UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE ESTATÍSTICA E MATEMÁTICA APLICADA
CURSO DE GRADUAÇÃO EM ESTATÍSTICA

ANTÔNIO ARTHUR SILVA DE LIMA
FRANCISCO LUAN RODRIGUES DE SOUSA
JULYET ALVES DO NASCIMENTO SILVA

PROJETO II - ESTATÍSTICA NÃO PARAMÉTRICA

FORTALEZA
2023

ANTÔNIO ARTHUR SILVA DE LIMA
FRANCISCO LUAN RODRIGUES DE SOUSA
JULYET ALVES DO NASCIMENTO SILVA

PROJETO II - ESTATÍSTICA NÃO PARAMÉTRICA

Relatório apresentado ao curso de Bacharelado em Estatística do Centro de Ciências da Universidade Federal do Ceará, como parte dos requisitos para a aprovação na disciplina de Estatística Não Paramétrica no semestre de 2023.2.
Prof.: Manoel Ferreira dos Santos Neto.

FORTALEZA
2023

Sumário

1 Introdução

A Estatística Não Paramétrica tem como foco a realização de testes de hipóteses e análises inferenciais que não pressupõem que os dados sigam uma distribuição específica ou que seus parâmetros sejam conhecidos. Desta maneira, é um campo amplamente utilizado quando não se conhece a verdadeira lei dos dados, mas quando é necessário realizar um estudo sobre os mesmos, ou quando tais dados violam alguma suposição inicial para a realização de testes paramétricos, como a de normalidade.

O presente trabalho busca realizar um estudo — aplicando métodos não paramétricos — a um conjunto de dados contendo a fração de vitórias de 21 jogadores em um jogo de fantasia. O cerne da pesquisa encontra-se em determinar se o resultado dos jogadores, que jogam contra a máquina em 112 tentativas, é obtido por pura sorte, sendo o caso esperado se a fração de vitórias for de 50%.

Com as análises e resultados obtidos, a Estatística pode proporcionar informações mais detalhadas sobre o comportamento dos usuários no jogo, atuando como guia para nortear ações da equipe de desenvolvedores e empresa responsável pelo *game*, visando melhorar a experiência do usuário, bem como aperfeiçoar o aproveitamento de recursos, direcionamento de público alvo e nivelamento da dificuldade do jogo.

2 Metodologia

Para a realização da pesquisa, foi disponibilizada através do **SIGAA**, uma base de dados em formato *.csv* contendo a numeração dos 21 jogadores e a proporção de vitórias de cada um. Como ferramenta de ação principal, foi utilizada a linguagem de programação **R**, juntamente com algumas de suas bibliotecas mais úteis de tratamento, análise e limpeza de dados, por exemplo, a biblioteca *tidyverse*. O código completo pode ser encontrado no Apêndice, e também de forma distribuída na seção de Análises, conforme cada teste.

Os testes verificados e/ou performados, todos feitos com a linguagem, foram:

1. **Paramétricos**
   (a) Teste t.
2. **Não paramétricos**
   (a) Teste de *Shapiro-Wilk*;
   (b) Teste dos Sinais;
   (c) Teste das Somas Ordenadas de *Wilcoxon*;
   (d) Teste Binomial.

2.1 Métodos paramétricos

O teste t é um teste paramétrico para média populacional, e é realizado quando as suposições de normalidade e aleatoriedade são satisfeitas, além de assumir que a variância populacional também seja desconhecida. A distribuição utilizada neste teste, como o próprio nome já revela, é a *t de Student*, descrita por William Gosset em 1908.

A estatística de teste é dada por

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}},$$

que nos leva a um intervalo de confiança para a média verdadeira:

$$IC[\mu; (1-\alpha)100\%] = \bar{x} \pm t_{(1-\frac{\alpha}{2}, n-1)} \frac{S}{\sqrt{n}}.$$

2.2 Métodos não paramétricos

O teste de *Shapiro-Wilk* verifica a normalidade de um conjunto de dados através da interpretação do valor-p, com o uso da estatística $W = \dfrac{(\sum_{i=1}^{n} a_{i,n} X_{(i)})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$.

