



**Universidade Estadual de Campinas
Instituto de Matemática, Estatística e Computação Científica
Departamento de Estatística**

Relatório Projeto Final de ME623

**Eliane Ramos de Siqueira RA:155233
Guilherme Pazian RA:160323**

Professora: Mariana Rodrigues Motta

Campinas-SP, 06 de Dezembro de 2016

1. Introdução

Atualmente muito se discute sobre eficiência no âmbito estudantil, nós, alunos, buscamos sempre a melhor maneira para se aproveitar o tempo de estudo a fim de se obter um maior rendimento no processo de aprendizagem, porém, muitas vezes temos dificuldades em manter a concentração mesmo com esforço, motivação e disciplina. A partir disto, muitos preocupam-se em entender e/ou controlar outros fatores além daqueles intrínsecos à cada aluno. Entre estes fatores, os mais comumente apontados são relacionados ao horário e ambiente de estudo, porém estes acabam sendo tomados por uma forma muito subjetiva, uma vez que não conseguimos definir claramente quais são os fatores que têm influência significativa no processo de aprendizagem.

Duas das habilidades mais importantes para se obter um bom rendimento no processo de aprendizado são “rapidez de raciocínio” e o “alto nível de concentração”. Um dos fatores que mais são apontados pelos alunos como grande influência no rendimento nos estudos é o período do dia (manhã, tarde e noite), e por isso, este será o fator de interesse deste experimento. Temos interesse também em obter resultados para o nosso perfil estudantil.

Entendemos que a área de estudo, estilo de vida e idade dos estudantes pode apresentar alguma influência nos resultados do experimento, por isso, dado nosso interesse, decidimos restringir este experimento ao ambiente em que estamos inseridos, ou seja, decidimos obter resultados somente para o perfil estudantil: “Estudantes de estatística da UNICAMP que tenham entre 20 e 26 anos”. Entendemos que tal perfil é bem razoável, pois contempla a grande maioria dos alunos do nosso curso e estes, quanto ao experimento, são semelhantes entre si, uma vez que a área de estudo é a mesma entre eles, o estilo de vida é similar e a faixa de idade definida não deve permitir que ocorram grandes diferenças de resultados entre os indivíduos devido à idade dos mesmos.

Dado o ambiente quantitativo em que nós estamos inseridos (Curso de Estatística), decidimos que é razoável quantificar a rapidez de raciocínio e o nível de concentração medindo o tempo de resolução de um Sudoku, uma vez que, resolver o Sudoku requer paciência, atenção e um pouco de lógica matemática para ajudar durante o processo de encaixar os números nos seus devidos lugares.

Por hora, não é de nosso interesse fazer inferência à alguma população (por isso não utilizaremos técnicas de amostragem), mas sim compreender a influência dos períodos do dia no rendimento no processo de aprendizagem de alunos com perfil semelhante ao nosso.

2. Objetivo

O objetivo do nosso experimento é verificar se existe influência significativa do período do dia (manhã, tarde, noite) na rapidez de raciocínio e no nível de concentração em alunos com perfil semelhante ao nosso (estudantes de estatística da UNICAMP que tenham entre 20 e 26 anos).

3. Planejamento

Queremos comparar o desempenho em três períodos do dia, manhã, tarde e noite, onde: o período da manhã corresponde ao período anterior ao meio dia, o período da tarde à após às 13:00hs e antes das 17:30hs e o período da noite após às 19:00hs. A variável resposta, tempo de resolução de um sudoku, será medida por um cronômetro. O Sudoku é um quebra cabeça baseado na colocação lógica de números de 1 a 9, o mesmo é formado por 9 linhas na horizontal, 9 colunas na vertical e 9 quadrados menores, subdivididos pelo quadrado maior. O jogo consiste em

encontrar o lugar certo para inserirmos um número de modo que ele não se repita na mesma linha, na mesma coluna ou no mesmo subquadrado (veja a imagem 1 em anexo).

Sabemos que a capacidade e rapidez de raciocínio varia de pessoa para pessoa e que o dia da semana pode influenciar no desempenho, porém não estamos interessados em medir tal variação, o que nos indica que é razoável considerar os dias e as pessoas como fatores ruídos. A fim de tentar minimizar a influência da relação supracitada sobre o experimento, vamos considerar os dias e as pessoas como blocos. Portanto, cada indivíduo será submetido à todos os tratamentos somente uma vez cada um.

