Estudo de Caso final - avaliação de algoritmos de otimização de custo e tempo em roteiros de viagem

Equipe 04

03 de Julho de 2017

Coordenador: Alessandro Cardoso
Relator: Bernardo Marques
Verificador: Danny Tonidandel
Monitor: Gustavo Vieira

1. Descrição do Problema

As ferramentas atuais para planejamento de viagens permitem que usuários busquem passagens e hotéis e construam seu próprio itinerário, mas não oferecem suporte para viagens com flexibilidade nas datas ou na ordenação de destinos. Para solucionar esse problema, foi proposto um novo sistema que busca as melhores combinações de passagens e hotéis para uma viagem, minimizando o custo e tempo de transporte. O sistema busca otimizar roteiros com datas flexíveis e não requer uma ordem fixa dos destinos, de forma que considera todas as possíveis configurações que atendem as especificações do usuário. Dessa forma, oferece boas opções de itinerários para viagens com esforço do usuário muito menor, representando uma melhoria significativa na experiência de planejamento de viagens.

O conjunto de restrições considera datas inicial e final, datas específicas para cada destino, ordenação dos destinos e características das acomodações e transportes. Pode ser considerado uma aplicação multiobjetivo do problema do caixeiro viajante com janelas de tempo e dependência temporal.

O experimento envolve comparações entre métodos aplicados em diferentes instâncias de problemas de roteamento. A variabilidade decorrente das características de cada problema de teste é uma possível fonte de variação espúria quando não considerada na análise dos resultados. Para a avaliar existência de diferenças significativas entre os algoritmos, bem como a influência das possíveis covariantes, tais como tamanho da instância e tempo de convergência, será utilizada a análise de covariância.

Os parâmetros experimentais desejados são:

- Nível de significância: $\alpha = 0.05$;
- Tamanho de efeito de interesse prático: $\delta^\star = 0.25$;
- Potência desejada: $(1 - \beta) = \pi \geq 0.8$.

2. Planejamento Experimental

A primeira etapa do experimento consistiu na geração dos dados experimetais a partir dos dois algoritmos de otimização. Foi realizado o cálculo do tamanho amostral pensando-se em utilizar o teste da análise de variância para avalizar a existência de diferenças significativas entre as classes de problemas. Embora a técnica permita analisar as médias e variâncias de observações de diferentes grupos, generalizando o teste-t para mais de dois deles [Campelo, 2015], ela é restrita apenas à indicação de existência ou não de diferenças entre os níveis avaliados, sem indicar quais níveis seriam diferentes [Montgomery and Runger, 2011], tampouco a influência de variáveis concomitantes.

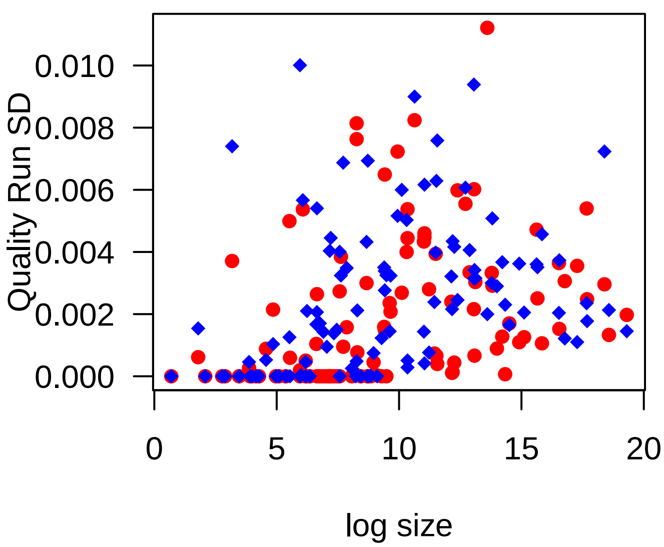
Afim de avaliar o desempenho dos dois algoritmos analisados e qual é o melhor tendo os dois parâmetros de comparação abaixo:

