

Reprodução da Análise Quantitativa de um Experimento

Diogo C. T. Batista¹, Elissandra G. Pereira¹

¹Universidade Federal do Paraná (UFPR)
Curitiba – Paraná – Brasil

diogocezar@ufpr.br, egpereira@inf.ufpr.br

Resumo. *Este relatório descreve a reprodução da análise quantitativa do experimento apresentado no artigo “Visual design guidelines for improving learning from dynamic and interactive digital text”. Apresentamos as etapas da análise dos dados no software estatístico e mostramos os resultados obtidos.*

1. Identificação do Artigo

Selecionamos o artigo utilizado anteriormente na atividade de planejamento e execução de um experimento controlado. O artigo comportava as necessidades deste trabalho, apresentando dados brutos de um experimento e seu resultado quantitativo e já estávamos familiarizados com o conteúdo.

O artigo escolhido foi “Visual design guidelines for improving learning from dynamic and interactive digital text” [Jin 2013], que faz uma Análise de Variância Multivariada (MANOVA). Na MANOVA, é possível analisar o efeito de uma ou mais variáveis independentes sobre um conjunto de variáveis dependentes [Chagas 2016]

2. Reconhecimento dos Testes Estatísticos do Experimento

No artigo, [Jin 2013] aplicou testes com estudantes. Os estudantes deveriam ler e tentar entender um texto digital por 20 minutos. Os estudantes completaram então um “teste de estrutura” (structure test), um teste de compreensão e um teste de usabilidade.

Para a reprodução e análise exigida no prazo de entrega deste trabalho, decidimos recortar apenas o resultado dos dois primeiros testes. Além disso, os resultados escolhidos foram aqueles relacionados às diretrizes em que tivemos maior interesse e foco em nosso exercício anterior sobre experimentos controlados.

Na Figura 1 estão dispostos os resultados obtidos por [Jin 2013] em seu artigo. As duas primeiras linhas (Structure e Comprehension) foram escolhidas para nossa reprodução.

Table 2

Mean score and standard deviation for four groups on the three dependent variables.

Dependent measures	Both design guidelines (n = 36)		Structure design guidelines (n = 35)		Selective-attention design guidelines (n = 35)		Control group (n = 35)		Main effect		Interaction effect
									Structure	Selective attention	
	M	SD	M	SD	M	SD	M	SD	F value P value d	F value P value d	
Structure ^a	6.04	2.74	5.07	2.38	3.31	2.25	2.30	1.69	50.179 .000 1.35	6.536 .012 .40	.003 .955
Comprehension ^b	13.75	2.79	11.94	3.00	11.28	3.31	10.46	2.90	15.208 .000 .64	6.771 .010 .44	.933 .336
Usability ^c	4.07	.41	3.80	.48	3.85	.38	2.37	.50	121.013 .000 .93	139.463 .000 1.00	65.299 .000
Ease of use	3.94	.46	3.81	.51	3.85	.43	2.42	.67	70.276 .000 .81	77.675 .000 .84	53.138 .000
Awareness of the structure	4.14	.56	4.05	.62	3.43	.50	2.39	.64	144.893 .000 1.51	33.489 .000 .55	23.312 .000
Content comprehension	4.20	.44	3.47	.56	4.10	.42	2.41	.58	46.287 .000 .59	201.957 .000 1.55	31.344 .000
Satisfaction of use	4.02	.53	3.86	.62	4.05	.60	2.25	.77	53.598 .000 .69	82.307 .000 .92	58.429 .000

^a The range of scores was 0 ~ 12.^b The range of scores was 0 ~ 20.^c The range of scores was 0 ~ 5.

Figura 1. Resultados utilizados como base

3. Reprodução utilizando o software estatístico

A simulação dos resultados qualitativos do experimento foi feita através da linguagem *Python*, da forma descrita a seguir.

