Análise da capacidade de medidores e de instrumentos de medidas

CE219 - Controle Estatístico de Qualidade

Prof. Cesar Taconeli
taconeli@ufpr.br

Prof. Walmes Zeviani
walmes@ufpr.br

Laboratório de Estatística e Geoinformação
Departamento de Estatística
Universidade Federal do Paraná

17 de maio, 2019

Medir para aprimorar

"If you cannot measure it, you cannot improve it."

– William Thomson

Fontes de variabilidade de medida

- Todos os métodos estudados até o momento, visando a qualidade de processos, baseiam-se em **medições de características** da qualidade avaliadas em unidades ou itens produzidos.
- Tais medições, naturalmente (e inevitavelmente), estão sujeitas a **variações atribuíveis** a diferentes fontes, sendo classificadas em duas categorias principais:
  - **Variabilidade real da característica**, i.e. inerente ao processo produtivo.
  - **Variabilidade inerente ao sistema de medida** que está relacionada aos instrumentos/procedimentos de medida/avaliação.

Objetivos do estudo de capacidade de medidor

Os principais objetivos dos estudos sobre a **capacidade** de sistemas de medidas são os seguintes:

- Isolar os **componentes da variabilidade** no sistema de medida.
- Determinar quanto da variabilidade total é decorrente do medidor ou instrumento.
- Avaliar se o instrumento ou medidor é capaz (isso é, se é adequado ao uso que se pretende).

Componentes da variabilidade

- Para a análise de sistemas de medida, vamos partir do seguinte modelo:

$$y = x + \epsilon,$$

em que y é a medida obervada, x é o verdadeiro valor da medida e ϵ é o erro de medida.

- É usual supor que:

$$x \sim \text{Normal}(\mu, \sigma_P^2), \quad \epsilon \sim \text{Normal}(0, \sigma_M^2),$$

em que o subíndice P e M remetem a processo e medida, respectivamente.

Componentes da mensuração

A **variabilidade atribuível** ao sistema de medição é normalmente associada aos seguintes dois componentes de mensuração:

- **Reprodutibilidade:** variabilidade decorrente do fato do medidor ser utilizado por diferentes operadores, condições de operação ou contorno, períodos de tempo.
- **Repetibilidade:** variabilidade do medidor quando aplicado repetidas vezes em circunstâncias semelhantes: mesmo operador, mesmas condições ambientais, etc.

Reprodutibilidade

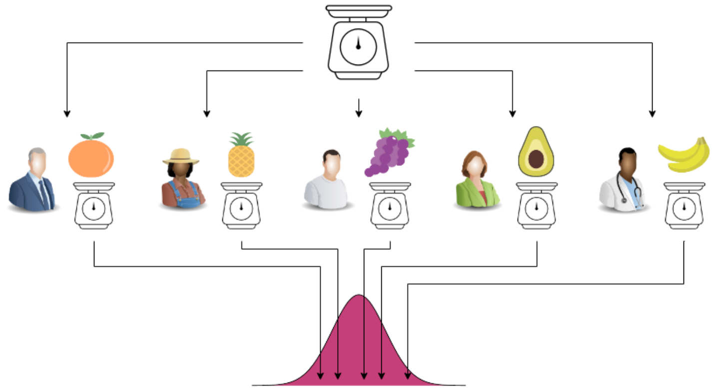


Figura 1. Ilustração de reprodutibilidade. Fonte: os autores. Feito com https://www.draw.io/.

Repetibilidade

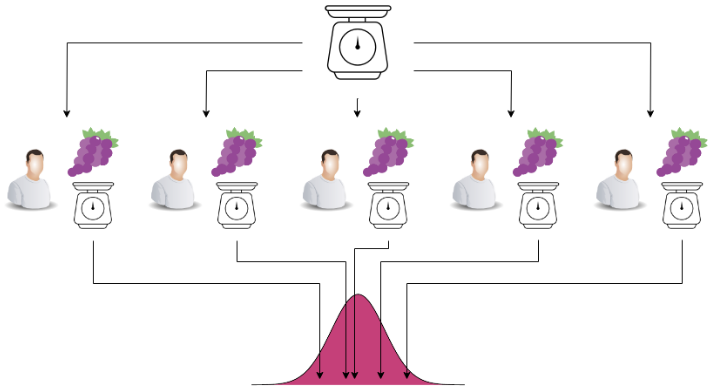


Figura 2. Ilustração de repetibilidade. Fonte: os autores. Feito com https://www.draw.io/.