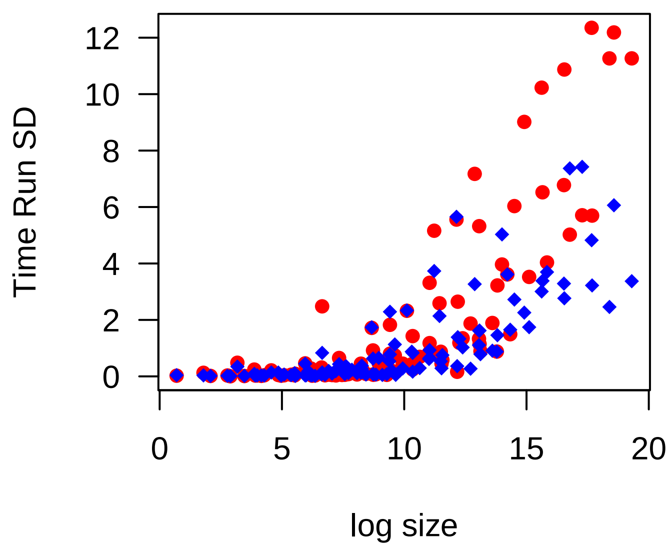
1. Tempo de execução dos algoritmos;
2. Qualidade relativa do resultado apresentado na resolução dos problemas.

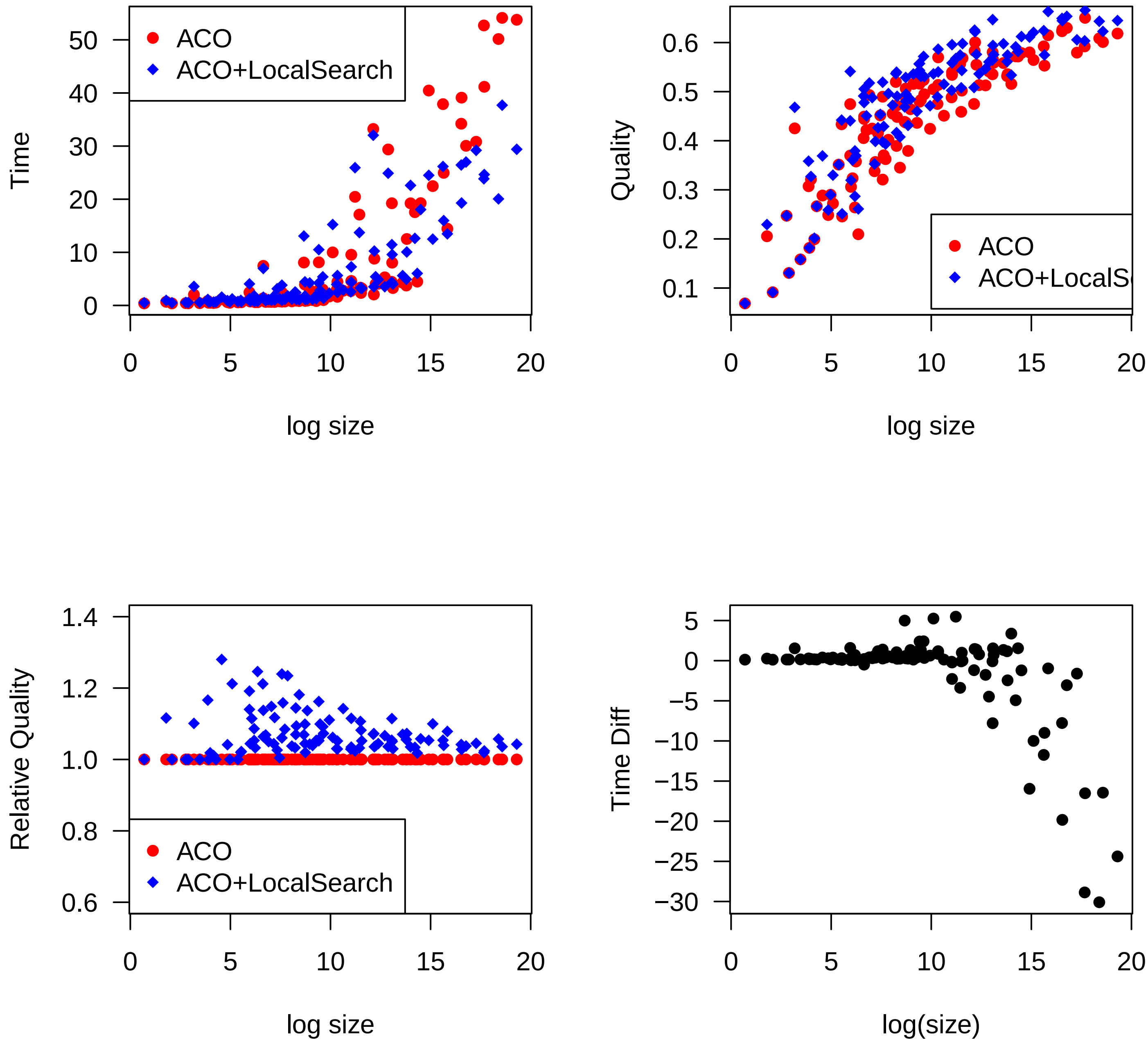
Foram geradas aleatoriamente as instâncias de problemas com diferentes tamanhos. Dessa forma é possível avaliar a variação do desempenho dos métodos à medida que a complexidade aumenta. O tamanho de um problema é determinado pelo número médio de opções de destinos do roteiro e do número médio de opções de transporte e acomodação para cada destino.

