

Estudo de Caso 03: Comparison of Rising Drilling Configurations

Equipe 04

24 de Junho de 2017

Coordenador: Bernardo Marques

Relator: Danny Tonidandel

Verificador: Gustavo Vieira

Monitor: Alessandro Cardoso

1. Descrição do Problema

Objetiva-se realizar uma comparação entre quatro tipos de tubos de perfuração (*drilling risers*), que consiste em uma espécie de conduíte ou tubo utilizado para servir como passagem temporária para o petróleo extraído em plataformas oceânicas. Mais especificamente, pretende-se comparar o tempo médio até a falha (*mean time to failure* ou *MTTF*) de quatro configurações diferentes (níveis A , B , C , D) de equipamentos, de forma a escolher a que forneça a menor probabilidade de falha, considerando-se um período de 20 anos. Para isto será comparada a configuração padrão (*Riser1*) com as outras três, buscando encontrar qual delas proverá o maior *MTTF*, que pode ser sumarizado na seguinte questão:

“Algum dos risers alternativos é melhor que o padrão?”

A plataforma de testes escolhida pela equipe de engenharia consiste na utilização de um modelo em escala para com um protocolo de tempo acelerado, em que há uma relação direta entre o tempo medido de cada observação (em minutos) e a configuração do sistema real. Todavia, o custo de cada observação é significativamente alto, cerca de US\$10.000 (dez mil dólares). A fim de minimizar os custos com amostras coletadas, tem-se, como alternativa, utilizar dados históricos existentes para a primeira configuração (*Riser1*).

Os parâmetros experimentais desejados são:

- Nível de significância: $\alpha = 0.05$;
- Tamanho de efeito de interesse prático: $\delta_t^* = 0.25$;
- Potência desejada: $(1 - \beta) = \pi \geq 0.85$.

2. Planejamento Experimental

Esta etapa consiste em investigar o comportamento dos níveis de fator a partir de uma análise exploratória, e, posteriormente, aplicar um teste de comparações múltiplas para as médias. A variância do processo é desconhecida e será estimada utilizando-se dados históricos de operação. A variância será considerada como sendo uniforme para todas as configurações.

Os dados experimentais utilizados foram obtidos através de simulação, por meio de um *aplicativo web*. A data de nascimento do segundo membro mais jovem da equipe (15/10/1992) foi o parâmetro utilizado como semente para o gerador de números da simulação. Os dados gerados informam uma tabela com os níveis de cada fator de interesse em uma coluna e os tempos (tomados em escala logarítmica) em coluna respectiva. Por este motivo não necessitam de transformações inversas.

2.1 Análise de Variância

Para avaliar o questionamento acima, será utilizada a ferramenta estatística para análise de variância denominada ANOVA.

A técnica ANOVA compara médias de diferentes amostras para verificar se estas possuem médias iguais ou não. Assim, essa técnica permite que vários grupos sejam comparados simultaneamente [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 a 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.

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

Esta etapa foi considerada como uma primeira verificação ao questionamento deste estudo por apresentar um custo de coleta de amostras menor do que o teste de comparações múltiplas discutido abaixo. Portanto, caso não seja detectada alguma diferença entre os níveis o estudo será concluído por meio do teste ANOVA.

2.2 Comparações Múltiplas

Caso a análise ANOVA identifique a existência de diferenças entre os níveis, deve-se proceder com testes de comparação múltiplas todos contra um, no intuito de identificar qual ou quais níveis apresentam tal diferença.

Desta forma, somente, 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 *Dunnnett*, onde os *Risers* propostos serão confrontados com a configuração padrão *Riser1* para verificar se alguma proposta traria ganho de *MTTF* frente à configuração já estabelecida.

