



MATEUS CARVALHO GONÇALVES
PEDRO ANTÔNIO DE SOUZA

**DEFINIÇÃO DE MELHOR(ES) ALGORITMO(S) POR MEIO DE TESTE NÃO
PARAMÉTRICO E *POST-HOC*: UM ESTUDO DE CASO**

LAVRAS - MG

2021

SUMÁRIO

1 INTRODUÇÃO	2
2 DEFINIÇÕES DE CONCEITOS	3
2.1 MÉTODO ESTATÍSTICO NÃO PARAMÉTRICO	3
2.2 PROBLEMA DE MINIMIZAÇÃO	3
2.3 ALGORITMOS DE APROXIMAÇÃO	3
2.4 PROBLEMA DE OTIMIZAÇÃO	3
2.5 ALGORITMO DE BUSCA LOCAL	4
2.6 FUNÇÃO OBJETIVO	4
3 METODOLOGIA	5
4 RESULTADOS E DISCUSSÃO	7
4.1 ANÁLISE DOS ALGORITMOS COM BUSCA LOCAL	7
4.2 ANÁLISE DOS ALGORITMOS SEM BUSCA LOCAL	7
4.3 ANÁLISE GERAL ALGORITMOS	8
5 CONCLUSÃO	9
6 REFERÊNCIAS BIBLIOGRÁFICAS	10
APÊNDICE A – Tabelas de atribuição dos postos para testes de Friedman	11
APÊNDICE B - Tabelas de resultados dos testes de Nemenyi	14
ANEXO A – Resultados de algoritmos aplicados em 30 matrizes simétricas oriundas de seis áreas de aplicação (A1 a A6) e t(s) denota tempo em segundos	16
ANEXO B – Tabela de distribuição do χ^2	17
ANEXO C – Critical Values of Studentized Range Distribution (q) for Familywise $\alpha = 0.05$	18

1 INTRODUÇÃO

A disciplina Complexidade e Projetos de Algoritmos aborda a análise dos custos de execução de diferentes algoritmos que solucionam um mesmo problema, a fim de definir qual possui melhor desempenho.

Em algoritmos que retornam a solução ótima para o problema de interesse, chamados de exatos, define-se como o melhor aquele que possui a execução mais rápida. Contudo, existem problemas computacionais impossíveis de serem resolvidos com algoritmos exatos, seja pela solução ótima ser não computável ou de custo inviável. Para esses problemas, utiliza-se algoritmos de aproximação a fim de encontrar, eficientemente, uma solução de valor aproximado ao valor ótimo (CARVALHO *et al.*, 2001).

Considerando um problema de minimização, o presente trabalho tem como objetivo definir, se possível, dentre os algoritmos A, R, B e K (com e sem aplicação de busca local), qual é o melhor. Visto que os quatro algoritmos são de aproximação e possuem complexidades assintóticas equivalentes, espera-se que o melhor seja aquele que retorna o menor valor para a função objetivo (FO). Porém, é necessário realizar uma análise estatística com valores obtidos nas execuções dos algoritmos para que seja possível definir com mais certeza essa relação de desempenho.

Para análises desse tipo, é recomendada a utilização de testes não paramétricos pois não requerem a caracterização de alguns parâmetros como a distribuição da população e ter uma amostra suficientemente grande (HESSE *et al.*, 2017). Os testes não paramétricos definem se os resultados possuem ou não uma diferença significativa entre si. Em caso de positivo, pode-se realizar testes *post-hoc* para determinar o melhor algoritmo.

Para isso, o estudo dispõe dos resultados da execução dos algoritmos citados no parágrafo anterior para 30 instâncias de um problema, apresentados no Anexo A. A tabela relaciona o resultado obtido por um algoritmo com uma determinada instância do problema. Alguns dos dados presentes são: o valor da função objetivo que se quer otimizar (FO_0), valor da função objetivo após aplicação de um algoritmo de otimização (FO) e algumas métricas que ajudam a definir qual algoritmo obteve melhor desempenho por área de aplicação (ρ e v).

O restante deste documento está organizado da seguinte forma: a Seção 2 discorre sobre conceitos importantes para a compreensão do texto; a Seção 3 descreve a metodologia de desenvolvimento do estudo; a Seção 4 contém os resultados da aplicação dos testes não paramétricos e discute as principais descobertas; por fim, a Seção 5 apresenta as conclusões do estudo.

2 DEFINIÇÕES DE CONCEITOS

Esta seção define, a partir de citações diretas, conceitos básicos fundamentais para o bom entendimento do texto. A escolha de citações diretas reflete sobre a relação entre a necessidade de expor os termos com clareza e o grau de familiaridade dos autores com o assunto.

2.1 MÉTODO ESTATÍSTICO NÃO PARAMÉTRICO

Um método estatístico é chamado de não paramétrico se não faz suposições sobre a distribuição da população ou o tamanho da amostra. Isso está em contraste com a maioria dos métodos paramétricos em estatísticas elementares que assumem que o conjunto de dados usado é quantitativo, a população tem uma distribuição normal e o tamanho da amostra é suficientemente grande. Em geral, as conclusões tiradas de métodos não paramétricos não são tão poderosas quanto as paramétricas. No entanto, como os métodos não paramétricos fazem menos suposições, eles são mais flexíveis, mais robustos e aplicáveis a dados não quantitativos. (HESSE *et al.*, 2017).