O teste dos sinais é um método não paramétrico que usa a mediana de um conjunto de dados ao invés da média. Neste procedimento, calcula-se a diferença entre o valor observado e o hipotético, atribuindo sinais à observação conforme tal diferença. O tamanho amostral é dado pela soma de sinais positivos e negativos. Em seguida, calcula-se a estatística baseado no tamanho amostral, para posterior comparação com o valor tabelado (que também irá depender do tamanho da amostra). As suposições para este teste são de não normalidade, aleatoriedade, continuidade da variável em questão, e assimetria.

Já no teste Binomial, queremos encontrar a probabilidade de sucesso de um evento em n ensaios de Bernoulli, supondo que temos uma amostra aleatória de dados normais. Assim, utilizamos o seguinte intervalo de confiança para essa proporção: $IC[p; (1-\alpha)100\%] = \hat{p} \pm z_{(1-\frac{\alpha}{2})} \dfrac{\sigma}{\sqrt{n}}$. Para a sua realização, e também de forma alternativa para confirmar os resultados realizados com os dados da proporção, uma nova coluna, chamada "Sucesso", foi adicionada.

Por fim, o teste das somas ordenadas de *Wilcoxon*, ou simplesmente teste de Wilcoxon, também é um teste que leva em consideração a mediana dos dados ao invés da média. Diferentemente do teste dos sinais, há aqui a suposição de simetria, além de também assumir que a variável de interesse é contínua, com a presença de aleatoriedade. Os primeiros passos são os mesmos dos do teste de sinais, porém, computa-se também o módulo da diferença entre os valores observados e hipotéticos, seguido da operação de ranquear tais valores. Sendo assim, encontra-se a estatística de teste por $W^+ = \sum R_d^+$, que é a soma dos ranques para as diferenças positivas. Também é possível que haja empates entre os valores absolutos das diferenças. Neste caso, é atribuído ao ranque a média dos ranques do grupo de observações empatadas, e por fim, somam-se todos os ranques. É possível ainda, quando em grades amostras, aproximar W^+ para

$$W^{+*} = \frac{W^+ - \dfrac{n(n+1)}{4}}{\sqrt{\dfrac{n(n+1)(2n+1)}{24}}}.$$

3 Análises

Queremos testar a hipótese de que os resultados dos jogadores foram obtidos por "pura sorte" — o que ocorre se houver 50% de vitórias/derrotas — e isso pode ser encontrado de duas formas:

1. Testar, de maneira geral, se a proporção média de vitórias é igual a 0,5 (56);
2. Testar, individualmente, se a proporção de vitórias é 0,5 (56).

Desta forma, construímos as seguintes hipóteses para realizar os testes:

$$H_0 : \mu = 0,5 \quad vs \quad H_1 : \mu \neq 0,5.$$

E para realizar os testes que utilizam a mediana:

$$H_0 : Md = 0,5 \quad vs \quad H_1 : Md \neq 0,5.$$

Na segunda situação, podemos realizar o teste binomial, pois queremos saber se a proporção do número de vitórias de cada jogador corresponde a 50% ou 0,5, ou seja, conhecemos a população, e iremos realizar um teste para uma proporção. Já na primeira situação, devemos verificar a distribuição dos nossos dados, e, a partir disso, decidir quais testes podemos realizar.

3.1 Testes iniciais

Antes da realização dos testes propriamente ditos, é fundamental saber que forma têm nossos dados, e se eles seguem normalidade. A tabela a seguir mostra o conjunto disponibilizado:

Jogador	P	Sucesso
1	0.31622024	35
2	0.37053571	41
3	0.40848214	46
4	0.50595238	57
5	0.51860119	58
6	0.36755952	41
7	0.26339286	29
8	0.45386905	51
9	0.29166667	33
10	0.31473214	35
11	0.21577381	24
12	0.42708333	48
13	0.53943452	60
14	0.50372024	56
15	0.43377976	49
16	0.47098214	53
17	0.35119048	39
18	0.47470238	53
19	0.46428571	52
20	0.62127976	70
21	0.36681548	41

Podemos criar rapidamente medidas sumário para as proporções ou para o número absoluto aproximado de vitórias. Fazendo isso no R, e utilizando medidas padrão, obtemos para a proporção:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Kurt.	Skew.
0.2158	0.3512	0.4271	0.4133	0.4747	0.6213	-0.7744	-0.0436

