centavos.

Assim, formula-se a hipótese de que a média das estimativas dos estudantes é igual ou não ao “valor real”:

$$\begin{cases} H_0 : \mu = 130, \\ H_0 : \mu \neq 130. \end{cases}$$

e

$$\begin{cases} H_0 : \mu = 9.10, \\ H_0 : \mu \neq 9.10. \end{cases}$$

Por razões históricas e por conveniência, o time julgou ser suficiente um nível de significância para o experimento de 5%, i.e., $\alpha = 0.05$, que implica em um grau de confiança $1 - \alpha$ de 95%.

2 - Análise Estatística

De antemão já é possível notar que é possível realizar uma análise estatística descritiva, porém esta análise só fornecerá informações sobre as opiniões dos estudantes, não da quantidade ou do valor real das moedas nos recipientes.

2.1 Validação das Premissas

Como não se tem informações sobre a variância da população, o time escolheu adotar o teste de *t student*, assumindo as premissas: 1) As estimativas dos estudantes se distribuem em torno do valor real. 2) As observações são independentes. 3) A distribuição das médias é normal.

Normalidade e independência são facilmente verificáveis por testes, porém, como a premissa de que a média dos “chutes” dos estudantes se aproxima do valor real não é facilmente testável *a priori*, sendo assim o ponto fraco da análise

