

Estudo de Caso 01: Desempenho de uma nova versão de Software

André Boechat, Mateus Pongelupe, Samuel Leite

24 de Setembro de 2018

Resumo

Este trabalho consiste no primeiro estudo de caso da disciplina de Planejamento e Análise de Experimentos. Nele, foram executados testes para avaliar se o desempenho de uma nova versão de software é superior ao da anterior. A média e a variância do custo de execução foram as variáveis escolhidas para fazer essa medida, sendo que a versão atual do software possui um custo de execução dado por uma distribuição conhecida. **Os resultados alcançados <!-- !TODO - Adicionar resultados ao resumo --> ...**

Planejamento do Experimento

Teste da média

A section detailing the experimental setup. This is the place where you will define your test hypotheses, e.g.:

$$\begin{cases} H_0 : \mu = 10 \\ H_1 : \mu < 10 \end{cases}$$

including the reasons behind your choices of the value for H_0 and the directionality (or not) of H_1 .

This is also the place where you should discuss (whenever necessary) your definitions of minimally relevant effects (δ^*), sample size calculations, choice of power and significance levels, and any other relevant information about specificities in your data collection procedures.

Teste da variância

A section detailing the experimental setup. This is the place where you will define your test hypotheses, e.g.:

$$\begin{cases} H_0 : \mu = 10 \\ H_1 : \mu < 10 \end{cases}$$

including the reasons behind your choices of the value for H_0 and the directionality (or not) of H_1 .

This is also the place where you should discuss (whenever necessary) your definitions of minimally relevant effects (δ^*), sample size calculations, choice of power and significance levels, and any other relevant information about specificities in your data collection procedures.

Coleta de Dados

A coleta de dados foi simulada a partir da rotina sugerida no caso de uso, com uma pequena modificação: uma *seed* foi definida para a execução do programa, de forma a garantir sua reprodutibilidade.

```

# Loading required package
library(ExpDE)
mre <- list(name = "recombination_bin", cr = 0.9)
mmu <- list(name = "mutation_rand", f = 2)
mpo <- 100
mse <- list(name = "selection_standard")
mst <- list(names = "stop_maxeval", maxevals = 10000)
mpr <- list(name = "sphere", xmin = -seq(1, 20), xmax = 20 + 5 * seq(5, 24))

# Setting seed so the program can be reproduced.
set.seed(1998)

# One sample
ExpDE(mpo, mmu, mre, mse, mst, mpr,
      showpars = list(show.iters = "none"))$Fbest

```

Em nossos experimentos, precisaremos coletar um número arbitrário N de amostras. Portanto, a partir das rotinas acima, foram criadas duas funções para essa coleta:

- *generate_sample* : Coleta uma única amostra.
- *generate_n_samples* : Coleta n amostras no formato de um *data.frame*.

Segue abaixo a codificação dessas funções, bem como um exemplo da chamada de *generate_n_samples* para $N = 10$:

```

#Generates one sample
generate_sample <- function() {
  return(ExpDE(mpo, mmu, mre, mse, mst, mpr,
               showpars = list(show.iters = "none"))$Fbest);
}

#Generates n samples on the data.frame format
generate_n_samples <- function(n) {
  cost <- replicate(n, generate_sample())
  return(data.frame(cost))
}

#Example for N=10
generate_n_samples(10)

```

Análise Exploratória dos Dados

The first step is to load and preprocess the data. For instance,

```

N <- 30
samples <- generate_n_samples(N)

```

To get an initial feel for the relationships between the relevant variables of your experiment it is frequently interesting to perform some preliminary (exploratory) analysis. This is frequently referred to as *getting a feel* of your data, and can suggest procedures (such as outlier investigation or data transformations) to experienced experimenters.

```

library(cowplot, warn.conflicts = FALSE)

theme_set(theme_cowplot(font_size=12))