Para que se proceda com o teste de comparações múltiplas, deve-se manter o controle sobre os erros do tipo-*I* em cada comparação. É preciso corrigir os valores de α para cada teste. A escolha para este caso é o método de correção de Bonferroni (2):

$$\alpha_{adj} = \frac{\alpha_{família}}{K}, \quad (2)$$

no qual $K = 3$ consiste no número de comparações a serem feitas que no caso do teste um-contra-todos de *Dunnnett* (eq.(3)):

$$K = a - 1; \quad (3)$$

sendo que $a = 4$ consiste no número de níveis e α (experimental) desejado $\alpha_{família} = 0.05$. Assim, o valor de $\alpha_{adj} = 0.0167$.

2.3 Definição do Tamanho Amostral

Para calcular o tamanho amostral serão utilizadas as mesmas relações para a comparação de duas amostras independentes emparelhadas (“todos contra um”), alterando-se apenas os valores de α para os valores corrigidos α_{adj} , considerando-se as múltiplas hipóteses e $a - 1$ comparações. O tamanho amostral para o teste unilateral para as configurações que serão comparadas ao *riser1* é calculado conforme a equação (4):

$$n_i = \left(1 + \frac{1}{K}\right) \left(\frac{\hat{\sigma}}{\delta^*}\right)^2 (t_{\alpha_{adj}} + t_{\beta})^2, \quad (4)$$

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*). Dessa forma, foi encontrado um primeiro valor para o tamanho amostral $n_i = 58$.

Contudo, para aumentar a potência do procedimento de múltiplas comparações, o tamanho amostral do grupo de controle deve ser calculado conforme a equação (5) [Campelo, 2015]:

$$n_0 = n_i \sqrt{K}, \quad (5)$$

Lembrando que K é o número de comparações, conforme atesta a relação (3), têm-se, portanto $n_i = 101$.

```
## Parsed with column specification:
## cols(
##   Riser = col_character(),
##   LogTTF = col_double()
## )
```

O cálculo do tamanho amostral foi igualmente realizado para a técnica ANOVA: isto pode ser feito a partir de sucessivas iterações em n até que

$$F_{(1-\alpha)} = F_{\beta;\phi}, \quad (6)$$

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

$$\phi = (n \sum_{i=1}^a \tau_i^2) / \hat{\sigma}^2. \quad (7)$$

O valor de τ , no caso de comparações “todos contra um” pode ser obtido a partir da relação (8)

$$\tau = \left(-\frac{(a-1)\delta^*}{a}, \frac{\delta^*}{a}, \frac{\delta^*}{a}, \frac{\delta^*}{a} \right), \quad (8)$$

que produz um valor de tamanho amostral para a análise de variância $n = 60$. Desta forma, tem-se a situação:

- Número de amostras necessárias para o ANOVA: $60 \times 3 + 50 = 230$
- Número de amostras necessárias a comparação todos contra um: $58 * 3 + 91 = 265$

Consequentemente, observa-se que o custo de aplicação a análise de variância será, para este caso, de \$2.300.000 e o das comparações \$2.650.000.

Serão necessárias 60 coletas para a realização do ANOVA e 58 para as comparações (excetuando-se o grupo controle), o que resultaria em uma sobreamostragem de 2 unidades por grupo. Entretanto, caso o ANOVA não forneça como resultado a diferença entre as configurações dos *risers*, seriam economizados cerca de \$350.000 fazendo o ANOVA antecipadamente. Tais informações sumarizadas na tabela 1.

Será seguido o caminho com a análise de variância em virtude da possível economia e do possível menor custo com a amostragem, caso o ANOVA venha a identificar alguma diferença entre as configurações possíveis de *risers*.

	Melhor caso	Pior caso
Com ANOVA	\$2.300.000	\$2.710.000
Sem ANOVA	\$2.650.000	\$2.650.000
Diferença	\$350.000	-\$60,000

Table 1: Análise de custo das opções.

2.4 Tratamento e Validação dos Dados

Considerando o experimento realizado, foi criada uma rotina para validação dos dados obtidos e identificação de erros, onde o tempo de *MTTF* na escala logarítma deve ser maior que 0.