2.2 PROBLEMA DE MINIMIZAÇÃO

Segundo Carvalho *et al.* (2001), “um problema de minimização está interessado nas soluções viáveis de valor mínimo”.

2.3 ALGORITMOS DE APROXIMAÇÃO

[...] um algoritmo de aproximação não é simplesmente uma heurística: ele garante encontrar, eficientemente, um elemento do domínio cujo valor guarda uma relação pré-estabelecida com o valor ótimo. (CARVALHO *et al.*, 2001).

2.4 PROBLEMA DE OTIMIZAÇÃO

Carvalho *et al.* (2001) afirma que “problemas de otimização têm o objetivo de encontrar um ponto ótimo (mínimo ou máximo) de uma função definida sobre um certo domínio”.

Ainda abordando problemas de otimização:

Em tais problemas, pode haver muitas soluções possíveis. Cada solução tem um valor, e desejamos encontrar uma solução com um valor ótimo (mínimo ou máximo). Chamamos tal solução uma solução ótima para o problema, em lugar de a chamarmos de a solução ótima, pois podem existir várias soluções que alcançam o valor ótimo. (CORMEN *et al.*, 2002).

2.5 ALGORITMO DE BUSCA LOCAL

Um algoritmo de busca local define, para cada solução, uma vizinhança composta por um conjunto de soluções com características “muito próximas”. Dada uma solução corrente, uma das formas de implementar um algoritmo de busca local é percorrer a vizinhança dessa solução em busca de outra com valor menor (para um problema de minimização). Se tal solução vizinha for encontrada, torna-se a nova solução corrente e o algoritmo continua. Caso contrário, a solução corrente é um ótimo local em relação à vizinhança adotada. (VIEIRA, 2006).

2.6 FUNÇÃO OBJETIVO

Holtz (2005) define função objetivo como “a função de uma ou mais variáveis de projeto que se quer otimizar, minimizando-a ou maximizando-a.”

3 METODOLOGIA

A resolução do problema de pesquisa foi dividida em duas etapas: (1) averiguar se é possível, estatisticamente, afirmar que algum algoritmo é melhor; e (2) definir qual é esse algoritmo. Vale lembrar que o passo dois só é executado se o primeiro for verdadeiro.

Deve-se consultar o Anexo A para melhor entendimento dos dados, onde os algoritmos são listados com os resultados de execução para cada instância.

Considerando sempre as 30 instâncias, este estudo analisou qual é o melhor algoritmo em três contextos diferentes: (1) entre os algoritmos com uso de busca local; (2) entre os algoritmos sem uso de busca local; (3) todos os algoritmos.

Para resolver o primeiro passo, o teste estatístico não paramétrico de Friedman (1937) foi aplicado. Seguindo a metodologia descrita no artigo citado anteriormente, os valores dos postos foram atribuídos considerando os valores da função objetivo.

Após essa atribuição, é necessário calcular a distribuição χ^2 definida pela seguinte equação:

$$\chi^2 = \frac{12}{np(p+1)} \sum_{j=1}^p \left(\sum_{i=1}^n r_{ij} \right)^2 - 3n(p+1) \quad (1)$$

Para efeito de simplificação dos cálculos apresentados na seção 4, o valor do somatório $\sum_{i=1}^n r_{ij}$ foi previamente calculado para cada i . Nas tabelas do Apêndice A, esse somatório é representado pela linha “Soma dos postos (r_j)”.

Então, é testada a hipótese nula de que não há diferença entre as distribuições dos resultados das execuções dos algoritmos. O intervalo de confiança utilizado foi de 95%, isto é, $\alpha = 0,05$. Para este teste, o valor de χ^2 é comparado com o valor crítico c_F correspondente ao intervalo de confiança e o grau de liberdade (veja o Anexo B). Se $\chi^2 > c_F$, então a hipótese nula é descartada e é possível definir qual é o melhor algoritmo.

Se o teste de Friedman mostrar significância, isto é, descartar a hipótese nula, é possível aplicar o teste post-hoc de Nemenyi (ZAIONTZ)¹. Esta é uma análise dois-a-dois que também tem como objetivo definir se há diferença estatisticamente significativa entre os elementos dos pares. Se essa diferença existir, então a definição do melhor algoritmo é factível. Vale lembrar, portanto, que é necessário aplicar os testes em todos os pares possíveis

¹ Foi utilizada uma fonte da web pois os autores não conseguiram encontrar o trabalho original de Peter Nemenyi, apesar dos esforços.

e, assim, pode-se chegar a um ou mais algoritmos considerados como melhores na amostra analisada.