```

```

plot.hist <- ggplot(samples, aes(x=cost)) +
  geom_histogram(colour="black", fill="white") + background_grid(major = 'xy')

plot.dens <- ggplot(samples, aes(x=cost)) +
  geom_density(alpha=.2, fill="#FF6666") +
  background_grid(major = 'xy')

plot_grid(plot.hist, plot.dens, labels = c('A','B'), ncol = 2)

```

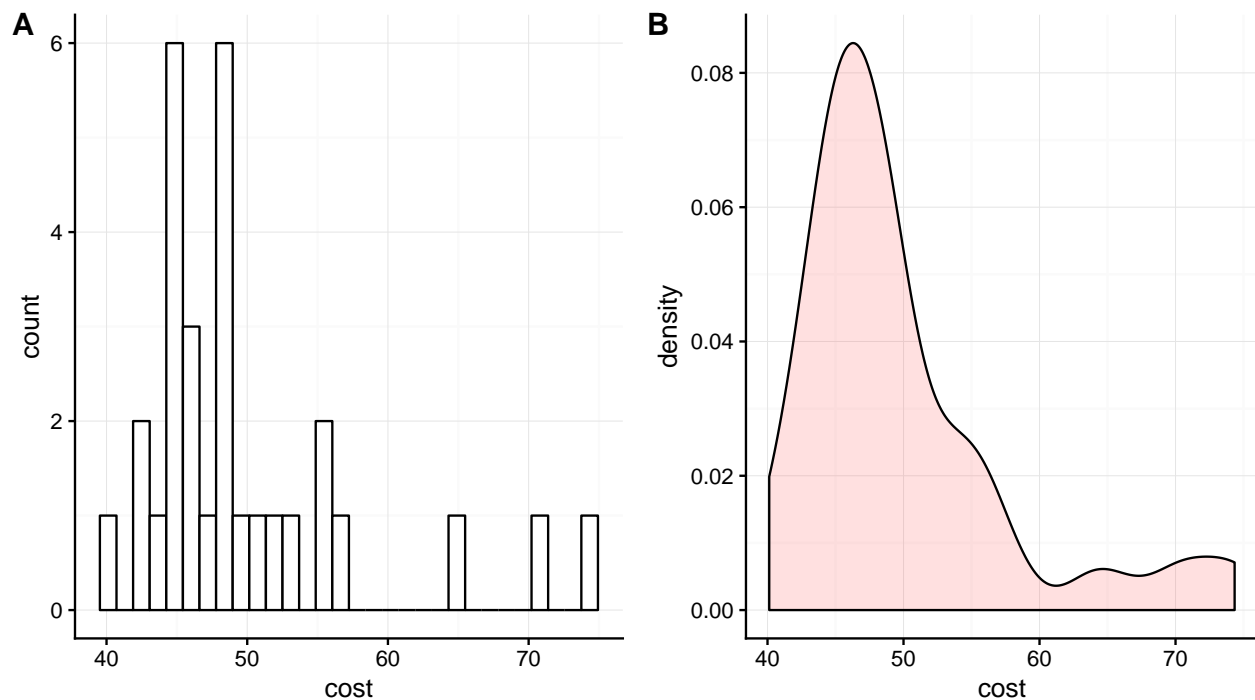


Figure 1: Exploring the effect of car transmission on mpg values

Your preliminary analysis should be described together with the plots. In this example, two facts are immediately clear from the plots: first, **mpg** tends to correlate well with many of the other variables, most intensely with **drat** (positively) and **wt** (negatively). It is also clear that many of the variables are highly correlated (e.g., **wt** and **disp**). Second, it seems like manual transmission models present larger values of **mpg** than the automatic ones. In the next section a linear model will be fit to the data in order to investigate the significance and magnitude of this possible effect.