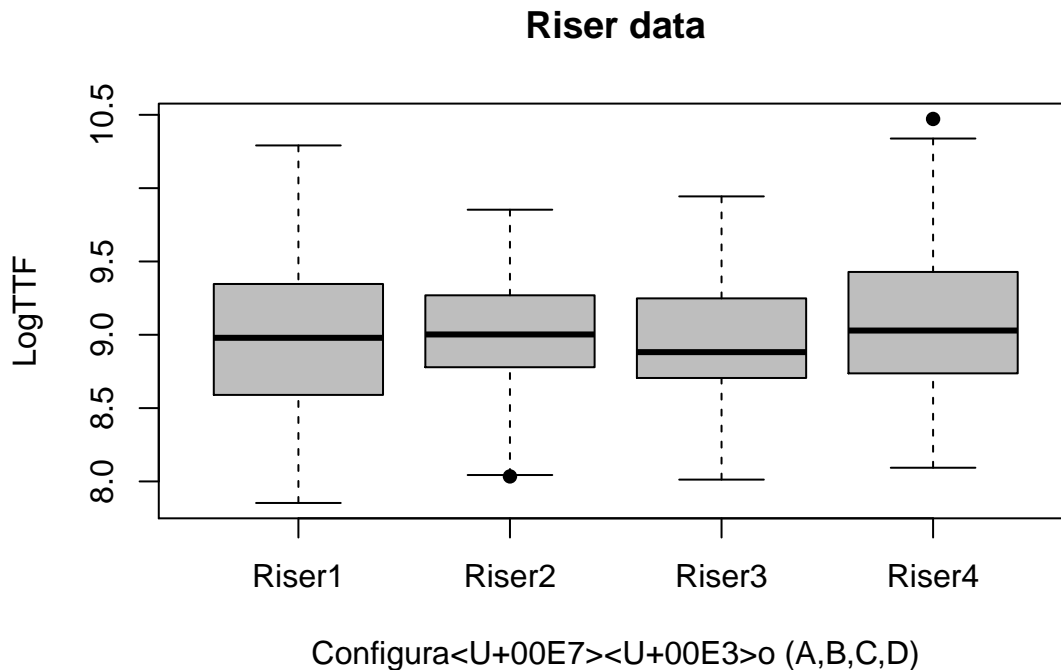
1. $\text{LogTTF} > 0$

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 Variância

```
## Parsed with column specification:
## cols(
##   Riser = col_character(),
##   LogTTF = col_double()
## )
```