Dito isso, o teste de Nemenyi usa os valores dos postos, calculados no teste de Friedman, e é definido pela seguinte equação:

$$q = \frac{r_{\max} - r_{\min}}{s.e.} \text{ onde } s.e. = \sqrt{\frac{np(p+1)}{12}} \quad (2)$$

Calculado o valor de q , compara-se este ao valor crítico c_N (veja o Anexo C). Se $q > c_N$, então existe diferença significativa entre os elementos do par. Nesse caso, o melhor algoritmo é definido pelo menor valor de r_j , isso porque os valores dos postos foram atribuídos de forma que a menor FO recebe 1 e assim por diante.

4 RESULTADOS E DISCUSSÃO

Nesta seção os resultados são apresentados e discutidos separadamente entre os três contextos descritos na Seção 3.

Considerando a equação 1, n é a quantidade de instâncias e p é a quantidade de algoritmos. Logo, temos $n = 30$ para todos os contextos, $p = 4$ para as seções 4.1 e 4.2 e $p = 8$ para a seção 4.3.

Além disso, sabe-se que o grau de liberdade para o teste de Friedman é obtido por meio da equação:

$$gl_F = p - 1 \quad (3)$$

e para o teste de Nemenyi:

$$gl_N = n - p \quad (4)$$

Lembrando que o intervalo de confiança utilizado é de 95% (ou seja $\alpha = 0,05$), os valores críticos c_F , de acordo com o Anexo B, são:

- Para $p = 4$, então $gl_F = 3$ e $c_F = 7,815$;
- Para $p = 8$, então $gl_F = 7$, $gl_N = 22$ e $c_F = 14,067$;

e conforme o Anexo C, os valores críticos c_N são:

- Para $p = 4$, temos $gl_N = 26$ e $c_N = 3,88$;
- Para $p = 8$, então $gl_N = 22$ e $c_N = 4,722$.

4.1 ANÁLISE DOS ALGORITMOS COM BUSCA LOCAL

A Tabela 1 do Apêndice A mostra a atribuição dos postos conforme o resultado da função objetivo para os algoritmos A, R, B e K, aplicando busca local.

Calculando, então, o valor de χ^2 por meio da equação 1, obtém-se 5,41. Como $5,41 \not> 7,815$, a hipótese nula não é rejeitada e não há diferença significativa entre as distribuições dos resultados testados. Portanto, foi provado que os algoritmos possuem resultados de desempenho estatisticamente semelhantes e conclui-se que não é preciso aplicar testes *post-hoc* para definir qual deles é o melhor.

4.2 ANÁLISE DOS ALGORITMOS SEM BUSCA LOCAL

A Tabela 2 do Apêndice A mostra a atribuição dos postos conforme o resultado da função objetivo para os algoritmos A, R, B e K, sem aplicação de busca local.

Então, calculando o valor de χ^2 por meio da equação 1, obtém-se 19,45. Como $19,45 > 7,815$, o teste de Friedman mostrou significância, rejeitando a hipótese nula, e é possível aplicar o método de teste *post-hoc* de Nemenyi.

A Tabela 3 do Apêndice B mostra os resultados de execução do teste de Nemenyi (resultado da equação 2 e também a hipótese de diferença significativa). Pelo entendimento dos autores, houve uma inconsistência nesses resultados. O algoritmo A revela-se melhor que B e K, mas não há diferença significativa com R; ao passo que R é estatisticamente igual a B e K. Porém, não é sabido se essa inconsistência é realmente possível ou se é um erro metodológico.

Partindo dessa discussão, e supondo que a inconsistência é um caso possível, entende-se que A emerge como o melhor algoritmo desse grupo. R vem em segundo lugar, mas com resultados similares tanto ao algoritmo superior quanto aos inferiores. Por fim, B e K possuem desempenho equivalente.

4.3 ANÁLISE GERAL ALGORITMOS

A Tabela 4 do Apêndice A mostra a atribuição dos postos conforme o resultado da função objetivo para os algoritmos A, R, B e K em todos os cenários (aplicando ou não busca local).

Calculando, então, o valor de χ^2 por meio da equação 1, obtém-se 50,975. Como $50,975 > 14,067$, a hipótese nula também deve ser rejeitada e é possível, por meio do teste *post-hoc* de Nemenyi, definir qual é o melhor algoritmo.

O resultado de aplicação do teste de Nemenyi para este contexto é mostrado na Tabela 5 do Apêndice B. Novamente, esta tabela guarda informações sobre o resultado da equação 2 e também sobre a hipótese de diferença significativa.

Como na Seção 4.2, foram encontradas três inconsistências: o algoritmo K_{BL}^2 deveria ser diferente de B; R deveria ser diferente de B; e R deveria ser diferente de K.

Considerando os resultados e as inconsistências como possíveis, pode-se concluir que os algoritmos A_{BL} , R_{BL} , B_{BL} e A são equivalentes e revelam-se como os melhores. Logo depois vem K_{BL} (muito próximo dos melhores) e R (mais próximo dos piores), sendo que R possui mais inconsistências. Finalmente, B e K são equivalentes, como na amostra discutida na Seção 4.2.

