

Estudo de caso: Grupo D 3

Gilmar Pereira, Maressa Tavares e Victor Ruela

30 de Setembro, 2019

1 Resumo

- 0) separar entre homens e mulheres em cada ano. Realizar um teste de hipóteses para cada sexo, comparando os dois anos
- 1) H_0 : não há diferença entre as médias H_1 : as médias são diferentes
- 2) Validar as premissas Normalidade Igualdade de variância (se diferente, o teste T não poderá ser usado) Independência
- 3) Tamanho de efeito: 5 (de acordo com tamanho das faixas, primeiro resultado do google)
Número de amostras conhecido: calcular a potência e estimar nível de significância adequado ($\alpha = 0.05$)
- 4) Como melhorar: verificar aumento da potência, levando a aumento da quantidade de amostras
- 5) Discutir sobre diferenças na potência entre homens e mulheres (menos mulheres na engenharia, poucas amostras)

2 Summary

O presente trabalho realizou o delineamento e executou os testes estatísticos para avaliar as diferenças no IMC médio entre duas populações de estudantes de pós-graduação em Engenharia elétrica, nos semestres de 2016-2 e 2017-2. As sub-populações masculina e feminina foram analisadas separadamente. Após os testes verificou-se que ...

3 Planejamento do experimento

3.1 Objetivo do experimento

O objetivo é estudar as diferenças entre o IMC médio entre duas turmas de estudantes de pós-graduação em Engenharia elétrica na UFMG, para dados coletados nos semestres de 2016-2 e 2017-2. A análise será dividida entre as duas sub-populações (homens e mulheres), uma vez que é esperado diferenças no IMC e também no tamanho amostral. Portanto, para verificar estas diferenças, as seguintes hipóteses serão testadas:

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$

Onde μ_1 e μ_2 são as turmas de 2016-2 e 2017-2, respectivamente. A mesma hipótese será testada separadamente para cada sub-população.

Os testes serão realizados para um tamanho de efeito $\delta^* = 5$, visto que essa é a diferença de valor entre os níveis de classificação IMC encontrados na literatura [1].

3.2 Análise Exploratória dos Dados

Antes de iniciar a análise exploratória dos dados, é preciso realizar um breve pré-processamento dos dados, visto que os dados de cada ano possuem formatos ligeiramente diferentes. Por exemplo, para os dados de 2016-2, é necessário remover as linhas referentes aos alunos de graduação. Ambos os arquivos foram padronizados para possuir o mesmo nome de coluna, e uma nova coluna com o cálculo do IMC foi adicionada

em cada. Novos datasets também foram criados contendo somente os dados de cada sexo e todos os dados, para auxiliar nas análises.

Para ter uma ideia inicial dos dados, suas estatísticas básicas são calculadas:

```
summary(data.2016.2)
```

```
## Gender      imc
## F: 7   Min.   :17.58
## M:21   1st Qu.:21.27
##        Median :23.52
##        Mean   :23.97
##        3rd Qu.:25.29
##        Max.   :37.55
```

```
summary(data.2017.2)
```

```
## Gender      imc
## F: 4   Min.   :17.36
## M:21   1st Qu.:20.83
##        Median :23.26
##        Mean   :23.35
##        3rd Qu.:25.18
##        Max.   :30.42
```

É possível fazer as seguintes observações que a quantidade de amostras do sexo feminino é bem menor do que a do masculino, o que pode afetar a potência dos testes a serem executados.

Para melhor visualização, um gráfico boxplot é gerado na figura abaixo. Nele é possível ver que há alguns outliers nos valores de IMC em alguns anos, que serão removidos posteriormente.