Como o artigo não fornece todas respostas do questionário, gerou-se valores aleatórios com base na média, desvio padrão e número de respostas obtidas, tanto para os testes com *Structure design guidelines*, *Selective-attention design guidelines* e o grupo de controle. Com isso, calculou-se o *valor-p* e o comparamos com os resultado apresentados pelo do autor do artigo. Utilizou-se a estratégia de executar o algoritmo 100 vezes e retirou-se a média de todos os *valor-p* encontrados, para assim podermos termos uma análise melhor para com o resultado descrito no artigo.

4. Algoritmo Criado

Para a elaboração do algoritmo utilizou-se a linguagem *Python*. Em conjunto com as bibliotecas *numpy* e *scipy*

O Código 1 mostra a definição de uma função genérica que obtém informações de um grupo qualquer e um grupo controle. São informadas as variáveis:

- `agent_average`: média do grupo agente;
- `agent_sigma`: desvio padrão do grupo agente;
- `agent_num`: quantidade de amostras do grupo agente;
- `control_average`: média do grupo de controle;
- `control_sigma`: desvio padrão do grupo de controle;
- `control_num`: quantidade de amostras do grupo de controle;
- `start`: menor resposta;
- `end`: maior resposta;

- `iterations`: quantas iterações serão realizadas para gerar a média;

```

1 import numpy as np
2 import scipy.stats as stats
3
4 def generate(agent_average, agent_sigma, agent_num, control_average, control_sigma,
5             control_num, start, end, iterations):
6     values = []
7     for x in range(iterations):
8         agent_dist = stats.truncnorm((start - agent_average) / agent_sigma, (end -
9                                     agent_average) / agent_sigma, loc=agent_average, scale=agent_sigma)
10        agent_values = agent_dist.rvs(agent_num)
11        control_dist = stats.truncnorm((start - control_average) / control_sigma, (end -
12                                       control_average) / control_sigma, loc=control_average, scale=control_sigma)
13        control_values = control_dist.rvs(control_num)
14        u_statistic, p_val = stats.mannwhitneyu(agent_values, control_values)
15        values.append(p_val)
16    return round(np.mean(values), 20)

```

Código 1. Definição da Função Genérica

Nessa função, obtém-se cada um dos parâmetros, e se gera em um range de *iterations* valores aleatórios que representem a média e o desvio padrão, considerando o grupo de controle nos cálculos.

Para cada uma das iterações, gera-se a distribuição aleatória `agent_dist` e `control_dist` e obtém-se os seus valores `agent_values` e `control_values`. Com isso utiliza-se a função `mannwhitneyu` que abstrai o teste de Mann-Whitney, usado para testar a heterogeneidade de duas amostras ordinais. Os valores retornados são agregados em um vetor, e por fim, retorna-se a média destes valores com um arredondamento.

Na sequência, o Código 2 mostra as definições da variáveis de acordo com o trabalho de [Jin 2013].

```

1 if __name__ == "__main__":
2     structure = {
3         "control": {
4             "average": 2.30,
5             "sigma": 1.69,
6             "n": 35,
7         },
8         "structure_design": {
9             "average": 5.07,
10            "sigma": 2.38,
11            "n": 35,
12            "start": 0,
13            "end": 12
14        },
15        "selective_attention": {
16            "average": 3.31,
17            "sigma": 2.25,
18            "n": 35,
19            "start": 0,
20            "end": 12
21        },
22    }
23    comprehension = {
24        "control": {
25            "average": 10.46,
26            "sigma": 2.90,
27            "n": 35,
28        },
29        "structure_design": {
30            "average": 11.94,