² A notação $_{BL}$ foi utilizada para indicar algoritmos com aplicação de busca local

5 CONCLUSÃO

Conclui-se, então, que os algoritmos com aplicação de busca local não possuem (grandes) diferenças entre si, sendo necessário apenas o teste de Friedman. Já nos casos em que a busca local não é aplicada, foi possível definir que A é o melhor algoritmo. E em uma avaliação geral, foi possível definir que o grupo de algoritmos A_{BL} , R_{BL} , B_{BL} e A têm desempenho equivalente e são os melhores - e K_{BL} possui performance um pouco inferior mas similar. Lembrando que essa discussão leva em consideração o desempenho sem dividir por área de aplicação.

Visto isso, percebe-se que a aplicação de busca local melhora significativamente o desempenho dos algoritmos.

Por fim, estudos futuros podem complementar este aplicando outros métodos de testes estatísticos não paramétricos e *post-hoc*, comparando os resultados obtidos por cada um. Outra abordagem que também se pode seguir, é a avaliação dos algoritmos por área de aplicação.

6 REFERÊNCIAS BIBLIOGRÁFICAS

CARVALHO, M. H. et al. **Uma Introdução Sucinta a Algoritmos de Aproximação**. IMPA - Instituto de Matemática Pura e Aplicada, 2001.

CORMEN, T. H. et al. **Algoritmos: teoria e prática**. 2. ed. Editora Campus, 2002.

FRIEDMAN, Milton. **The use of ranks to avoid the assumption of normality implicit in the analysis of variance**. Journal of the American Statistical Association, v. 32, n. 200, p. 675-701, 1937.

HESSE, C. A.; OFOSU, J. B.; NORTEY, E. N. **Introduction to nonparametric statistical methods**. Accra: Akrong Publications Ltd, 2017.

HOLTZ, G. C. da C. **Traçado automático de envoltórias de esforços em estruturas planas utilizando um algoritmo evolucionário**. 2005. Dissertação de Mestrado. PUC-Rio.

VIEIRA, C. E. C. **Heurísticas para o Problema das p-Medianas Conectadas**. 2006. Tese de Doutorado. PUC-Rio.

ZAIONTZ, C. Friedman Test Post-hoc Analysis. **Real Statistics Using Excel**. Disponível em:
<https://www.real-statistics.com/anova-repeated-measures/friedman-test/friedman-test-post-hoc-analysis>. Acesso em 17 set. 2021.

APÊNDICE A – Tabelas de atribuição dos postos para testes de Friedman

Tabela 1: Atribuição dos postos para algoritmos com aplicação de busca local

Instância	A	R	B	K
1	1	3,5	2	3,5
2	1	4	3	2
3	1	4	2	3
4	1,5	1,5	4	3
5	1,5	3	1,5	4
6	1,5	1,5	3	4
7	1	2	4	3
8	1	2	3	4
9	1	2	3	4
10	1	3	4	2
11	1	2	3	4
12	4	1,5	1,5	3
13	4	1	2	3
14	4	3	1	2
15	1	2	4	3
16	4	3	2	1
17	1	3	4	2
18	3	4	1	2
19	3,5	1,5	1,5	3,5
20	1	2	4	3
21	4	3	2	1
22	4	2	2	2
23	1	2	3	4
24	1	3	4	2
25	4	1,5	3	1,5
26	2	3	1	4
27	1	3	2	4
28	2	4	3	1
29	1	2	3	4
30	4	1,5	3	1,5
Soma dos postos (r_j)	62	74,5	79,5	84

Tabela 2: Atribuição dos postos para algoritmos sem aplicação de busca local

Instância	A	R	B	K
1	1,5	1,5	3,5	3,5
2	1	3	4	2
3	2	1	4	3
4	1	2	4	3
5	1,5	3	1,5	4
6	4	1	2	3
7	1	2	4	3
8	1	2	3	4
9	1	3	2	4
10	1	3	4	2
11	1	2	3	4
12	3	1,5	1,5	4
13	3	1	2	4
14	1	4	2	3
15	1	2	4	3
16	1	4	2	3
17	1	3	4	2
18	3	4	1	2
19	3,5	2	1	3,5
20	1	2	4	3
21	1,5	4	3	1,5
22	2,5	2,5	2,5	2,5
23	1	2	3	4
24	1	3	4	2
25	2	2	4	2
26	2	3	1	4
27	1	3	2	4
28	2	4	3	1
29	1	2	3	4
30	3	1	2	4
Soma dos postos (r_j)	50,5	73,5	84	92