Statistical Analysis

Your statistical analysis should come here. This is the place where you should fit your statistical model, get the results of your significance test, your effect size estimates and confidence intervals.

```

model<-aov(mpg~am*disp,data=mtcars)
summary(model)

```

##		Df	Sum Sq	Mean Sq	F value	Pr(>F)
##	am	1	405.2	405.2	47.948	1.58e-07 ***
##	disp	1	420.6	420.6	49.778	1.13e-07 ***
##	am:disp	1	63.7	63.7	7.537	0.0104 *

```
## Residuals    28  236.6    8.4
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Checking Model Assumptions

The assumptions of your test should also be validated, and possible effects of violations should also be explored.

```
par(mfrow=c(2,2), mai=.3*c(1,1,1,1))
plot(model,pch=16,lty=1,lwd=2)
```

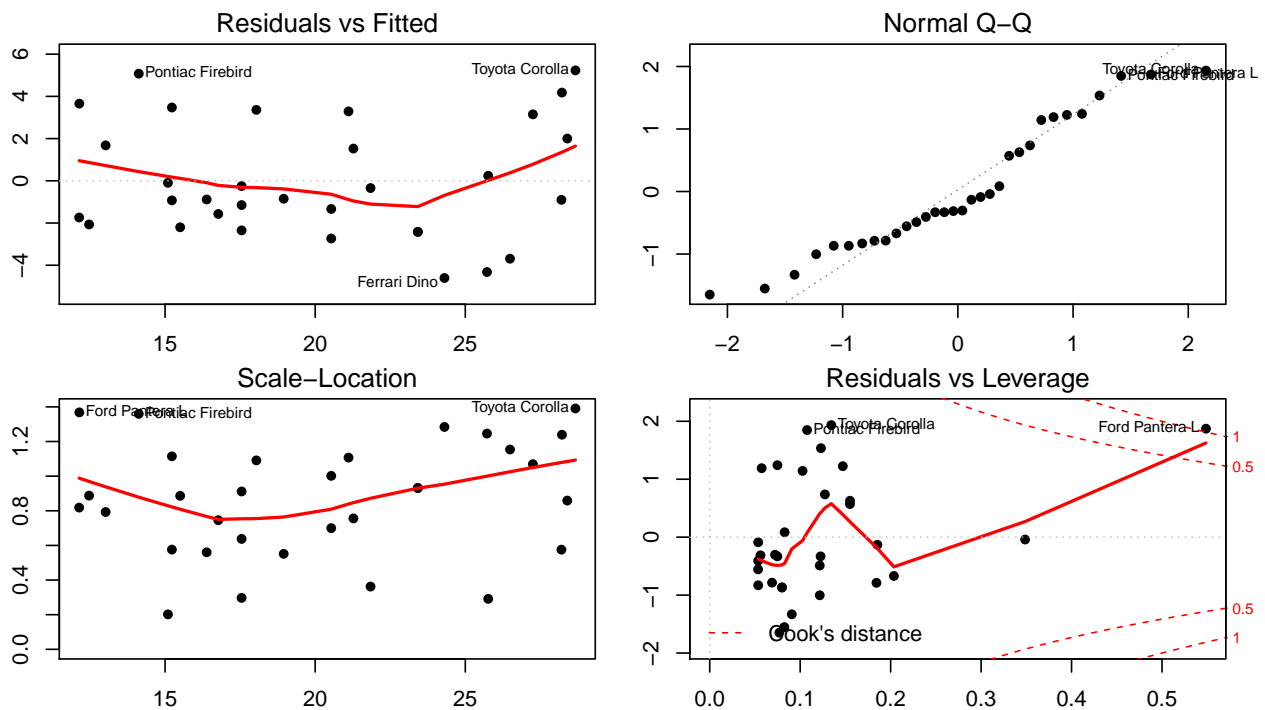


Figure 2: Residual plots for the anova model

Conclusions and Recommendations

The discussion of your results, and the scientific/technical meaning of the effects detected, should be placed here. Always be sure to tie your results back to the original question of interest!