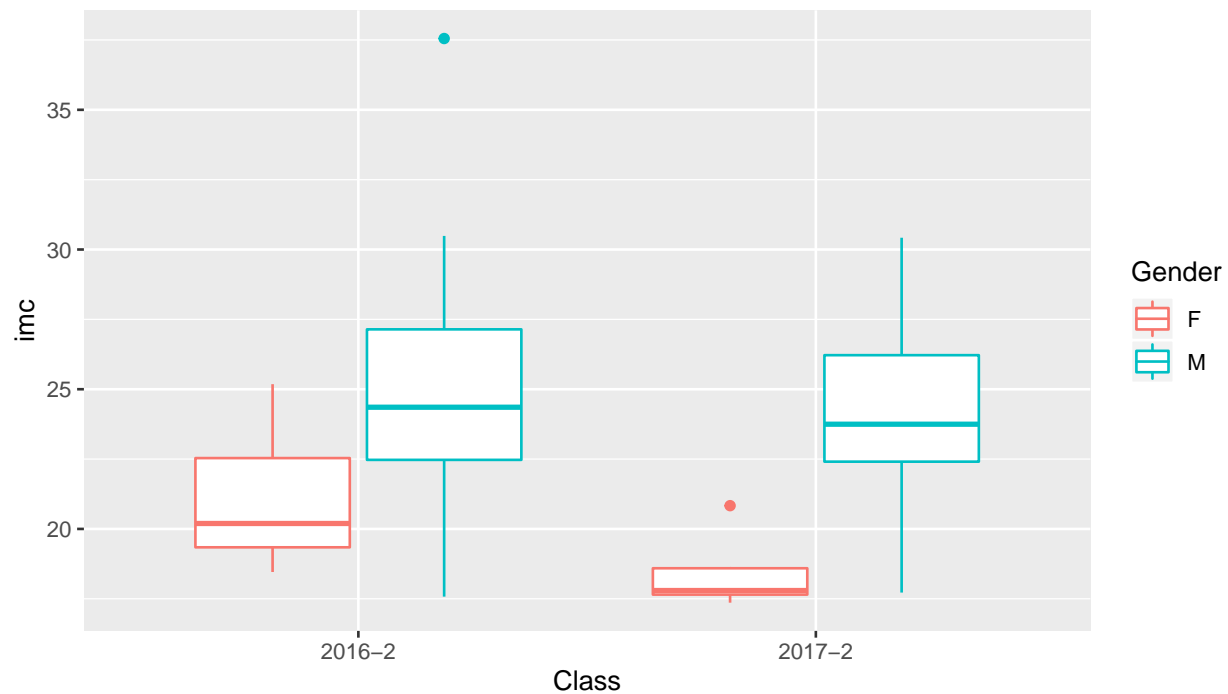


Figure 1: Boxplot dos dados

3.2.1 Estimativa da significância e potência

Um nível de significância $\alpha = 0.05$ é estabelecido para uma potência de $\pi = 1 - \beta = 0.8$. Dado que o número de amostras é conhecido para cada turma e gênero, avalia-se a seguir se a quantidade de amostras é suficiente para atender os parâmetros especificados.

```
(sig.est.males <- pwr.t.test(d = d.cohen.males,
  n = n.males,
  sig.level = NULL,
  type = "two.sample",
  alternative = "two.sided",
  power = pi))
```

```
##
##      Two-sample t test power calculation
##
##              n = 21
##              d = 1.280529
##      sig.level = 0.002185197
##              power = 0.8
##      alternative = two.sided
##
## NOTE: n is number in *each* group
```

```
(sig.est.females <- pwr.t.test(d = d.cohen.females,
  n = n.females,
  sig.level = NULL,
  type = "two.sample",
  alternative = "two.sided",
  power = pi))
```

```
##
##      Two-sample t test power calculation
##
##              n = 4
##              d = 2.437794
##      sig.level = 0.04525425
##              power = 0.8
##      alternative = two.sided
##
## NOTE: n is number in *each* group
```

Portanto, é possível ver que há amostras suficientes para a realização dos testes de cada sub-população. É importante notar que, dado que há mais amostras de estudantes do sexo masculino, a significância mínima estimada foi bem menor. Um outro resultado interessante, causado pela diferença do número de amostras, é que a potência do teste para alunos do sexo feminino será significativamente menor em relação à do sexo masculino. Assumindo $\alpha = 0.05$, a potência é estimada a seguir.

```
(pwr.est.females <- pwr.t.test(d = d.cohen.females,
  n = n.females,
  sig.level = 0.05,
  type = "two.sample",
  alternative = "two.sided",
  power = NULL))
```

```
##
##      Two-sample t test power calculation
```

```
##
##           n = 4
##           d = 2.437794
##       sig.level = 0.05
##           power = 0.8177638
##       alternative = two.sided
##
## NOTE: n is number in *each* group
(pwr.est.males <- pwr.t.test(d = d.cohen.males,
  n = n.males,
  sig.level = 0.05,
  type = "two.sample",
  alternative = "two.sided",
  power = NULL))
```

```
##
##       Two-sample t test power calculation
##
##           n = 21
##           d = 1.280529
##       sig.level = 0.05
##           power = 0.9815966
##       alternative = two.sided
##
## NOTE: n is number in *each* group
```

Conforme esperado, a potência do teste masculino é 17% maior que o feminino, indicando uma maior possibilidade de erros do tipo 2.