Tabela 4: Atribuição dos postos para todos algoritmos

Instância	Com busca local				Sem busca local			
	A	R	B	K	A	R	B	K
1	1	6,5	3	6,5	3	3	6,5	6,5
2	1	7	5	3	2	6	8	4
3	1	7,5	2,5	5,5	4	2,5	7,5	5,5
4	1,5	1,5	6	5	3	4	8	7
5	2,5	5,5	2,5	7,5	2,5	5,5	2,5	7,5
6	2	2	4	5	8	2	6	7
7	1,5	3,5	7,5	5,5	1,5	3,5	7,5	5,5
8	1	3	4	7	2	5	6	8
9	1	3	2	6	4	7	5	8
10	1	3	7	2	4	6	8	5
11	1	3	4	5	2	6	7	8
12	8	1,5	1,5	3	6	4,5	4,5	7
13	8	1	2	3	6	4	5	7
14	8	3	1	2	4	7	5	6
15	1	2	5	3	4	6	8	7
16	6	5	1	4	2	8	3	7
17	1	4	5	3	2	7	8	6
18	5	6	1	2	7	8	3	4
19	3,5	1,5	1,5	3,5	7,5	6	5	7,5
20	2	3	7	5	1	4	8	6
21	8	3	2	1	4,5	7	6	4,5
22	8	2	2	2	5,5	5,5	5,5	5,5
23	1	3	5	7	2	4	6	8
24	1	5	7	3	2	6	8	4
25	8	1,5	3	1,5	5	5	7	5
26	4	5	1	7	3	6	2	8
27	1	5	2	7	3	6	4	8
28	2	6	4	1	5	8	7	3
29	2	3	5	6	1	4	7	8
30	7	1,5	4	1,5	6	3	5	8
Soma dos postos (r_j)	99	107,5	107,5	123,5	112,5	159,5	179	191,5

APÊNDICE B - Tabelas de resultados dos testes de Nemenyi

Tabela 3: Resultados do teste de Nemenyi para algoritmos sem busca local

Comparação	q	Valor crítico (c_n)	Possui diferença significativa? ($q > c_n$)
A - R	3,252691193	3,88	NÃO
A - B	4,737615434		SIM
A - K	5,868986284		SIM
R - B	1,48492424		NÃO
R - K	2,61629509		NÃO
B - K	1,13137085		NÃO

Tabela 5: Resultados do teste de Nemenyi para contexto geral (todos os algoritmos)

Comparação	q	Valor crítico (c_n)	Possui diferença significativa? ($q > c_n$)
$A_{BL} - R_{BL}$	0,6335525936	4,722	NÃO
$A_{BL} - B_{BL}$	0,6335525936		NÃO
$A_{BL} - K_{BL}$	1,826122182		NÃO
$A_{BL} - A$	1,00623059		NÃO
$A_{BL} - R$	4,509403755		NÃO
$A_{BL} - B$	5,96284794		SIM
$A_{BL} - K$	6,894542931		SIM
$R_{BL} - B_{BL}$	0		NÃO
$R_{BL} - K_{BL}$	1,192569588		NÃO
$R_{BL} - A$	0,3726779962		NÃO
$R_{BL} - R$	3,875851161		NÃO
$R_{BL} - B$	5,329295346		SIM
$R_{BL} - K$	6,260990337		SIM
$B_{BL} - K_{BL}$	1,192569588		NÃO
$B_{BL} - A$	0,3726779962		NÃO
$B_{BL} - R$	3,875851161		NÃO
$B_{BL} - B$	5,329295346		SIM
$B_{BL} - K$	6,260990337		SIM
$K_{BL} - A$	0,8198915917		NÃO
$K_{BL} - R$	2,683281573		NÃO
$K_{BL} - B$	4,136725758		NÃO
$K_{BL} - K$	8,896813174		SIM
$A - R$	3,503173165		NÃO
$A - B$	4,95661735		SIM
$A - K$	5,888312341		SIM
$R - B$	1,453444185		NÃO
$R - K$	2,385139176		NÃO
$B - K$	0,9316949906		NÃO

Observação: a notação $_{BL}$ foi utilizada para indicar algoritmos com aplicação de busca local.

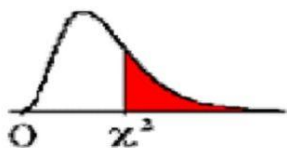
ANEXO A – Resultados de algoritmos aplicados em 30 matrizes simétricas oriundas de seis áreas de aplicação (A1 a A6) e t(s) denota tempo em segundos