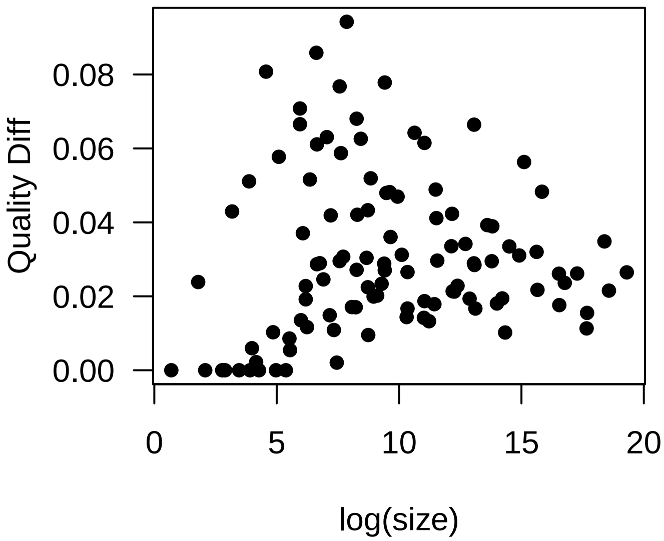
Busca-se a verificação de diferenças estatísticas significativas entre as configurações testadas. (((Uma análise exploratória qualitativa dos dados foi feita como etapa preliminar para os testes.)))