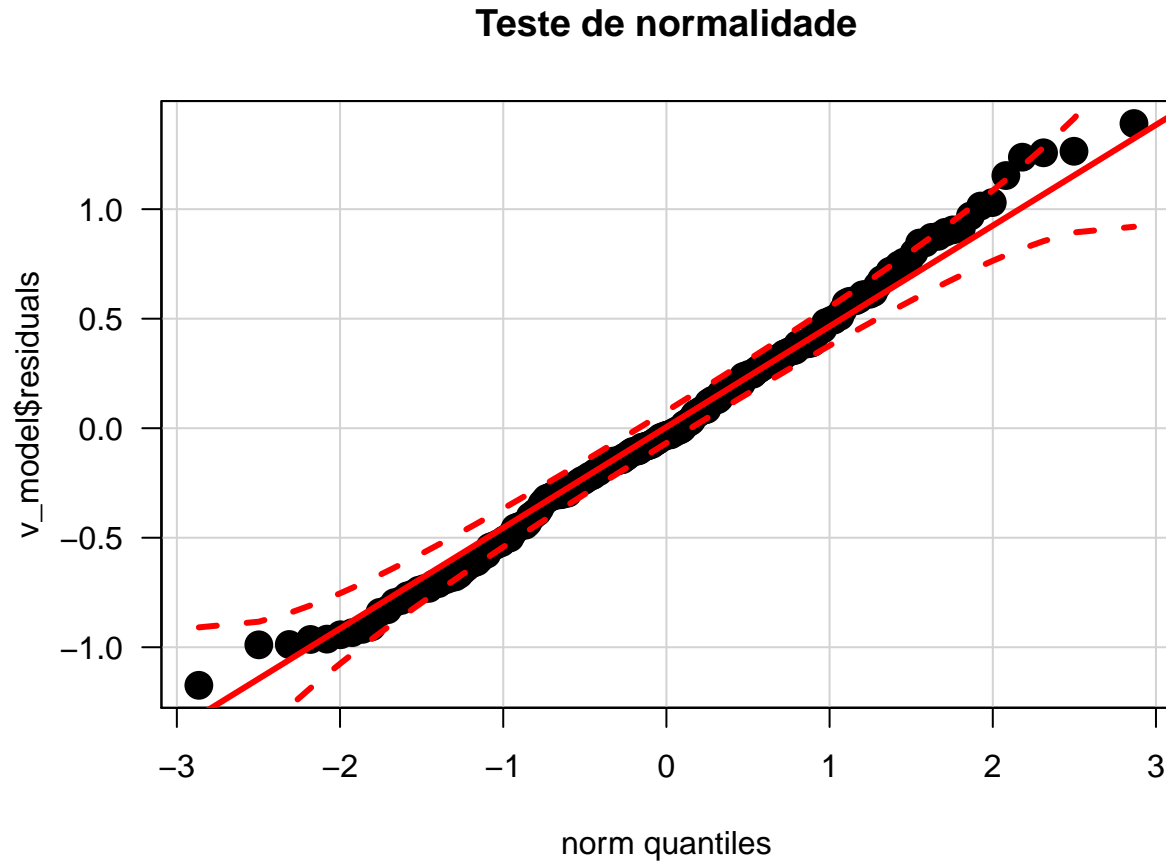


Como pode ser observado, o gráfico anterior não indica visualmente diferença significativa entre os níveis. Corroborar com isto, o $F\text{-valor} = 0.608$ do teste ANOVA indicando pela não rejeição da hipótese de que não há diferença entre os grupos.

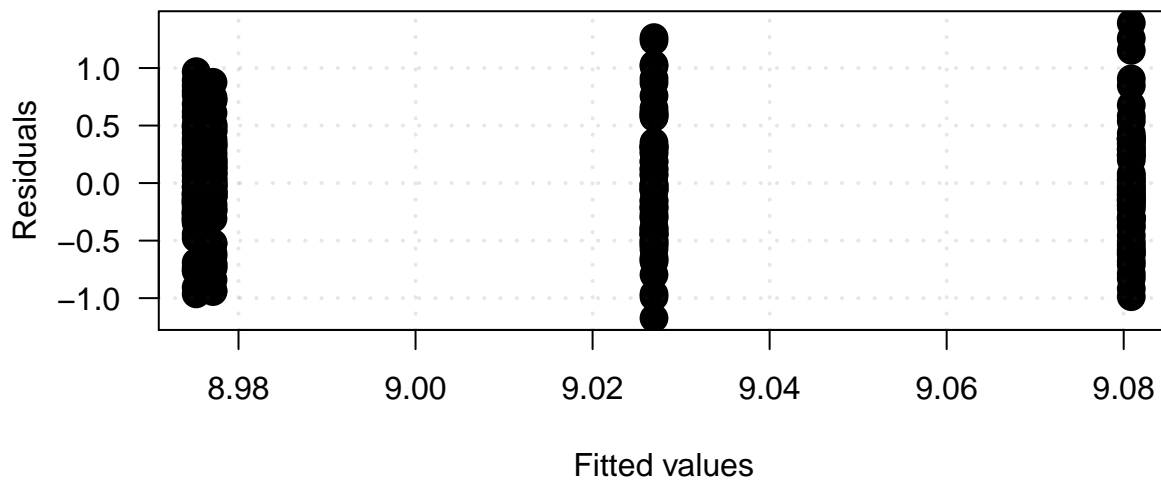
Este resultado também indica que não há necessidade de testes de comparação múltiplas, que poderiam identificar quais os níveis diferentes do grupo de controle (*Riser1*).