A premissa de normalidade foi realizada por meio do teste de Shapiro-Wilk

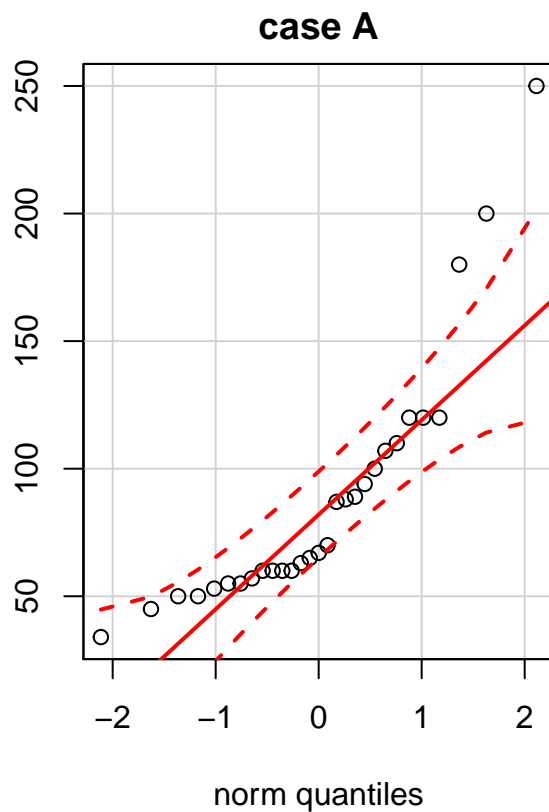
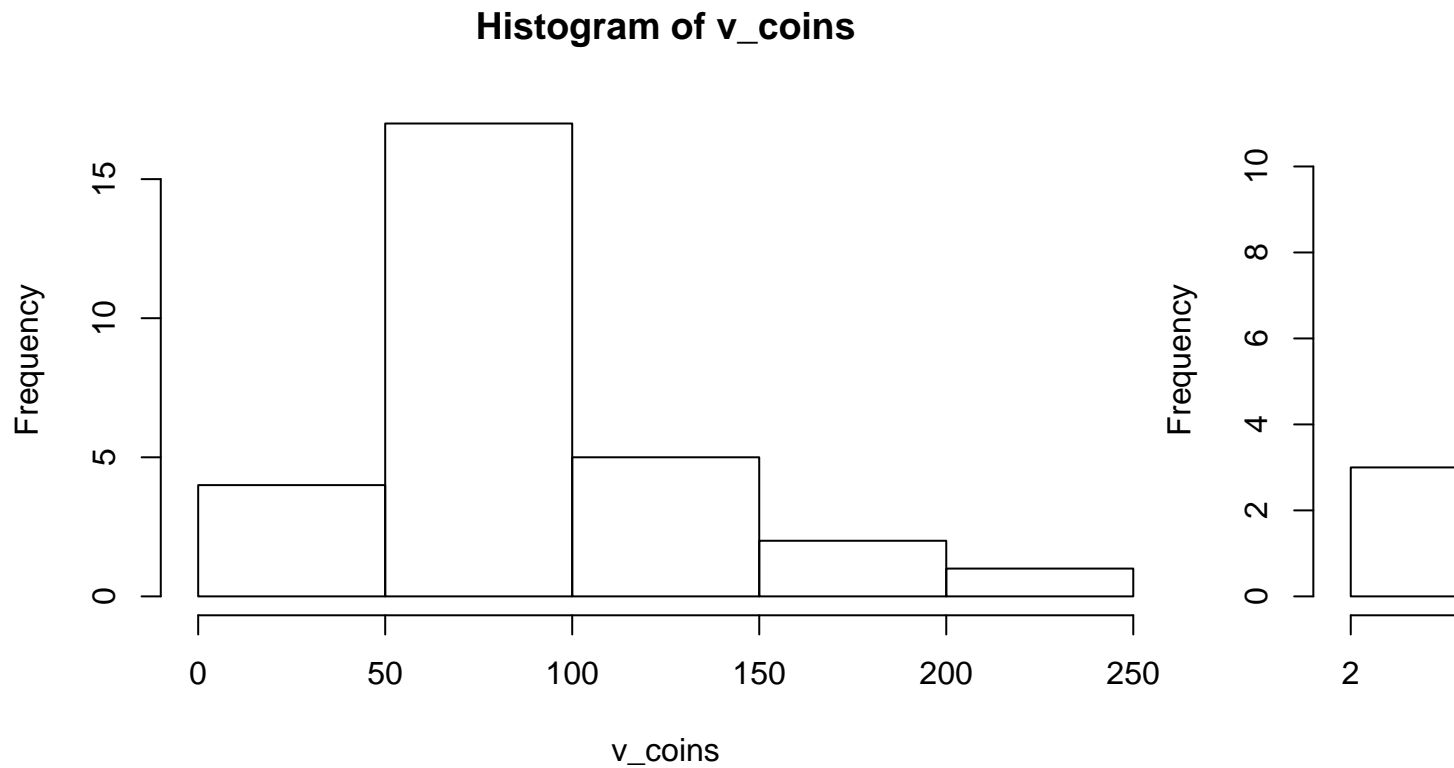
Para o caso *A*, do número de moedas:

```
##
##  Shapiro-Wilk normality test
##
## data:  v_coins
## W = 0.79755, p-value = 7.551e-05
```

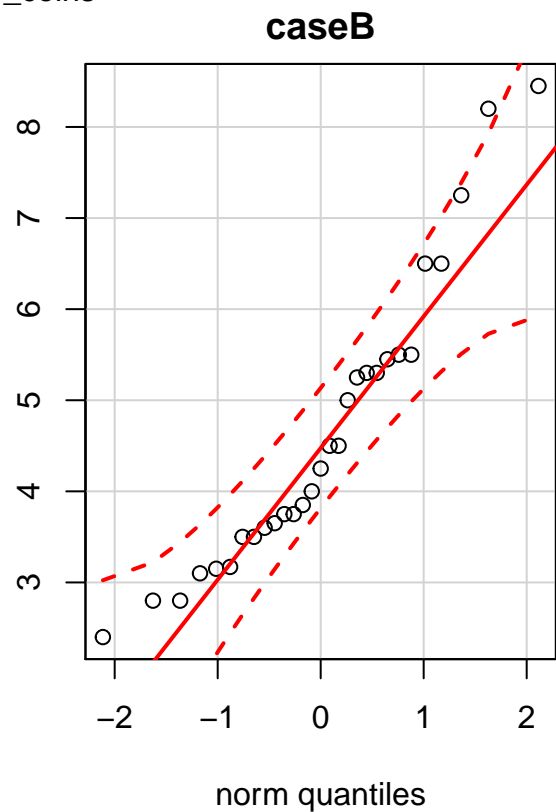
E para o caso *B*, para o valor:

```
##
##  Shapiro-Wilk normality test
##
## data:  v_value
## W = 0.92156, p-value = 0.0334
```