2.1. Análise Exploratória NÃO ENTENDI ESTA SEÇÃO

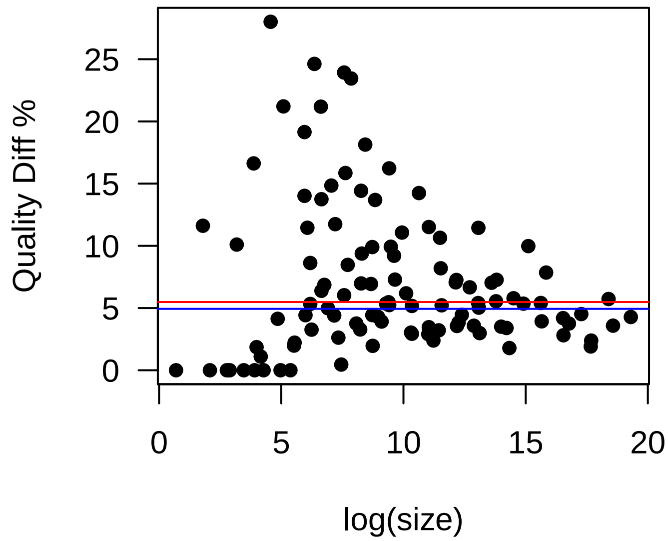


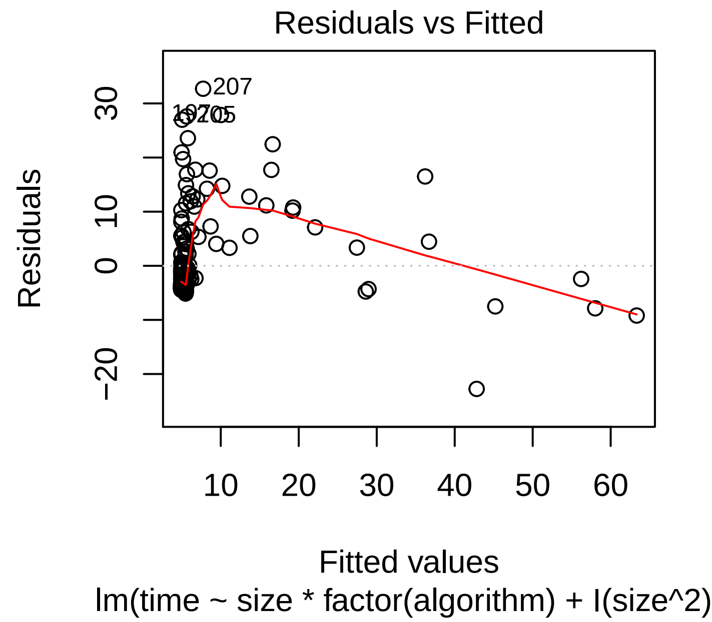


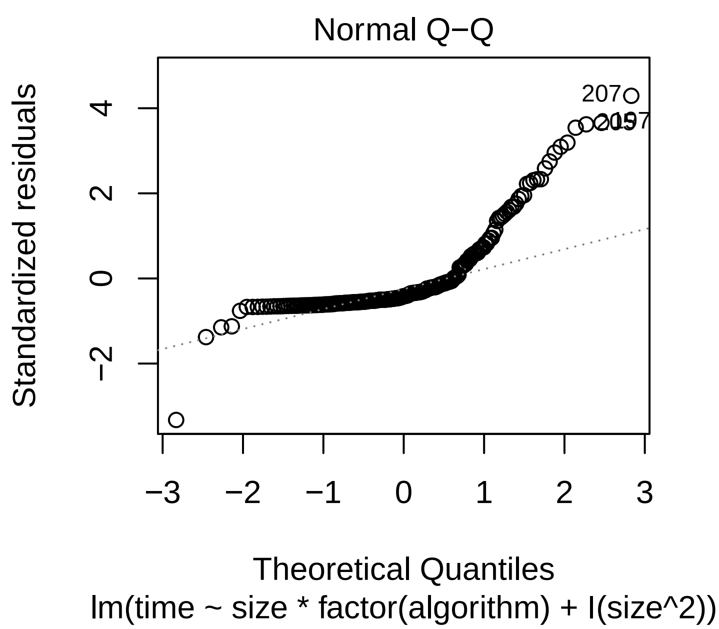

Time
50
40
30
20
10
0
ACO
ACO+LocalSearch
0
5
10
15
20
log size
Quality
0.6
0.5
0.4
0.3
0.2
0.1
ACO
ACO+LocalS
0
5
10
15
20
log size
Relative Quality
1.4
1.2
1.0
0.8
0.6
ACO
ACO+LocalSearch
0
5
10
15
20
log size
Time Diff
5
0
−5
−10
−15
−20
−25
−30
0
5
10
15
20
log(size)




```
## Anova Table (Type II tests)
##
## Response: time
##                         Sum Sq  Df  F value    Pr(>F)
## size                    9639.3   1 164.3244 < 2.2e-16 ***
## factor(algorithm)        119.5   1   2.0376    0.1549
## I(size^2)               4623.7   1  78.8223 2.976e-16 ***
## size:factor(algorithm)   952.7   1  16.2407 7.787e-05 ***
## Residuals              12377.3 211
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```