4. Delineamento Experimental

Para realizar as análises, o delineamento que vamos utilizar é o **Quadrado latino**, onde os fatores ruídos serão: “Dia da Semana” e “Pessoa”. Vamos considerar três dias da semana, segunda, quarta e sexta e três pessoas diferentes (Pessoa A, Pessoa B e Pessoa C). Devido à pequena dimensão do quadrado latino (3x3) teremos poucos graus de liberdade para o erro, o que pode acarretar equívocos na interpretação dos resultados. A fim de resolver esta situação, iremos realizar 3 replicações nos quadrados latinos, mantendo os mesmos dias da semana em todas as replicações e variando as pessoas. Ou seja, utilizaremos então o delineamento **Quadrados latinos com Replicação**. Onde o fator ruído “Dia da semana” manteremos os mesmos e fator ruído pessoa agora terá nove níveis (A,B,C,D,E,F,G, H,I) (veja a tabela com a representação dos quadrados latinos em anexo).

Dado o delineamento de experimento escolhido, usaremos o seguinte modelo:

$$Y_{ijk} = \mu + \alpha_{i(l)} + \tau_j + \beta_k + \psi_l + \varepsilon_{ijkl} \begin{cases} i = 1, 2, 3 \\ j = 1, 2, 3 \\ k = 1, 2, 3 \\ l = 1, 2, 3 \end{cases}$$

onde $\alpha_{i(l)}$ representa o efeito do i-ésimo dia na replicação l. τ_j representa o efeito do j-ésimo período do dia, β_k representa o efeitos da k-ésima pessoa e μ é a média global. ε_{ijk} é o erro experimental com $\varepsilon_{ijk} \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$.

5. Análises

Pela figura do boxplot abaixo, podemos notar que o tempo de resolução do sudoku parece não diferir para dias diferentes. Porém, há uma diferença significativa de pessoa para pessoa. Já o período, parece interferir no tempo de resolução, uma vez que no período da tarde, os tempos são menores. Nota-se também a presença de dois outliers, para os períodos da tarde e noite, e para os dias 1 e 3 (Segunda-feira e Sexta-feira).