Conceitos básicos da capacidade de um medidor

Considerando o modelo supramencionado e admitindo que as duas fontes de variação (produto e medição) sejam **independentes**, decorre que a variabilidade total σ_T^2 pode ser decomposta em:

$$\sigma_T^2 = \sigma_P^2 + \sigma_M^2.$$

Adicionalmente, a variabilidade do sistema de medida pode ser decomposta em:

$$\sigma_M^2 = \sigma_{Repe}^2 + \sigma_{Repr}^2.$$

Análise de sistemas de medida

- A **análise da qualidade de sistemas de medida** pode se basear em resultados de experimentos planejados especificamente com tal finalidade.
- Em geral, em tais experimentos dispomos de uma amostra aleatória de p peças, medidas por uma amostra aleatória de o operadores. Cada operador mede cada peça n vezes.
- As medições são feitas de forma **casualizada** e, preferencialmente, por teste cego.
- Tais experimentos são chamados **experimentos fatoriais**.

Ilustração

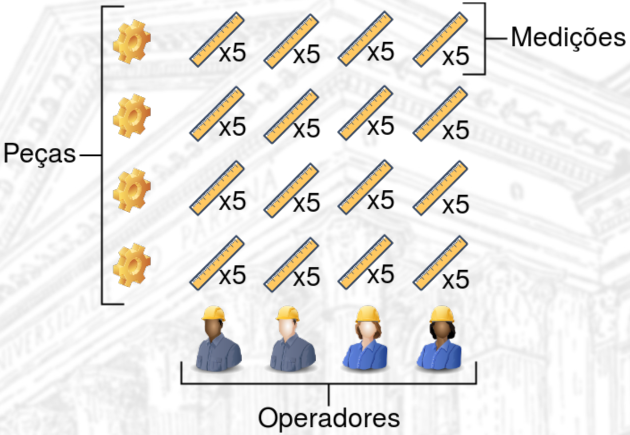


Figura 3. Ilustração de um experimento de sistema de medição com 4 operadores, 4 peças e 5 medições por operador e peça. Cada operador mede as 4 peças 5 vezes de forma aleatorizada. Fonte: os autores. Feito com `https://www.draw.io/`.

Exemplo

Tabela 1. Dados de um experimento de um sistema de medida com 20 peças, 3 operadores e 2 medições.

Peça	Operador 1		Operador 2		Operador 3	
	Teste 1	Teste 2	Teste 1	Teste 2	Teste 1	Teste 2
1	21	20	20	20	19	21
2	24	23	24	24	23	24
3	20	21	19	21	20	22
4	27	27	28	26	27	28
5	19	18	19	18	18	21
6	23	21	24	21	23	22
7	22	21	22	24	22	20
			...			
15	29	30	30	28	31	30
16	26	26	25	26	25	27
17	20	20	19	20	20	20
18	19	21	19	19	21	23
19	25	26	25	24	25	25
20	19	19	18	17	19	17

O método da análise de variância

Um modelo apropriado para o presente fenômeno é definido por:

$$y_{ijk} = \mu + P_i + O_j + (PO)_{ij} + \epsilon_{ijk}$$

em que:

- P_i, $i = 1, 2, \ldots, p$, representam os efeitos de **peças**.
- O_j, $j = 1, 2, \ldots, o$, representam os efeitos dos **operadores**.
- $(PO)_{ij}$ representa o efeito da **interação** (efeito combinado) de peças e operadores.
- ϵ_{ijk}, $k = 1, 2, \ldots, n$, o efeito do **erro**.

O método da análise de variância

No seguinte modelo

$$y_{ijk} = \mu + P_i + O_j + (PO)_{ij} + \epsilon_{ijk},$$

todos os termos indexados (ou efeitos) são **variáveis aleatórias independentes** para os quais assumimos distribuição Normal de média zero e variâncias:

- $\text{Var}(P_i) = \sigma^2_P$.
- $\text{Var}(O_j) = \sigma^2_O$.
- $\text{Var}((PO)_{ij}) = \sigma^2_{PO}$.
- $\text{Var}(\epsilon_{ijk}) = \sigma^2$.