```

```

31     "sigma": 3.00,
32     "n": 35,
33     "start": 0,
34     "end": 20
35 },
36 "selective_attention": {
37     "average": 11.28,
38     "sigma": 3.31,
39     "n": 35,
40     "start": 0,
41     "end": 20
42 },
43 }
44 result = {
45     "structure": {
46         "structure_design": generate(
47             structure["structure_design"]["average"],
48             structure["structure_design"]["sigma"],
49             structure["structure_design"]["n"],
50             structure["control"]["average"],
51             structure["control"]["sigma"],
52             structure["control"]["n"],
53             structure["structure_design"]["start"],
54             structure["structure_design"]["end"],
55             100
56         ),
57         "selective_attention": generate(
58             structure["selective_attention"]["average"],
59             structure["selective_attention"]["sigma"],
60             structure["selective_attention"]["n"],
61             structure["control"]["average"],
62             structure["control"]["sigma"],
63             structure["control"]["n"],
64             structure["selective_attention"]["start"],
65             structure["selective_attention"]["end"],
66             100
67         )
68     },
69     "comprehension": {
70         "structure_design": generate(
71             comprehension["structure_design"]["average"],
72             comprehension["structure_design"]["sigma"],
73             comprehension["structure_design"]["n"],
74             comprehension["control"]["average"],
75             comprehension["control"]["sigma"],
76             comprehension["control"]["n"],
77             comprehension["structure_design"]["start"],
78             comprehension["structure_design"]["end"],
79             100
80         ),
81         "selective_attention": generate(
82             comprehension["selective_attention"]["average"],
83             comprehension["selective_attention"]["sigma"],
84             comprehension["selective_attention"]["n"],
85             comprehension["control"]["average"],
86             comprehension["control"]["sigma"],
87             comprehension["control"]["n"],
88             comprehension["selective_attention"]["start"],
89             comprehension["selective_attention"]["end"],
90             100
91         )
92     }
93 }
94 print(result)

```

Código 2. Definições das Variáveis

5. Resultados

Após a execução do algoritmo, obteve-se resultados disponíveis no Código 3.

```
1 {  
2   'structure': {  
3     'structure_design': 0.0001331037517740879,  
4     'selective_attention': 0.050186023592034855  
5   },  
6   'comprehension': {  
7     'structure_design': 0.07122011277566724,  
8     'selective_attention': 0.17129878841709156  
9   }  
10 }
```

Código 3. Saída do script

A Tabela 1 apresenta uma comparação dos resultados obtidos na nossa reprodução da análise e os originais do artigo. Pode-se notar que o resultado obtido para Structure Design (para o tipo Structure) possui equivalência com os apresentados em [Jin 2013]. Os outros três valores possuem maior diferença com os originais devido às limitações do método aplicado de geração de valores aleatórios.

Experimento	Tipo	pvalue (autor)	pvalue (gerado)
Structure Design	Structure	.000	.000
Selective Attention	Structure	.012	.050
Structure Design	Comprehension	.000	.071
Selective Attention	Comprehension	.010	0.17

Tabela 1. Resultados obtidos na análise

6. Conclusão

A identificação de um artigo que apresente todos os dados brutos necessários para reprodução exigida no trabalho foi uma dificuldade. Como forma de mitigar esta dificuldade, escolhemos o artigo estudado anteriormente para o planejamento e execução de um experimento controlado.

O artigo selecionado apresenta uma rica descrição de diferentes métodos de teste, avaliação e análise em um experimento, além do conteúdo relevante para IHC sobre diretrizes de design em materiais de ensino digital. O estudo feito para este trabalho da disciplina permitiu uma exploração enriquecedora do artigo.

Apesar das limitações causadas pela ausência dos dados brutos, o resultado da reprodução conseguiu replicar de forma satisfatória um dos resultados obtidos pelo trabalho de [Jin 2013]. Percebemos que apesar da reprodução não ter sido ideal, pudemos treinar e estudar de forma adequada o procedimento de análise dos dados.

Referências

Chagas, E. (2016). Módulo 5 - análise multivariada no spss.

Jin, S.-H. (2013). Visual design guidelines for improving learning from dynamic and interactive digital text. *Computers & Education*, 63:248 – 258.