Instância	n	E	FO ₀	com busca local						sem busca local														
				A			B			K			A			R			B			K		
				FO	t(s)	R	FO	t(s)	R	FO	t(s)	K	FO	t(s)	R	FO	t(s)	K	FO	t(s)	R	FO	t(s)	K
				A1																				
1	525825	3674625	525820	510	1,2	514	1,4	513	1,2	514	1,3	513	0,8	513	0,6	514	0,3	514	0,7	514	0,3	514	0,7	
2	123440	3085406	4501	2253	1,6	2374	1,5	2353	1,6	2338	1,5	2270	0,3	2364	0,3	2386	0,2	2344	0,3	2344	0,3	2344	0,3	
3	1465137	21005389	36616	40171	2,2	42770	2,6	40178	2,4	40768	2,7	40571	3,9	40178	3,6	42770	2,2	40768	3,6	40768	3,6	40768	3,6	
	ρ_{FO}	e ρ_t	1031,02	0,01	10	0,23	10	0,15	10	0,16	10	0,12	3	0,15	2	0,23	0	0,16	2	0,16	2	0,16	2	
	v_{FO}		1,50	0,08		0,17		0,11		0,12		0,10		0,11		0,17		0,12		0,12		0,12		
A2																								
4	82654	574458	80916	220	2,0	220	2,2	233	2,1	232	2,3	222	0,2	228	0,1	240	0,1	238	0,2	238	0,2	238	0,2	
5	102158	711558	102138	251	1,2	256	1,4	251	1,2	262	1,4	251	0,3	256	0,2	251	0,1	262	0,2	251	0,1	262	0,2	
6	102158	711558	102138	263	1,6	263	1,5	264	1,6	266	1,4	279	0,3	263	0,1	268	0,1	271	0,2	271	0,2	271	0,2	
7	204316	1423116	204276	251	1,6	261	1,5	277	1,6	271	1,4	251	0,5	261	0,2	277	0,1	271	0,3	271	0,3	271	0,3	
8	1228045	8580313	1226000	857	2,2	865	2,6	907	2,4	920	2,6	859	2,4	915	1,7	919	0,9	924	1,6	924	1,6	924	1,6	
	ρ_{FO}	e ρ_t	3402,50	0,00	85	0,07	89	0,22	87	0,26	88	0,07	12	0,16	5	0,29	0	0,31	6	0,31	6	0,31	6	
	v_{FO}		4,9903	0,0000		0,0001		0,0003		0,0004		0,0002		0,0002		0,0005		0,0006		0,0006		0,0006		
A3																								
9	101492	1647264	100407	4608	0,4	4699	0,2	4649	0,2	4811	0,4	4719	0,3	4812	0,2	4760	0,2	4927	0,3	4927	0,3	4927	0,3	
10	259789	4242673	19627	18694	1,0	19580	1,0	20763	0,5	18714	1,0	19606	0,8	20535	0,8	21776	0,4	19627	0,8	19627	0,8	19627	0,8	
11	589446	21758924	588744	16933	4,2	17470	3,8	17478	3,8	17591	4,2	17197	3,5	17742	3,2	17750	1,5	17865	3,5	17865	3,5	17865	3,5	
12	726713	5080961	1921	1153	0,8	1129	1,0	1129	0,6	1131	1,0	1142	0,7	1141	0,8	1141	0,5	1143	0,8	1143	0,8	1143	0,8	
13	1102824	89306020	1036475	27140	10,2	21802	11,3	22467	4,2	23585	11,3	25102	8,5	23729	9,4	24453	3,5	25670	9,4	25670	9,4	25670	9,4	
14	1157456	48538952	948032	19778	6,6	17676	6,5	16265	3,0	17649	6,6	17995	5,5	19621	5,4	18054	2,5	19591	5,5	19591	5,5	19591	5,5	
	ρ_{FO}	e ρ_t	159,14	0,48	8	0,19	8	0,18	1	0,25	9	0,36	6	0,50	6	0,49	0	0,57	6	0,57	6	0,57	6	
	v_{FO}		4,34	0,02		0,04		0,10		0,01		0,06		0,10		0,16		0,06		0,06		0,06		
A4																								
15	715176	4817870	65837	3092	2,0	3134	1,1	3405	0,7	3292	1,1	3367	1,7	3413	0,9	3707	0,6	3584	0,9	3584	0,9	3584	0,9	
16	923136	40373538	17279	15873	3,6	15497	3,6	13467	2,4	14962	3,6	14669	3,0	16883	3,0	14672	2,0	16300	3,0	16300	3,0	16300	3,0	
17	943695	77651847	925946	34550	8,4	38571	18,0	38837	8,2	38556	9,5	35087	7,0	39170	15,0	39440	6,8	39155	7,9	39155	7,9	39155	7,9	
18	952203	42493817	686979	10077	4,8	10221	4,8	9698	2,4	9766	4,8	10318	4,0	10466	4,0	9930	2,0	10000	4,0	10000	4,0	10000	4,0	
19	1391349	64131971	81578	81709	7,4	80997	7,4	80997	3,7	81709	9,8	85694	6,2	84947	6,2	84746	3,1	85694	8,2	85694	8,2	85694	8,2	
20	1437960	60236322	26018	24382	7,2	24451	7,2	29395	3,6	27521	7,2	24138	6,0	24700	6,0	29695	3,0	27802	6,0	27802	6,0	27802	6,0	
21	1498023	59374451	29036	31811	7,2	26262	7,2	26254	3,6	26075	7,2	28943	6,0	29151	6,0	29142	3,0	28943	6,0	28943	6,0	28943	6,0	
22	1508065	52259885	2634	4264	4,6	3624	4,8	3624	2,2	3624	4,6	3944	3,8	3944	4,0	3944	1,8	3944	3,8	3944	3,8	3944	3,8	
	ρ_{FO}	e ρ_t	116,4	1,1	11	0,7	11	0,8	2	0,8	10	0,9	7	1,3	7	1,3	0	1,4	7	1,4	7	1,4	7	
	v_{FO}		3,29	0,72		0,37		0,42		0,45		0,53		0,68		0,72		0,75		0,75		0,75		
A5																								
23	115967	1033473	103933	595	0,2	600	0,2	617	0,1	710	0,2	597	0,2	602	0,2	619	0,1	712	0,2	712	0,2	712	0,2	
24	116158	8516500	112127	3168	1,4	3707	0,5	4070	0,4	3451	1,4	3181	1,2	3723	0,4	4087	0,3	3466	1,2	3466	1,2	3466	1,2	
25	213360	3021648	20021	1135	0,7	1131	0,7	1132	0,4	1131	0,7	1133	0,6	1133	0,6	1134	0,3	1133	0,6	1133	0,6	1133	0,6	
26	259156	4429042	212262	6659	1,1	7074	1,1	6516	0,6	7395	1,1	6599	0,9	7139	0,9	6576	0,5	7463	0,9	7463	0,9	7463	0,9	
	ρ_{FO}	e ρ_t	256,35	0,03	7	0,26	4	0,32	1	0,42	7	0,02	6	0,28	3	0,34	0	0,44	6	0,44	6	0,44	6	
	v_{FO}		3,8788	0,0009		0,0075		0,0083		0,0078		0,0006		0,0081		0,0089		0,0083		0,0083		0,0083		
A6																								
27	399130	1216334	398620	141159	0,8	141862	0,8	141288	0,5	148920	0,8	141570	0,7	142275	0,7	141699	0,4	149354	0,7	149354	0,7	149354	0,7	
28	445315	7479343	206333	43637	2,4	43892	2,4	43741	1,3	43473	2,4	43821	2,0	44078	2,0	43926	1,1	43657	2,0	43657	2,0	43657	2,0	
29	1000005	3105536	987649	412116	1,1	415051	1,1	445472	0,7	447971	1,1	408400	0,9	418863	0,9	449564	0,6	452085	0,9	452085	0,9	452085	0,9	
30	1505785	27130349	676026	206726	8,4	201534	8,4	204363	4,8	201534	8,4	206416	7,0	201837	7,0	204671	4,0	211378	7,0	211378	7,0	211378	7,0	
	ρ_{FO}	e ρ_t	9,34	0,04	4	0,03	4	0,11	1	0,15	4	0,04	3	0,05	3	0,13	0	0,22	3	0,22	3	0,22	3	
	v_{FO}		2,724	0,012		0,011		0,043		0,060		0,010		0,017		0,050		0,080		0,080		0,080		