```
##               Df Sum Sq Mean Sq F value Pr(>F)
## algorithm      1    120   119.5   0.923  0.338
## Residuals    214  27709   129.5
```

2.2 Análise de Variância - ANOVA

Para a avaliar existência de diferenças significativas entre os dois parâmetros, será utilizado o teste estatístico ANOVA.

Essa técnica analisa médias e variâncias de observações de diferentes grupos para verificar se existe diferença estatístixa significativa entre as médias desses grupos. Assim, generaliza o teste t para mais de dois grupos, permitindo que sejam comparados simultâneamente [Campelo, 2015], [Montgomery and Runger, 2011]. Em outras palavras, a análise de variância é utilizada quando se quer decidir se os níveis apresentam médias diferentes em relação é uma média global μ. As diferenças entre as médias dos níveis e a média global é $\tau_i \; \forall \; i$.

Vale ressaltar que a técnica é restrita apenas à indicação de existência ou não de diferenças entre os níveis avaliados, sem indicar quais níveis seriam diferentes. Além disso, quando a análise de variância tem como resultado um indicativo de refutação da hipótese nula (1) é que podem ser evidenciados os indícios de diferenças entre os níveis.

$$\begin{cases} H_0: & \tau_i = 0, \quad \forall i \\ H_1: & \exists \;\; \tau_i \neq 0 \end{cases} \tag{1}$$

Esta etapa foi considerada como uma primeira verificação ao questionamento deste estudo por apresentar uma abordagem que pode minimizar o custo de coleta de amostras menor do que o teste de comparações múltiplas discutido na Seção 2.2. Portanto, caso não seja detectada alguma diferenças entre os níveis, as comparações múltiplas são dispensáveis, o que reduz o custo do experimento.

2.3 Comparações Múltiplas

Caso a análise ANOVA identifique a existência de diferenças entre os parâmetros analisados nos algoritmos, deve-se proceder com testes de comparação múltiplas. As comparações serão realizadas em relação ao algoritmo *ACO*, de maneira a verificar se a outra alternativa proposta é melhor. Desta forma, caso a análise de variância indique a existência de diferenças entre os níveis, será aplicado o teste de comparações múltiplas um-contra-todos (one-vs-all) de *Dunnett*, onde o *ACO com busca local* será confrontado com o algoritmo *ACO*.

Para realizar o teste de comparações múltiplas, deve-se manter o controle sobre os erros do tipo-I em cada comparação de maneira com que ele não se acumule a cada teste sucessivo. Assim, os valores de α são corrigidos para cada teste através do método de correção de Bonferroni (2):

$$\alpha_{adj} = \frac{\alpha_{família}}{K}\,, \tag{2}$$

no qual $K = 1$ consiste no número de comparações a serem feitas que no caso do teste um-contra-todos de *Dunnett*. O número de comparações nesse caso é dado pela Equação (3):

$$K = a - 1\,; \tag{3}$$

na qual $a = 2$ indica oo número de níveis. Assim, utiliza-se o nível de significância ajustado $\alpha_{adj} = 0.0166667$.

2.4 Definição do Tamanho Amostral para ANOVA

O experimento realizado possui duas etapas: ANOVA para verificar diferença entre as médias dos algoritmos, seguida de comparções múltiplas one-vs-all (se necessário). Cada uma dessas etapas requer diferentes tamanhos amostrais.

O cálculo do tamanho amostral para a técnica ANOVA pode ser feito iterativamente até encontrar o número n tal que:

$$F_{(1-\alpha)} = F_{\beta;\phi}\,, \tag{4}$$

em que ambas distribuições F têm $(a-1)$ graus de liberdade no numerador e $a(n-1)$ no denomiador. O parâmetro de não-centralidade ϕ é dado por:

$$\phi = \frac{\left(n\sum_{i=1}^{a}\tau_i^2\right)}{\hat{\sigma}^2}\,. \tag{5}$$