Agora, podemos criar histogramas para a proporção e valor absoluto (aproximado) de vitórias dos jogadores:

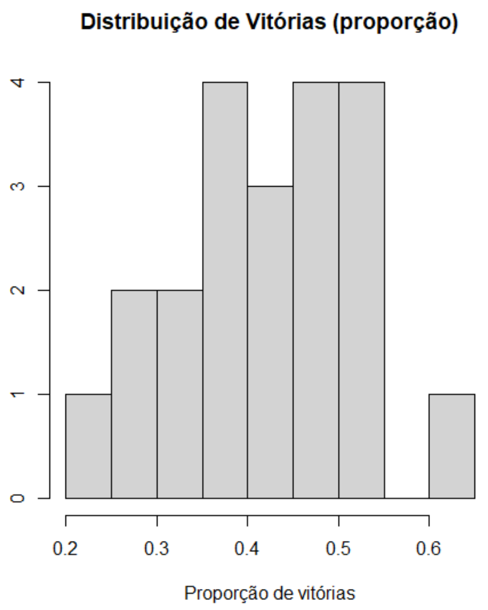

Figura 3.1: Distribuição da proporção de vitórias

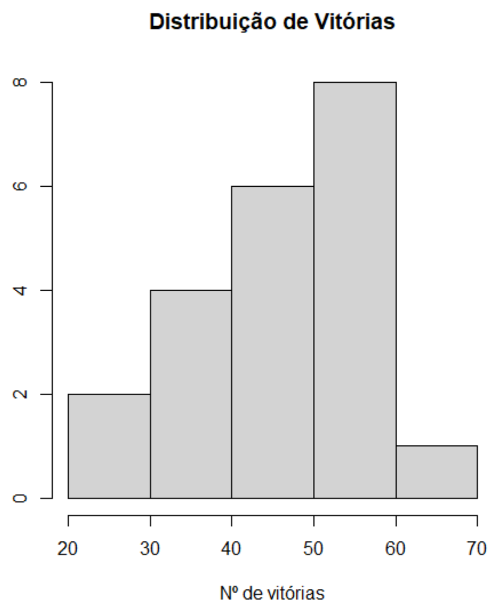

Figura 3.2: Distribuição absoluta de vitórias

Vemos que nossa distribuição é aproximadamente simétrica em torno do 0,4 para a proporção, e em torno de 50 para os valores absolutos.

Iremos agora verificar a suposição de normalidade por meio do teste de *Shapiro-Wilk*:

```
> shapiro.test(df$P)

Shapiro-Wilk normality test

data:  df$P
W = 0.98738, p-value = 0.9912

> shapiro.test(df$Vitorias)

Shapiro-Wilk normality test

data:  df$Sucesso
W = 0.98409, p-value = 0.9716
```

Assim, não podemos rejeitar a suposição de normalidade do número de vitórias dos jogadores.

3.2 Teste dos Sinais

O Teste dos Sinais é um procedimento não paramétrico para testar a hipótese nula de que a mediana de uma distribuição corresponde a certo valor hipotético k.

As suposições para este teste são:

1. Dados **não** são normalmente distribuídos;
2. Ter uma Amostra Aleatória Simples (A.A.S.);
3. Variável de interesse é contínua;
4. Distribuição é assimétrica.

Sendo assim, não será possível utilizar o teste dos sinais, pois há violações de normalidade e continuidade.

3.3 Teste t

No caso geral, temos uma distribuição normal, com média $\hat{p}$ de vitórias, na qual desconhecemos o desvio-padrão. Dessa forma, podemos utilizar um teste t, onde queremos testar se a média da proporção de vitórias dos 21 jogadores corresponde a 0.5.

Através do R, realizamos o teste bilateral e unilateral, com 95% de confiança:

```
> # caso bilateral
> t.test(df$P, mu = 0.5)

One Sample t-test

data:  df$P
t = -3.9289, df = 20, p-value = 0.0008307
alternative hypothesis: true mean is not equal to 0.5
95 percent confidence interval:
 0.3673241 0.4593483
sample estimates:
mean of x
0.4133362

> # caso unilateral
> t.test(df$P, mu = 0.5, alternative = "less")

One Sample t-test

data:  df$P
t = -3.9289, df = 20, p-value = 0.0004154
alternative hypothesis: true mean is less than 0.5
```

```
95 percent confidence interval:
      -Inf 0.4513799
sample estimates:
mean of x
0.4133362
```

Conforme os resultados acima, há significância estatística, pois rejeitamos a hipótese nula de que a média das proporções de vitórias seja igual 0.5.