4 Resultados

4.1 Validação das premissas

Para realizar as inferências estatísticas sobre o IMC das duas populações é necessário validar as premissas antes de executar o teste. Neste caso, como as variâncias das duas populações é desconhecida, utiliza-se a distribuição t para o teste de hipóteses e para os intervalos de confiança [2]. A seguir são apresentados os testes realizados para validar as premissas exigidas pelo teste t. Para facilitar as análises optou-se por separar o grupo na população feminina e masculina.

4.1.1 Subpopulação Feminina

1 - Normalidade:

Para avaliar a normalidade dos dados das duas subpopulações, utilizou-se o gráfico quantil-quantil e o teste de Shapiro Wilk com $\alpha = 0.05$, como apresnetado a seguir.

```
## [1] 6 1
```

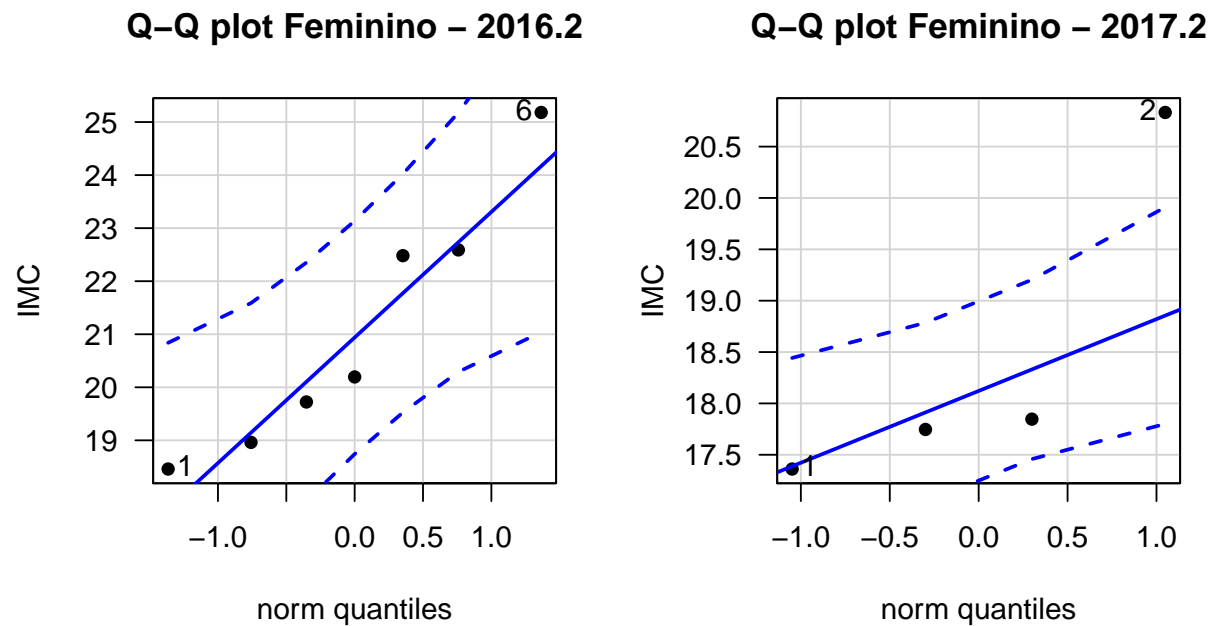


Figure 2: Gráfico quantil-quantil das populações - Feminino

```
## [1] 2 1
shapiro.test(data.2016.2.Females$imc)
```

```
##
## Shapiro-Wilk normality test
##
## data: data.2016.2.Females$imc
## W = 0.91974, p-value = 0.4674
```

```
shapiro.test(data.2017.2.Females$imc)
```

```
##
## Shapiro-Wilk normality test
##
## data: data.2017.2.Females$imc
## W = 0.7475, p-value = 0.03659
```

Pela análise do teste de Shapiro-Wilk e do gráfico concluímos que não há evidências para rejeitar a normalidade dos dados de 2016.2, pois o p-valor do teste de Shapiro Wilk foi maior que 0.05. Por outro lado, os dados de 2017.2 apresentam um ponto fora dos limites e o p-valor foi inferior a 0.05, por isso, há evidências para rejeitar a normalidade dos dados.