O valor de τ para comparações "todos contra um" pode ser obtido a partir da relação (6)

$$\tau = \left(-\frac{(a-1)\delta^\star}{a}, \frac{\delta^\star}{a}, \frac{\delta^\star}{a}, \frac{\delta^\star}{a}\right)\,. \tag{6}$$

O resultado desse teste indica um tamanho amostral $n =$????? observações em cada grupo.

Para calcular o tamanho amostral das comparações múltiplas one-vs-all, serão utilizadas as mesmas relações para a comparação de duas amostras independentes emparelhadas, alterando-se apenas os valores de α para os valores corrigidos α_{adj}, considerando-se as múltiplas hipóteses e $a-1$ comparações. Em comparações desse tipo, a potência é maximizada utilizando um número maior de observações para o grupo de controle, dado pela Equação (7) [Campelo, 2015]:

$$n_0 = n_i\sqrt{K}\,, \tag{7}$$

onde \$K =1 \$ é o número de comparações, n_0 é o tamanho amostral do grupo controle e n_i o número de observações dos demais grupos. Esse valor é calculado conforme a Equação (8):

$$n_i = \left(1+\frac{1}{K}\right)\left(\frac{\hat{\sigma}}{\delta^\star}\right)^2(t_{\alpha_{adj}} + t_\beta)^2\,, \tag{8}$$

em que $t_{\alpha_{adj}}$ e t_β são dependentes de n.

Para solucionar o problema da dependência de n, os termos $t_{\alpha_{adj}}$ e t_β foram substituídos por $z_{\alpha_{adj}}$ e z_β e a equação foi testada iterativamente até a convergência (implementação em anexo no arquivo *calcN.R*).

O resultado obtido indica um tamanho amostral $n_0 =$??? para o grupo controle e $n_i =$??? o segundo.

Serão necessárias ???? coletas para a realização do ANOVA e ???? para as comparações (excetuando-se o grupo controle), o que resultaria em uma sobreamostragem de 2 unidades por grupo. Entretanto, o segundo experimento só é necessário caso o ANOVA indique diferença entre os parâmetros analisados.

2.5 Análise de Covariância (ANCOVA)

A análise de covariância é uma técnica − assim como a chamada "blocagem" ou pareamento (em testes-t) − bastante útil para a melhora da precisão de um experimento [Montgomery, 1984]. Em diversos aspectos é similar à análise de variância (ANOVA), porém permite ter controle sobre a influência do covariante nas variâveis dependentes.

A covariável complementa o controle local e, obviamente, necessita estar correlacionada com a variável de resposta para que se possa fazer uso de tal análise. E quando a análise de variância é realizada com uma ou mais covariáveis, é usual chamar a análise de ANCOVA.

Table 1: ANCOVA como um "modelo ajustado" da ANOVA.

Fontes de variação	$\sum$ **de quadrados**	**Graus de Lib**	**média quadrática**	F_0
regressão	(S_{xy}/S_{xx})	1		
Tratamentos	$SS'_E - SS_E -$ $(S_{xy})^2/S_{xx} - [E_{yy} - (E_{xy})^2/E_{xx}]$	$a-1$	$\frac{SS'_E - SS_E}{a-1}$	$\frac{1}{MS_E}\frac{SS'_E - SS_E}{a-1}$
Erro	$SS_E = E_{yy} - (E_{xy})^2/E_{xx}$	$a(n-1)-1$	$\frac{SS_E}{a(n-1)-1}$	
Total	S_{yy}	$an-1$		

Na ANCOVA a variável dependente é contínua (e.g. tempo, velocidade, etc.), enquanto que a variável independente é normalmente categória (e.g. "masculino/feminino", "fumante/não-fumante" etc.). A análise de variância, por sua vez, poderá converter-se na ANCOVA quando for adicionado um covariante, que consiste em outra variável, que pode ser tanto categórica quanto contínua.

Existem aliás, duas razões principais para sair-se da ANOVA e passar a ser considerada a análise de covarância, i.e., razões técnicas para adicionar um covariante, quais sejam:

- Reduzir fatores de variabilidade inter-grupos, ou alcançar um nível maior de entendimento a partir da variância desconhecida;