3.2 Validação das Premissas

Normalidade

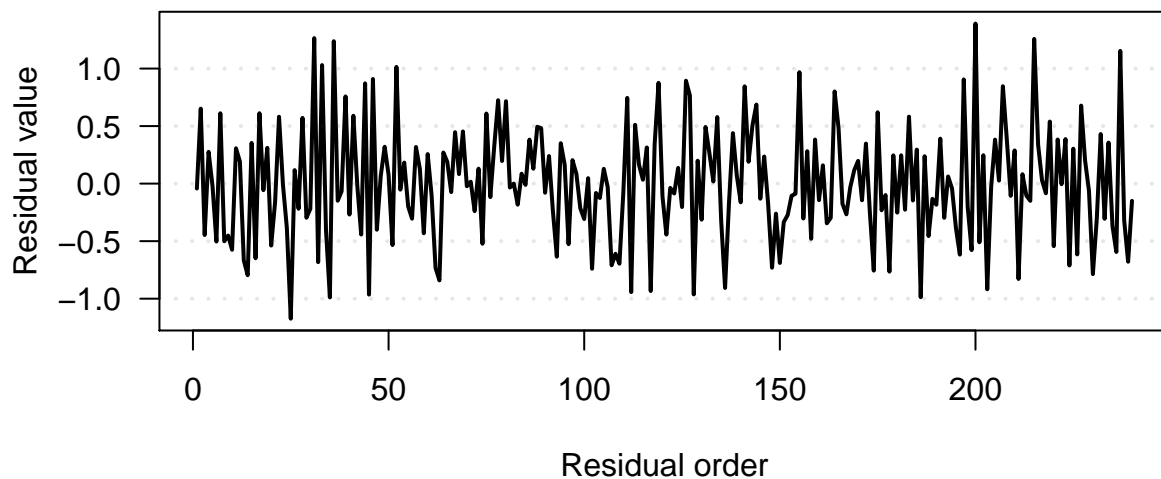


O $p\text{-valor} = 0.299$ encontrado no teste de *Shapiro-Wilk* indica pela rejeição normalidade das amostras. No entanto, o qq plot mostra que as violações de normalidade são relativamente pequenas. A análise ANOVA é robusto a pequenas variações de normalidade [Campelo, 2015]. Desta forma considerou-se a premissa de normalidade atendida.

Homocedasticidade

O teste de igualdade de variância dos resíduos de *Fligner-Killeen* apresentou um $p\text{-valor} = 0.1416$ o que também indica pela rejeição da homocedasticidade das amostras.

Contudo, podemos observar no gráfico acima, uma variância relativamente baixa entre os resíduos das amostras. Novamente com base de que a análise ANOVA é robusto a pequenas variações também de homocedasticidade como indicado por [Campelo, 2015], considerou-se a premissa atendida.

Independência

O plot dos valores ordenados de diferenças de tempo entre os algoritmos não apresenta nenhum indício de dependência temporal dos valores. O teste de autocorrelação serial Durbin-Watson apresenta $p = 0.002$, o que reforça a hipótese de que não há autocorrelação serial entre as amostras.

4. Discussão e Conclusões

Os testes realizados levam às seguintes conclusões:

Variância entre grupos é explicada pela variância intra grupo. Não há indício de diferença significativa entre eles.

Recomenda-se manter riser 1. Custo do experimento é significativo (\$2.300.000), mas previniu um custo potencialmente maior de trocar o Riser.

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