Diante disso, é necessário fazer uma transformação nos dados de 2017.2 para tentar normalizar os dados, a fim de possibilitar a realização dos testes estatísticos.

```
qqnorm(log(data.2017.2.Females$imc), pch = 16, las = 1)
qqline(log(data.2017.2.Females$imc))
```

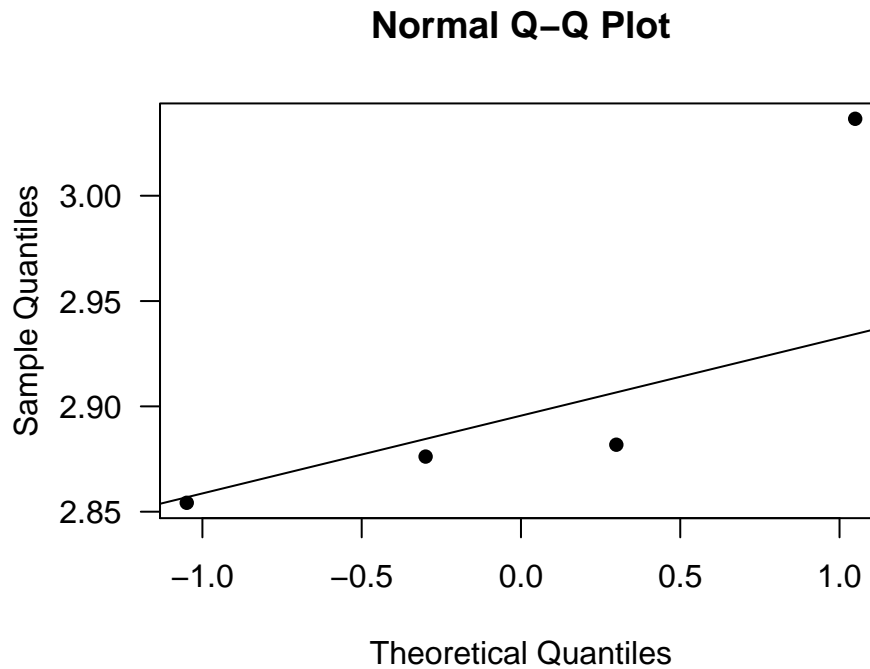


Figure 3: Gráfico quantil-quantil para transformação logarítmica dos dados

```
shapiro.test(log(data.2016.2.Females$imc))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  log(data.2016.2.Females$imc)
## W = 0.93097, p-value = 0.5592
```

Embora o gráfico dos quantis não deixe claro a normalidade o teste de Shapiro mostra que, após o ajuste logarítmico dos dados, não há evidências para rejeitar a normalidade.

2- Igualdade de Variâncias:

A segunda premissa a ser avaliada é a igualdade de variâncias das duas populações, homocedasticidade. Para tanto, utilizou-se o teste F com a função `var.test` e $\alpha = 0.05$, que tem como hipótese nula a igualdade da variância das duas amostras.

```
var.test(data.2016.2.Females$imc,data.2017.2.Females$imc)
```

```
##
##  F test to compare two variances
##
## data:  data.2016.2.Females$imc and data.2017.2.Females$imc
## F = 2.2688, num df = 6, denom df = 3, p-value = 0.5353
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##   0.1539783 14.9715344
## sample estimates:
## ratio of variances
```

```
##                2.268827
```

Pelo teste F conclui-se que não há evidências para rejeitar a hipótese nula de igualdade das variâncias das duas populações, portanto, as amostras são consideradas homocedásticas.