Fonte: https://campusvirtual.ufla.br/presencial/pluginfile.php/743199/mod_resource/content/3/t1.pdf

ANEXO B – Tabela de distribuição do χ^2

Distribuição Qui-Quadrado



A tabela fornece os valores “c” tais que
 $P(\chi^2 > c) = p$

onde “n” é o número de graus de liberdade e
 “p” é a probabilidade de sucesso.

gl	0,995	0,990	0,975	0,950	0,900	0,500	0,100	0,050	0,025	0,010	0,005
1	0,000	0,000	0,001	0,004	0,016	0,455	2,706	3,841	5,024	6,635	7,879
2	0,010	0,020	0,051	0,103	0,211	1,386	4,605	5,991	7,378	9,210	10,597
3	0,072	0,115	0,216	0,352	0,584	2,366	6,251	7,815	9,348	11,345	12,838
4	0,207	0,297	0,484	0,711	1,064	3,357	7,779	9,488	11,143	13,277	14,860
5	0,412	0,554	0,831	1,145	1,610	4,351	9,236	11,070	12,833	15,086	16,750
6	0,676	0,872	1,237	1,635	2,204	5,348	10,645	12,592	14,449	16,812	18,548
7	0,989	1,239	1,690	2,167	2,833	6,346	12,017	14,067	16,013	18,475	20,278
8	1,344	1,646	2,180	2,733	3,490	7,344	13,362	15,507	17,535	20,090	21,955
9	1,735	2,088	2,700	3,325	4,168	8,343	14,684	16,919	19,023	21,666	23,589
10	2,156	2,558	3,247	3,940	4,865	9,342	15,987	18,307	20,483	23,209	25,188
11	2,603	3,053	3,816	4,575	5,578	10,341	17,275	19,675	21,920	24,725	26,757
12	3,074	3,571	4,404	5,226	6,304	11,340	18,549	21,026	23,337	26,217	28,300
13	3,565	4,107	5,009	5,892	7,042	12,340	19,812	22,362	24,736	27,688	29,819
14	4,075	4,660	5,629	6,571	7,790	13,339	21,064	23,685	26,119	29,141	31,319
15	4,601	5,229	6,262	7,261	8,547	14,339	22,307	24,996	27,488	30,578	32,801
16	5,142	5,812	6,908	7,962	9,312	15,338	23,542	26,296	28,845	32,000	34,267
17	5,697	6,408	7,564	8,672	10,085	16,338	24,769	27,587	30,191	33,409	35,718
18	6,265	7,015	8,231	9,390	10,865	17,338	25,989	28,869	31,526	34,805	37,156
19	6,844	7,633	8,907	10,117	11,651	18,338	27,204	30,144	32,852	36,191	38,582
20	7,434	8,260	9,591	10,851	12,443	19,337	28,412	31,410	34,170	37,566	39,997
21	8,034	8,897	10,283	11,591	13,240	20,337	29,615	32,671	35,479	38,932	41,401
22	8,643	9,542	10,982	12,338	14,041	21,337	30,813	33,924	36,781	40,289	42,796
23	9,260	10,196	11,689	13,091	14,848	22,337	32,007	35,172	38,076	41,638	44,181
24	9,886	10,856	12,401	13,848	15,659	23,337	33,196	36,415	39,364	42,980	45,559
25	10,520	11,524	13,120	14,611	16,473	24,337	34,382	37,652	40,646	44,314	46,928
26	11,160	12,198	13,844	15,379	17,292	25,336	35,563	38,885	41,923	45,642	48,290
27	11,808	12,879	14,573	16,151	18,114	26,336	36,741	40,113	43,195	46,963	49,645
28	12,461	13,565	15,308	16,928	18,939	27,336	37,916	41,337	44,461	48,278	50,993
29	13,121	14,256	16,047	17,708	19,768	28,336	39,087	42,557	45,722	49,588	52,336
30	13,787	14,953	16,791	18,493	20,599	29,336	40,256	43,773	46,979	50,892	53,672
35	17,192	18,509	20,569	22,465	24,797	34,336	46,059	49,802	53,203	57,342	60,275
40	20,707	22,164	24,433	26,509	29,051	39,335	51,805	55,758	59,342	63,691	66,766
45	24,311	25,901	28,366	30,612	33,350	44,335	57,505	61,656	65,410	69,957	73,166
50	27,991	29,707	32,357	34,764	37,689	49,335	63,167	67,505	71,420	76,154	79,490
55	31,735	33,570	36,398	38,958	42,060	54,335	68,796	73,311	77,380	82,292	85,749
60	35,534	37,485	40,482	43,188	46,459	59,335	74,397	79,082	83,298	88,379	91,952
70	43,275	45,442	48,758	51,739	55,329	69,334	85,527	90,531	95,023	100,425	104,215
80	51,172	53,540	57,153	60,391	64,278	79,334	96,578	101,879	106,629	112,329	116,321
90	59,196	61,754	65,647	69,126	73,291	89,334	107,565	113,145	118,136	124,116	128,299
100	67,328	70,065	74,222	77,929	82,358	99,334	118,498	124,342	129,561	135,807	140,169
110	75,550	78,458	82,867	86,792	91,471	109,334	129,385	135,480	140,917	147,414	151,948
120	83,852	86,923	91,573	95,705	100,624	119,334	140,233	146,567	152,211	158,950	163,648