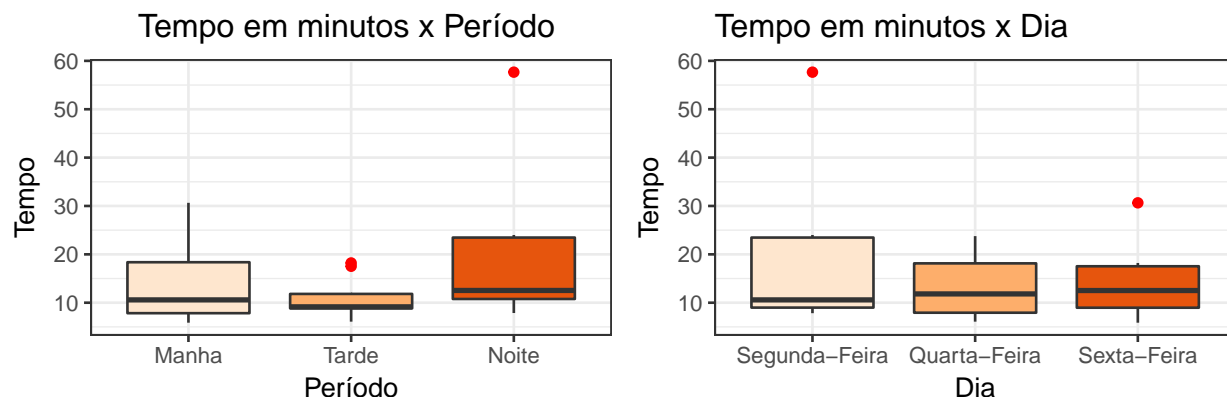


Figura 2: Boxplot Comparativo

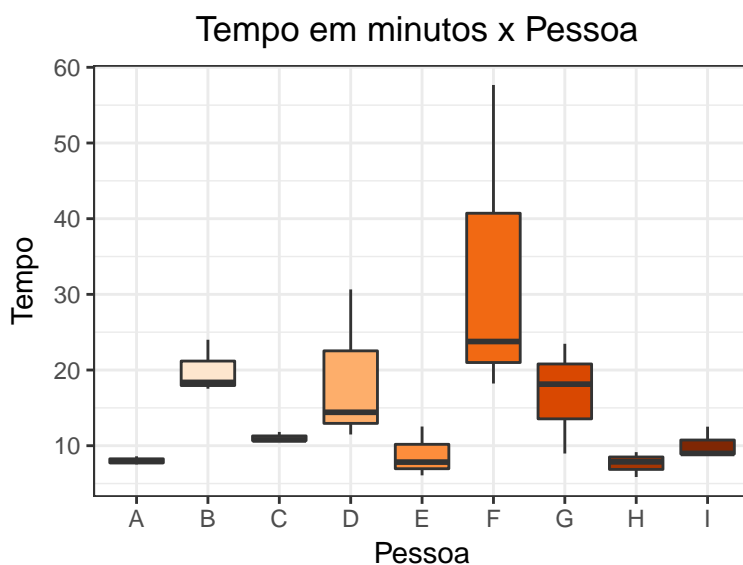


Figura 3: Boxplot Comparativo

Na tabela abaixo, temos os valores para a ANOVA, onde queremos testar: $H_0 : \mu_1 = \mu_2 = \mu_3$ vs $H_1 : \mu_i \neq \mu_j$ para pelo menos um par (i, j) . Observe que o teste de igualdade de médias para o período do dia acusou um p-valor de 0.386, o que indica, à níveis de significância usuais (0.01 a 0.10) que não existe influência significativa do período do dia (manhã, tarde, noite) em relação ao tempo de solução de um sudoku. Mais especificamente, não foi constatado, neste experimento, indícios de que exista influência significativa do período do dia (manhã, tarde, noite) na rapidez de raciocínio e no nível de concentração em alunos com perfil semelhante ao nosso (estudantes de estatística da UNICAMP que tenham entre 20 e 26 anos), levando-se em consideração que estas habilidades (rapidez de raciocínio e no nível de concentração) foram medidos via tempo de resolução de um sudoku. Porém não pode-se confiar nestes resultados sem antes fazer a análise de resíduos, se estes indicarem um bom ajuste os resultados são válidos caso contrário, os resultados não são válidos.

Tabela 2: Resultados para a ANOVA

Observando os gráficos abaixo, identificamos uma tendência no gráfico “Residual vs Fitted” o que indica uma heterocedasticidade no modelo. O gráfico “Normal Q-Q” apresenta quase todos os pontos muito próximos da reta teórica, o que indica normalidade nos resíduos, neste mesmo gráfico observa-se também dois pontos discrepantes,

	GL	Soma de Quadrados	Quadrado Médio	Estatística F	P-Valor
replicacao	2.00	395.94	197.97	1.456	0.266
peessoa	2.00	163.37	81.68	0.601	0.562
Periodo	2.00	277.32	138.66	1.02	0.386
replicacao:Dia	6.00	221.82	36.97	0.272	0.941
Residuals	14.00	1903.54	135.97		

estes estão consideravelmente distanciados dos demais e da reta teórica, portanto estes, são possíveis outliers. Dadas as observações feitas acima, concluímos que o modelo não teve um bom ajuste, portanto os resultados e conclusões obtidos a partir da tabela ANOVA não são confiáveis. Podemos citar alguns fatores externos aos quais tiveram influência não controlada adequadamente, um dos mais importantes é que a experiência anterior dos participantes quanto ao jogo não era constante, e alguns, por falta de experiência no jogo acabaram tendo desempenho crescente durante o experimento. Outro fator que não pode ser controlado adequadamente foi o ambiente em que se solucionava o sudoku, pois aspectos como silêncio, podem influenciar no desempenho dos participantes do experimento. Contudo, mesmo não apresentando resultados que são de fato confiáveis, conseguimos obter informações relativas ao assunto abordado, e este experimento pode servir de referencia para outros futuros experimentos com objetivo semelhante a este. Como sugestão para próximos experimentos com objetivos semelhantes ao nosso, sugerimos adotar medidas a fim de assegurar que os participantes realizem o experimento em ambientes similares, assim como, selecionar participantes com o mesmo nível de experiência no teste que será realizado.