3 - Independência

Considerando que as populações referem-se a dois grupos distintos da pós-graduação, sabe-se a priori, que as amostras são independentes. Para concluir com precisão em relação à independência realizou-se o teste Qui-quadrado com $\alpha = 0.05$.

```
data.chi.Female <- c(data.2016.2.Females$imc, data.2017.2.Females$imc)
chisq.test(data.chi.Female)
```

```
##
## Chi-squared test for given probabilities
##
## data:  data.chi.Female
## X-squared = 3.0047, df = 10, p-value = 0.9813
```

Como era esperado, o teste Qui-quadrado reafirma a independência entre as duas amostras, como o p-valor = 0.9813.

Diante dessas análises, percebemos que provavelmente o tamanho pequeno das amostras de IMC feminino pode ter influenciado nas conclusões em relação às premissas. Portanto, esse fator pode influenciar também nos testes de hipóteses que serão apresentados nas próximas seções.

4.1.2 Subpopulação Masculina

Assim como na validação das premissas da subpopulação feminina, foram realizados os testes para a subpopulação masculina, como descrito a seguir.

1 - Normalidade:

Para validar a normalidade dos dados utilizou-se o gráfico quantil-quantil e o teste de Shapiro Wilk com $\alpha = 0.05$.

```
## [1]  4 12
```

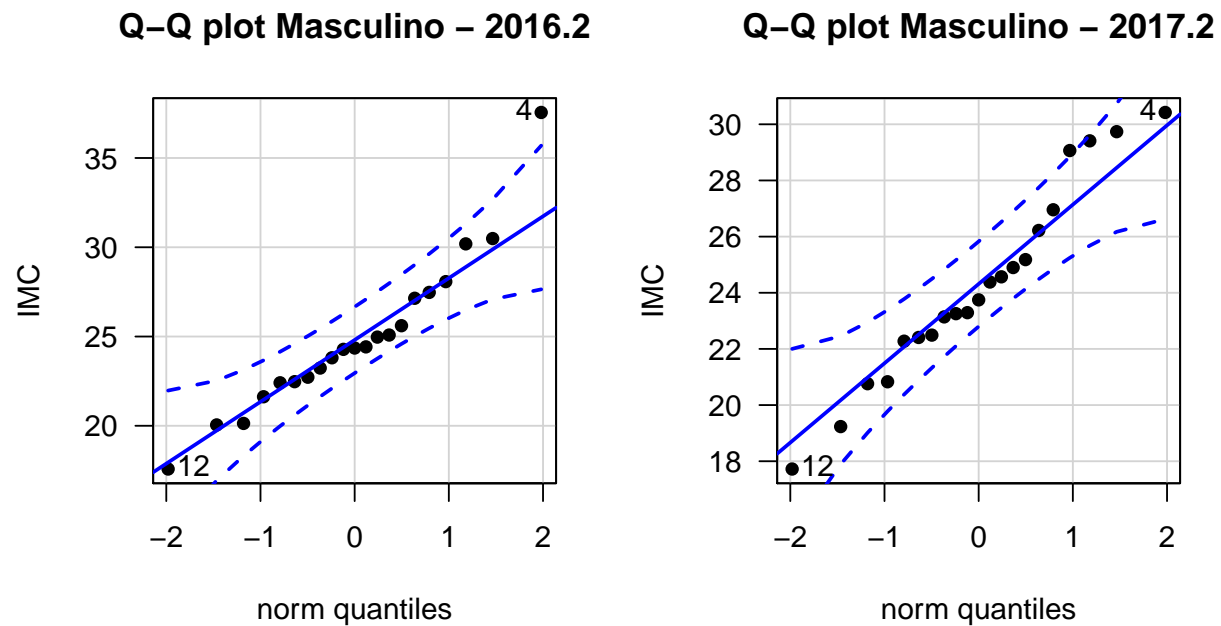


Figure 4: Gráfico quantil-quantil das populações - Feminino

```
## [1] 12 4
shapiro.test(data.2016.2.Males$imc)

##
## Shapiro-Wilk normality test
##
## data: data.2016.2.Males$imc
## W = 0.92833, p-value = 0.1275
shapiro.test(data.2017.2.Males$imc)

##
## Shapiro-Wilk normality test
##
## data: data.2017.2.Males$imc
## W = 0.96494, p-value = 0.6206
```

Pela análise do teste de Shapiro-Wilk e dos gráficos concluímos que não há evidências para rejeitar a hipótese nula para as duas amostras, pois em ambas o p-valor do teste foi superior a 0.05. Portanto, para a subpopulação masculina dos dois semestres os dados estão normalmente distribuídos.