Para os dois casos, o valor de W_α



v_coins



###

2. Independência

B. Análise Descritiva

C. Teste de Hipóteses

```
##
## One Sample t-test
##
## data:  v_coins
## t = -4.5118, df = 28, p-value = 0.0001052
## alternative hypothesis: true mean is not equal to 130
## 95 percent confidence interval:
##  69.78405 107.38836
## sample estimates:
## mean of x
##  88.58621

##
## One Sample t-test
##
## data:  v_value
## t = -15.153, df = 28, p-value = 5.061e-15
## alternative hypothesis: true mean is not equal to 9.1
## 95 percent confidence interval:
##  4.033574 5.240219
## sample estimates:
## mean of x
##  4.636897
```

Potência dos Tests

A potência é 0.374285

Considerações finais

Surowiecki, em seu estudo, lembra que a diferença não só contribui trazendo novas perspectivas para o ambiente, mas também ajuda os integrantes a expressarem mais livremente suas opiniões - sejam elas divergentes ou não [2, p. 38-39]. Isto revela o problema da coleta aberta no segundo caso, pois, não importa a magnitude do erro: mesmo que a “intuição” sugira o contrário, as pessoas dificilmente dariam respostas muito discrepantes da maioria. Gregory Berns, em seu *Iconoclast: A Neuroscientist Reveals How to Think Differently*[7] questiona inclusive a influência do grupo sobre a percepção das pessoas. Embora os estudantes garantirem terem dados a melhor resposta de acordo com suas observações, eles provavelmente questionavam suas convicções. Pode ser que alguns duvidassem daquilo que estavam vendo. Aparentemente as percepções permanecem intactas, mas a “fé” das pessoas nos seus sentidos, esta sim, parece ser moldada pela influência externa, alterando as decisões tomadas. E, no final, como o próprio prof. Campelo costuma afirmar em sala de aula, o que importa mesmo são as decisões. Vale ressaltar que o grupo experienciou um certo “alívio” ao saber que a experiência era por isso, de certo modo, uma pequena farsa.

V. Atividades Desempenhadas

Referências

- [1] Steiner, E. B. *Turns Out the Internet Is Bad at Guessing How Many Coins Are in a Jar*. Wired Magazine: USA, 2017. Disponível em <https://www.wired.com/2015/01/coin-jar-crowd-wisdom-experiment-results/>
 - [2] Surowiecki, J. *The Wisdom of Crowds*. Anchor Books: New York, 2004.
 - [3] Galton, F. *Vox Populi*. Nature: England, mar. 1907.
 - [4] Campelo, F. *Estudo de caso 01*. Arquivo da disciplina Design and Analysis of Experiments. Disponível em <https://goo.gl/b3IeAn>.
 - [5] Ramirez, J.G. *Statistical Intervals: Confidence, Prediction, Enclosure*. Disponível em <http://goo.gl/NJz7ot>
 - [6] D.C. Montgomery, *Design and Analysis of Experiments*, 5th ed., John Wiley & Sons, 2001.
 - [7] Berns, G. *Iconoclast: A Neuroscientist Reveals How to Think Differently*. USA: Harvard Business press, 2008.
- Peng, Roger D., *R Programming for Data Science*, Lulu, Inc.
- Peng, Roger D., *Report Writing for Data Science in R*, Lulu, Inc.
- Felipe Campelo, *Lecture Notes on Design and Analysis of Experiments*, 2015.