- Isolar o efeito que ocorre quando o controle experimental não permite que o experimentador elimine, de maneira razoável, explicações alternativas para uma relação observada entre variáveis independentes e dependentes, o que é chamado de "confusão" *confounding*. Em outras palavras, esta variável é algo que pode estar influenciando o experimento mas que não está, a principio, no modelo original, e é uma potencial fonte de viéses no experimento.

A ANCOVA, permite, portanto, um controle do erro experimental, aumentando sua precisão. Vale ressaltar que a técnica não é restrita apenas à indicação de existência ou não de diferenças entre os níveis avaliados.

Considerando um experimento com um fator e uma covariável, o modelo estatístico da análise de covariância pode ser estabelecido como [Montgomery and Runger, 2011]:

$$y_{ij} = \mu + \tau_i + \beta(x_{ij} - \bar{x_{..}}) + \epsilon_{ij} \, , \tag{9}$$

em que: μ é uma constante, τ_i é o efeito do i-ésimo "tratamento", X_{ij} é o valor observado da covariável e $\bar{x_{..}}$ a média dos valores x_{ij} (da covariável). Além disso, β consiste no coeficiente de regressão linear entre a covariável X e a variável de resposta Y, com $\beta \neq 0$ e ϵ_{ij} é uma componente aleatória de erro, com $i = 1, \ldots, a$ e $j = 1, \ldots, n$. Neste caso a relação deve ser linear. Vale ressaltar que este modelo pressupõe que a variável de resposta e a covariável estão relacionadas linearmente. Para efeito de simplificação, contudo, a análise de covariância pode ser vista como uma análise "ajustada" de variância, e pode ser sumarizada na tabela (**??**):

em que:

$$
\begin{aligned}
S_{yy} &= \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - \frac{y_{..}^2}{an}\,, && (10)\\
S_{xx} &= \sum_{i=1}^{a}\sum_{j=1}^{n} x_{i.}^2 - \frac{x_{..}^2}{an}\,, && (11)\\
S_{xy} &= \sum_{i=1}^{a}\sum_{j=1}^{n} x_{ij}y_{ij} - \frac{(x_{..})(y_{..})}{an}\,, && (12)\\
T_{yy} &= \frac{1}{n}\sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{an}\,, && (13)\\
T_{xx} &= \frac{1}{n}\sum_{i=1}^{a} x_{i.}^2 - \frac{x_{..}^2}{an}\,, && (14)\\
T_{xy} &= \frac{1}{n}\sum_{i=1}^{a} x_{i.}y_{i.} - \frac{(x_{..})(y_{..})}{an}\,, && (15)\\
E_{yy} &= S_{yy} - T_{yy}\,, && (16)\\
E_{xx} &= S_{xx} - T_{xx}\,, && (17)\\
E_{xy} &= S_{xy} - T_{xy}\,. && (18)
\end{aligned}
$$

Assim, rejeita-se a hipótese nula $H_0 : \beta = 0$ se $F_0 > F_{\alpha,1,a(n-1)-1}$.

2.6 Tratamento e Validação dos Dados

Os dados de coletados não apresentaram normalidade quanto à sua distribuição. Para contornar tal constatação, os dados passaram por transformação para uma escala XXXXXXX.

Considerando o experimento realizado, foi criada uma rotina para validação dos dados obtidos e identificação de erros, onde o tempo de execução e valor de qualidade das respostas dos algoritmos na escala (((XXXXXlogarítma))) devem ser maior que 0.

1. LogTempo > 0
2. LogQualidade > 0

Caso os valores de uma execução não atendam essas condições, ela seria descartada. No entanto, nenhuma das amostras apresentou tal problema.

3. Análise Estatística

3.1 Análise de Covariância

3.2 Validação das Premissas

Normalidade

Homocedasticidade

Independência

4. Discussão e Conclusões

Referências

Felipe Campelo. Lecture notes on design and analysis of experiments. https://github.com/fcampelo/Design-and-Analysis-of-Experiments, 2015. Version 2.11, Chapter 7; Creative Commons BY-NC-SA 4.0.

D. C. Montgomery. *Design and analysis of experiments*, volume 7. Wiley New York, 1984.

D. C. Montgomery and G. C. Runger. *Applied Statistics and Probability for Engineers*, volume 5. John Wiley and Sons, 2011.