2- Igualdade de Variâncias:

Para validar a igualdade de variâncias das duas amostras utilizou-se o teste F com $\alpha = 0.05$. A hipótese nula considera a igualdade da variância das duas amostras.

```
var.test(data.2016.2.Males$imc,data.2017.2.Males$imc)

##
## F test to compare two variances
```



```
##
## data: data.2016.2.Males$imc and data.2017.2.Males$imc
## F = 1.5839, num df = 20, denom df = 20, p-value = 0.3119
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
## 0.6426853 3.9034665
## sample estimates:
## ratio of variances
## 1.583888
```

Pelo teste F conclui-se que as populações são homocedásticas, pois o p-valor foi superior a 0.05, logo, não há evidências para rejeitar a hipótese nula.

3 - Independência

Assim como na subpopulação feminina, sabe-se a priori, que as amostras são independentes. Para concluir com precisão em relação à independência realizou-se o teste Qui-quadrado com $\alpha = 0.05$.

```
data.chi.male <- c(data.2016.2.Males$imc,data.2017.2.Males$imc)
chisq.test(data.chi.male)
```

```
##
## Chi-squared test for given probabilities
##
## data: data.chi.male
## X-squared = 24.96, df = 41, p-value = 0.9772
```

Como era esperado, o teste Qui-quadrado reafirma a independência entre as duas amostras, como o p-valor = 0.9772.

Ao contrário do que aconteceu com as amostras da subpopulação feminina, nesse caso o tamanho das amostras foram maiores, e por isso, não houve nenhum problema em relação às premissas do teste.

5 Resultados

Nesta seção são apresentados os resultados do teste para as hipóteses estabelecidas anteriormente com $\alpha = 0.05$. Os resultados estão subdivididos entre a população feminina e masculina devido às diferenças de IMC que existem entre eles.

5.1 Teste de Hipóteses Subpopulação Feminina

```
#teste para custo médio:
(mean.t.teste <- t.test(data.2016.2.Females,
                        data.2017.2.Females,
                        var.equal = TRUE))

##
## Two Sample t-test
##
## data: data.2016.2.Females and data.2017.2.Females
## t = 1.9308, df = 9, p-value = 0.08556
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.4527037 5.7283762
## sample estimates:
## mean of x mean of y
## 21.08443 18.44660
```

5.2 Teste de Hipóteses Subpopulação Masculina

```
#teste para custo médio:
(mean.t.teste <- t.test(data.2016.2.Males,
                        data.2017.2.Males,
                        var.equal = TRUE))

##
## Two Sample t-test
##
## data: data.2016.2.Males and data.2017.2.Males
## t = 0.53979, df = 40, p-value = 0.5923
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -1.784943 3.085836
## sample estimates:
## mean of x mean of y
## 24.93595 24.28551

pwr.t2n.test(n1 = 7, n2= 4, d = 0.3, sig.level = 0.05, power = NULL,
             alternative = "two.sided")

##
## t test power calculation
##
## n1 = 7
## n2 = 4
## d = 0.3
## sig.level = 0.05
## power = 0.07141217
## alternative = two.sided

pwr.t2n.test(n1 = 21, n2= 21, d = 0.8, sig.level = 0.05, power = NULL,
             alternative = "two.sided")

##
## t test power calculation
##
## n1 = 21
## n2 = 21
## d = 0.8
## sig.level = 0.05
## power = 0.7155555
## alternative = two.sided
```

6 Discussão e Conclusões

Neste trabalho foi feito um estudo estatístico ...

7 Divisão das Atividades

- Victor - Verificador e Monitor
- Maressa - Relatora
- Gilmar - Coordenador

Referências

- [1] *Tabela de imc*. <http://www.calculoimc.com.br/tabela-de-imc/>.
- [2] D. C. Montgomery and G. C. Runger, *Applied statistics and probability for engineers, (with cd)*. John Wiley & Sons, 2007.