Assim, baseado neste teste, os resultados dos jogos não são obtidos por pura sorte, mas sim, pela presença ou falta de alguma habilidade dos jogadores.

3.4 Teste de Wilcoxon

Realizamos o teste de Wilcoxon para uma única amostra para avaliar se a fração de vitórias dos usuários é significativamente diferente de 0,5, que representaria um desempenho puramente aleatório. Os testes foram desenvolvidos em linguagem R como apresentado a seguir:

```
    > # vetor de dados
> dados_projeto <- c(0.316220238095238,
+                    0.370535714285714,
+                    0.408482142857143,
+                    0.505952380952381,
+                    0.51860119047619,
+                    0.367559523809524,
+                    0.263392857142857,
+                    0.453869047619048,
+                    0.291666666666667,
+                    0.314732142857143,
+                    0.21577380952381,
+                    0.427083333333333,
+                    0.539434523809524,
+                    0.503720238095238,
+                    0.433779761904762,
+                    0.470982142857143,
+                    0.351190476190476,
+                    0.474702380952381,
+                    0.464285714285714,
+                    0.621279761904762,
+                    0.366815476190476)
> # tamanho da amostra
> length(dados_projeto)
[1] 21
> # teste de Wilcoxon para uma amostra (diferente)
> resultado <- wilcox.test(dados_projeto, mu = 0.5, alternative = "two.sided")
> # resultados do teste:
> print(resultado)

Wilcoxon signed rank exact test
```

```
data:  dados_projeto
V = 25, p-value = 0.0008516
alternative hypothesis: true location is not equal to 0.5

> # teste de Wilcoxon para uma amostra (menor que)
> resultado3 <- wilcox.test(dados_projeto, mu = 0.5, alternative = "less")
> # resultados do teste
> print(resultado3)

Wilcoxon signed rank exact test

data:  dados_projeto
V = 25, p-value = 0.0004258
alternative hypothesis: true location is less than 0.5
```

Os resultados do teste são os seguintes:

- Estatística de teste (V) = 25
- Valor-p = 0.0004258
- Hipótese alternativa: A verdadeira localização (mediana) é menor do que 0,5.

Com base no valor-p de 0.0004258, há evidências estatísticas para rejeitar a hipótese nula. Isso sugere que a fração de vitórias dos usuários não é igual a 0,5 e é significativamente menor. Portanto, os resultados indicam que o desempenho dos usuários não é aleatório e está inclinado para um desempenho inferior a 0,5. Isso pode sugerir que os usuários têm dificuldade em vencer o jogo em relação ao valor de referência de 0,5.

3.5 Teste Binomial

Para o caso em que a análise é feita individualmente para cada jogador, temos que

$$X = \begin{cases} 1, & \text{se o jogador ganhou} \\ 0, & \text{c.c.} \end{cases}$$

Agora, seja S $= \sum_{i=1}^{n} X_i$. Sabemos que $S \sim Bin(n,p) \approx N(np, npq)$. Dessa forma, podemos fazer um teste para a proporção, usando o teste binomial exato ou aproximado pela normal, com fator de correção. Para o teste, temos como hipóteses $H_0 : p = 0,5$ vs $H_1 : p \neq 0,5$, e 95% de confiança. Iremos realizar o teste no R:

```
tab = data.frame(valor_p = NA, li = NA, ls = NA)
for(i in 1:length(df$Sucesso)){
  b = prop.test(df$Sucesso[i], 112, p=0.5)
  l = c(round(b$p.value,4), round(b$conf.int[1],5), round(b$conf.int[2],5))
  tab = rbind(tab, l)
}
```

```
tab = tab[-1,]; tab
row.names(tab) = NULL; tab
xtable(tab)
```

O código acima gera a seguinte tabela, que contém os respectivos valores-p para cada jogador, bem como os intervalos de confiança para p.

