

Estudo de caso: Grupo D 3

Gilmar and Maressa Nunes R. Tavares and Victor

3 de Setembro, 2019

Sumário

O presente trabalho tem como objetivo delinear e executar testes estatísticos para avaliar uma nova versão de um software, em relação aos resultados obtidos na versão anterior. Tendo em vista que a última versão possui uma distribuição de custos com média $\mu = 50$ e variância $\sigma = 100$, dados da população, objetiva-se verificar se a nova versão apresenta resultados melhores para tais características. Para tanto, utilizou-se o teste z com nível de significância $\alpha = 0,01$ e $\alpha = 0,05$, para os testes de média e variância, respectivamente. Após os testes verificou-se que....

Experimental design

Teste da média

Nesta seção são apresentados os detalhes do delineamento dos testes que foram executados para comparar os resultados das duas versões do software. Inicialmente, o teste foi executado para as médias do custo computacional, assim, para verificar se a nova versão é melhor que a anterior, formulou-se as seguintes hipóteses:

$$\begin{cases} H_0 : \mu = 50 \\ H_1 : \mu < 50 \end{cases}$$

Como a média da população para a primeira versão é $\mu = 50$, considourou-se como hipótese nula (H_0) a ausência de melhoria do software, isto é, a segunda versão apresenta a mesma performace da versão anterior, com média igual $\mu = 50$. Por outro lado, a hipótese alternativa considera que houve melhorias entre as versões, portanto, a média é menor que 50 (H_1).

Descrição do Problema

Para a versão atual de um dado sistema, sabe-se que sua distribuição de custos de execução possui média populacional de $\mu = 50$ e variância $\sigma^2 = 100$. Uma nova versão desse software foi desenvolvida, portanto uma análise estatística deve ser feita para investigar os ganhos de desempenho obtidos em relação à versão atual. Espera-se que sejam testados a média e variância dos custos de execução

Para simular a geração de dados da nova versão, a biblioteca *ExpDE* [1] foi utilizada. Ela é declarada da seguinte forma:

```
# Set-up the data generating procedure
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))

set.seed(1234) # to generate always the same results

# define functions for data generation
get.single.sample <- function(mpo, mmu, mre, mse, mst, mpr){
  generator <- ExpDE(mpo, mmu, mre, mse, mst, mpr, showpars = list(show.its = "none"))
  return(generator$Fbest)
```

```

}

get.n.samples <- function(mpo, mmu, mre, mse, mst, mpr, N){
  my.sample <- numeric(N)
  for (i in seq(N)){
    my.sample[i] <- get.single.sample(mpo, mmu, mre, mse, mst, mpr)
  }
  return(my.sample)
}

```

As funções `get.single.sample` e `get.n.samples` foram criadas para facilitar o entendimento da função de geração de dados.

Planejamento do Experimento

Teste do custo médio

Para este teste, são estabelecidos os seguintes objetivos:

- Nível de significância desejado $\alpha = 0.01$. Logo, o nível de confiança desejado é $1 - \alpha = 0.99$
- Efeito relevante mínimo de $\delta^* = 4$
- Potência desejada $\pi = 1 - \beta = 0.8$

Como estamos interessados em saber se existem ganhos em termos do custo médio, e dado que a média populacional da versão atual é $\mu_0 = 50$, define-se a seguinte hipótese nula e alternativa:

$$\begin{cases} H_0 : \mu = 50 \\ H_1 : \mu < 50 \end{cases}$$

Teste da variância do custo

Para este teste, são estabelecidos os seguintes objetivos:

- Nível de significância desejado $\alpha = 0.01$. Logo, o nível de confiança desejado é $1 - \alpha = 0.99$
- Usar as mesmas observações coletadas para o teste da média.

Como estamos interessados em saber se existem ganhos em termos de variância média, e dado que a variância populacional da versão atual é $\sigma^2 = 100$, define-se a seguinte hipótese nula e alternativa:

$$\begin{cases} H_0 : \sigma^2 = 100 \\ H_1 : \sigma^2 < 100 \end{cases}$$

Análise Exploratória dos Dados

Análise Estatística

Teste sobre a média do custo

Cálculo do tamanho amostral

Baseado nas informações preliminares do problema, $\sigma^2 = 100$, $\delta^* = 4$ e $\pi = 0.8$, e dado que estamos considerando uma hipótese alternativa unilateral para a média amostral, o cálculo do tamanho amostral pode ser estimado com a função `power.t.test`:

```

# define current system parameters
current_mu <- 50
current_var <- 100

```

```

# define mean cost test parameters
sig_level_mean <- 0.01
delta <- 4
beta <- 0.2
pi <- 1 - beta
ci_mean <- 1 - sig_level_mean

# use the function invisible() to suppress the function console output
invisible(sample_size_calc <- power.t.test(delta = delta,
      sd = sqrt(current_var),
      sig.level = sig_level_mean,
      power = pi,
      alternative = "one.sided",
      type = "one.sample"))

# round to the next integer
N <- ceiling(sample_size_calc$n)

```

Resultando em um tamanho amostral de:

```
## [1] N = 66
```

Teste sobre a variância do custo

Validação das Premissas

Discussão e Conclusões

Divisão das Atividades

Victor - Reporter Maressa - Coordenadora Gilmar - Verificador e Monitor

Referências

[1] M. B. Felipe Campelo, “CRAN - package expde - modular differential evolution for experimenting with operators.” <https://cran.r-project.org/web/packages/ExpDE/index.html>, Jan-2018.