ANEXO C – Critical Values of Studentized Range Distribution (q) for Familywise $\alpha = 0.05$

Denominator DF	Number of Groups (a.k.a. Treatments)							
	3	4	5	6	7	8	9	10
1	26.976	32.819	37.081	40.407	43.118	45.397	47.356	49.070
2	8.331	9.798	10.881	11.734	12.434	13.027	13.538	13.987
3	5.910	6.825	7.502	8.037	8.478	8.852	9.177	9.462
4	5.040	5.757	6.287	6.706	7.053	7.347	7.602	7.826
5	4.602	5.218	5.673	6.033	6.330	6.582	6.801	6.995
6	4.339	4.896	5.305	5.629	5.895	6.122	6.319	6.493
7	4.165	4.681	5.060	5.359	5.606	5.815	5.997	6.158
8	4.041	4.529	4.886	5.167	5.399	5.596	5.767	5.918
9	3.948	4.415	4.755	5.024	5.244	5.432	5.595	5.738
10	3.877	4.327	4.654	4.912	5.124	5.304	5.460	5.598
11	3.820	4.256	4.574	4.823	5.028	5.202	5.353	5.486
12	3.773	4.199	4.508	4.748	4.947	5.116	5.262	5.395
13	3.734	4.151	4.453	4.690	4.884	5.049	5.192	5.318
14	3.701	4.111	4.407	4.639	4.829	4.990	5.130	5.253
15	3.673	4.076	4.367	4.595	4.782	4.940	5.077	5.198
16	3.649	4.046	4.333	4.557	4.741	4.896	5.031	5.150
17	3.628	4.020	4.303	4.524	4.705	4.858	4.991	5.108
18	3.609	3.997	4.276	4.494	4.673	4.824	4.955	5.071
19	3.593	3.977	4.253	4.468	4.645	4.794	4.924	5.037
20	3.578	3.958	4.232	4.445	4.620	4.768	4.895	5.008
21	3.565	3.942	4.213	4.424	4.597	4.743	4.870	4.981
22	3.553	3.927	4.196	4.405	4.577	4.722	4.847	4.957
23	3.542	3.914	4.180	4.388	4.558	4.702	4.826	4.935
24	3.532	3.901	4.166	4.373	4.541	4.684	4.807	4.915
25	3.523	3.890	4.153	4.358	4.526	4.667	4.789	4.897
26	3.514	3.880	4.141	4.345	4.511	4.652	4.773	4.880
27	3.506	3.870	4.130	4.333	4.498	4.638	4.758	4.864
28	3.499	3.861	4.120	4.322	4.486	4.625	4.745	4.850
29	3.493	3.853	4.111	4.311	4.475	4.613	4.732	4.837
30	3.487	3.845	4.102	4.301	4.464	4.601	4.720	4.824

Fonte: <https://www.stat.purdue.edu/~lingsong/teaching/2018fall/q-table.pdf>