Jogador	Valor-p	Li	Ls
1	0.00	0.23	0.41
2	0.01	0.28	0.46
3	0.07	0.32	0.51
4	0.92	0.41	0.60
5	0.78	0.42	0.61
6	0.01	0.28	0.46
7	0.00	0.18	0.35
8	0.40	0.36	0.55
9	0.00	0.21	0.39
10	0.00	0.23	0.41
11	0.00	0.14	0.30
12	0.16	0.34	0.53
13	0.51	0.44	0.63
14	1.00	0.41	0.59
15	0.22	0.34	0.53
16	0.64	0.38	0.57
17	0.00	0.26	0.44
18	0.64	0.38	0.57
19	0.51	0.37	0.56
20	0.01	0.53	0.71
21	0.01	0.28	0.46

Através do teste, nota-se que pouco mais da metade dos jogadores obtiveram seus resultados por mero acaso, enquanto o restante possui alguma falta ou presença de habilidades que influenciam o resultado do jogo.

4 Conclusão

Com a realização dos testes paramétricos e não paramétricos, tanto para as proporções de vitórias quanto para seus respectivos valores absolutos aproximados, e utilizando análises de forma conjunta e individual, foi possível observar que os testes são significantes, ou seja, podemos rejeitar a hipótese nula de que os resultados dos jogadores são obtidos por pura sorte, de modo que podemos fazer novas suposições sobre fatores de influência, tais quais: nível de dificuldade do jogo, presença ou falta de habilidade dos jogadores, problemas no software do jogo, dentre outros.

Referências

[1] J. V. Deshpande, Uttara Naik-Nimbalkar e Isha Dewan. *Nonparametric statistics: theory and methods*. New Jersey: World Scientific, 2018. ISBN: 9789814663571.

[2] *One sample Wilcoxon signed-rank test*. URL: https://statkat.com/stat-tests/one-sample-wilcoxon-signed-rank-test.php (acesso em 08/11/2023).

[3] Ted Hessing PV Ramana. *1 Sample Wilcoxon Non Parametric Hypothesis Test*. en-US. Ago. de 2017. URL: https://sixsigmastudyguide.com/1-sample-wilcoxon-non-parametric-hypothesis-test/ (acesso em 08/11/2023).

[4] *RPubs - Estatística não paramétrica*. URL: https://rpubs.com/santosneto/enp04 (acesso em 08/11/2023).

[5] *Student's t-test*. en. Page Version ID: 1177707120. Set. de 2023. URL: https://en.wikipedia.org/w/index.php?title=Student%27s_t-test&oldid=1177707120 (acesso em 08/11/2023).

Apêndice

```
library(tidyverse)
library(DescTools)
library(xtable)

df = read.csv2('dados_projeto.csv', col.names = c('Jogador', 'P'), sep = ',')
View(df)
df$P = as.numeric(df$P)

df = df %>% mutate(Sucesso = round(P*112))

hist(df$P)
shapiro.test(df$P)

hist(df$Sucesso)
shapiro.test(df$Sucesso)

# descrição
x = df %>% summarise(Min = min(P),
                "1st Qu." = quantile(P, .25),
                Median = median(P),
                Mean = mean(P),
                "3rd Qu." = quantile(P,.75),
                Max = max(P),
                Curt = Kurt(P),
                Skew = Skew(P)) %>% round(4) %>% xtable()

digits(x)
digits(x) = c(4,4,4,4,4,4,4,4,4)
x

# teste t

## proporção de vitórias
t.test(df$P, mu = 0.5)
t.test(df$P, mu = 0.5, alternative = "less")

## nº absoluto de vitórias
t.test(df$Sucessos, mu = 56)
t.test(df$Sucessos, mu = 56, alternative = "less")

# teste de wilcoxon
wilcox.test(df$P, mu = 0.5, conf.int = .95, alternative = "two.sided")
wilcox.test(df$P, mu = 0.5, conf.int = .95, alternative = "less")

# teste binomial
tab = data.frame(valor_p = NA, li = NA, ls = NA)
```

```
for(i in 1:length(df$Sucesso)){
  b = prop.test(df$Sucesso[i], 112, p=0.5)
  l = c(round(b$p.value,4), round(b$conf.int[1],5), round(b$conf.int[2],5))
  tab = rbind(tab, l)
}

tab = tab[-1,]; tab
row.names(tab) = NULL; tab
xtable(tab)
```