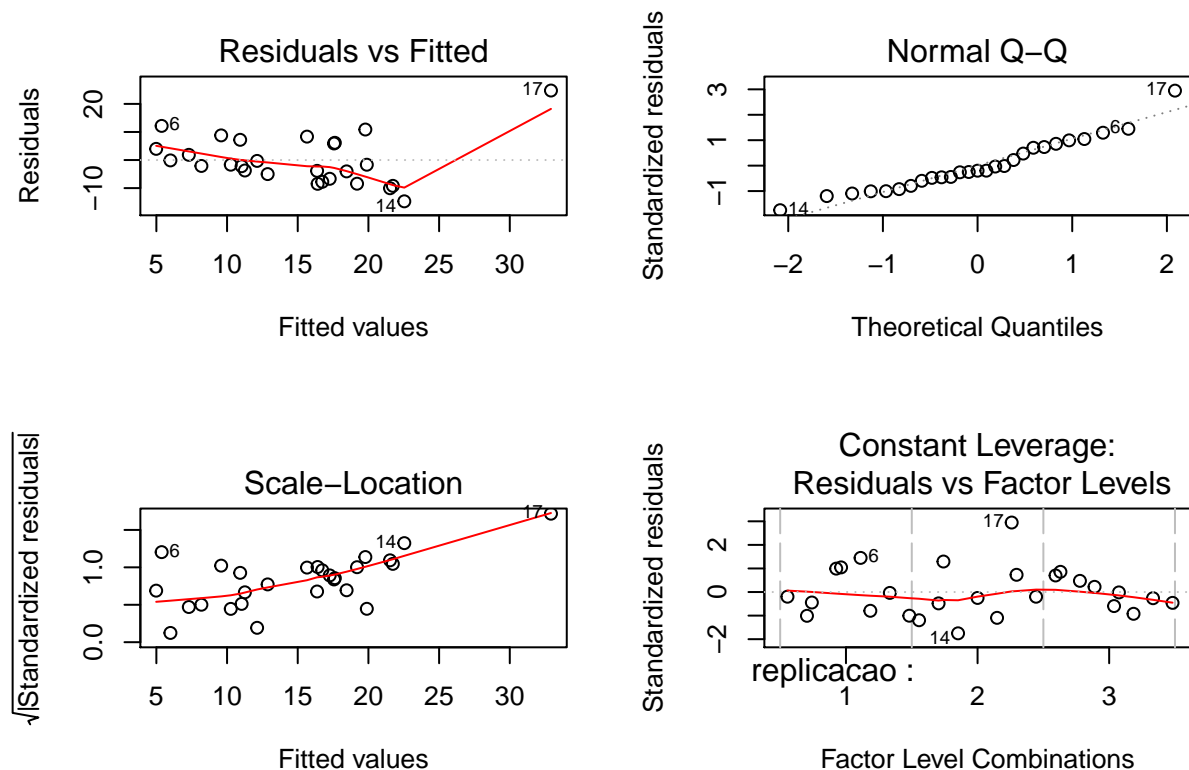


Figura 4: Análise de Resíduo

Anexo

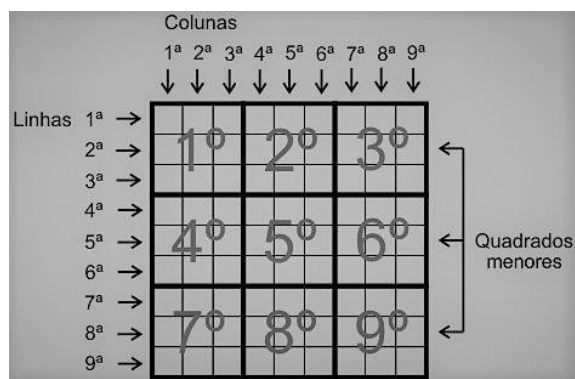


Figura 1: Estrutura do Sudoku

1º Qua-					2º Qua-					3º Qua-				
drado	Dia da				drado	Dia da				drado	Dia da			
Latino:	Semana	Pessoa			Latino:	Semana	Pessoa			Latino:	Semana	Pessoa		
		A	B	C			D	E	F			G	H	I
	Seg	T	N	M		Seg	T	M	N		Seg	N	T	M
	Qua	N	M	T		Qua	N	T	M		Qua	M	N	T
	Sex	M	T	N		Sex	M	N	T		Sex	T	M	N

Seg: Segunda-feira Qua: Quarta-feira Sex:Sexta-feira Pi= Pessoa i. (i=1,...,9) M: Período da manhã T: Período da Tarde N: Período da Noite

Observação: Os quadrados latinos acima foram confeccionados a partir do código para aleatorização de quadrados latinos disponibilizado na página do curso.