O conjunto de variâncias que compõe a variância de Y é chamado **componentes de variância**.

Ilustração

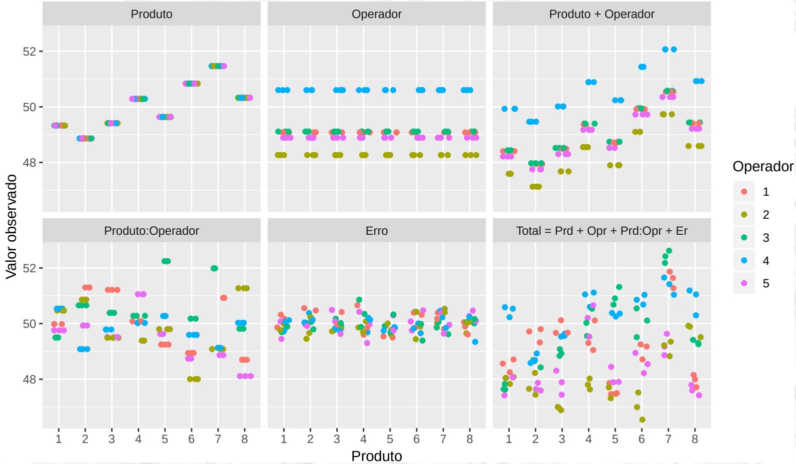


Figura 4. Componentes de variância de um sistema de medida. Fonte: os autores.

Somas de quadrados

- Os componentes de variância são estimados a partir de **somas de quadrados** calculadas com base nos dados amostrais, de tal forma que:

$$\mathrm{SQ}_{\mathrm{Total}} = \mathrm{SQ}_P + \mathrm{SQ}_O + \mathrm{SQ}_{PO} + \mathrm{SQ}_E.$$

- As somas de quadrados **expressam as variabilidades** inerentes a cada fonte de variação (operadores, peças, etc).
- Cada soma de quadrados gera uma respectiva **média quadrática**, definidas da seguinte forma:

$$\mathrm{MQ}_P = \frac{\mathrm{SQ}_P}{p-1}, \qquad \mathrm{MQ}_O = \frac{\mathrm{SQ}_O}{o-1},$$
$$\mathrm{MQ}_{PO} = \frac{\mathrm{SQ}_{PO}}{(p-1)(o-1)}, \qquad \mathrm{MQ}_E = \frac{\mathrm{SQ}_E}{po(n-1)}.$$

Estimação dos componentes de variância

Pelo método da análise de variância os estimadores dos componentes de variância são dados por:

$$
\begin{aligned}
\hat{\sigma}^2 &= \mathrm{MQ}_E,\\
\hat{\sigma}^2_{PO} &= \frac{\mathrm{MQ}_{PO} - \mathrm{MQ}_E}{n},\\
\hat{\sigma}^2_{O} &= \frac{\mathrm{MQ}_{O} - \mathrm{MQ}_{PO}}{pn},\\
\hat{\sigma}^2_{P} &= \frac{\mathrm{MQ}_{P} - \mathrm{MQ}_{PO}}{on}.
\end{aligned}
$$

Estimação da repetibilidade e reprodutibilidade

- Definimos:
  - σ^2 como o componente **repetibilidade** e
  - $\sigma^2_O + \sigma^2_{OP}$ como o componente **reprodutibilidade** do sistema de medida.
- Assim, a variância associada aos componentes RR (variância do medidor) fica dada por:

$$\sigma^2_{\text{Medidor}} = \sigma^2_O + \sigma^2_{OP} + \sigma^2.$$

- A contribuição da variação do medidor na variação total é definida por:

$$\frac{\sigma^2_{\text{Medidor}}}{\sigma^2_{\text{Total}}},$$

que será mínima quanto mais preciso o sistema de medida.

Exercício

Considere os dados apresentados anteriormente sobre medidas de 3 operarios em 20 peças (itens). Faça a análise de variância e determine os componentes de variância. Calcule a variância atribuida a repetibilidade e reprodutibilidade.

```r
library(tidyverse)

# Importa o arquivo de dados.
tb <- read_csv2("../data/medidas.csv", col_names = TRUE)
tb <- tb %>%
    mutate(operador = factor(operador),
           item = factor(item))
glimpse(tb)
```

```
# Observations: 120
# Variables: 4
# $ item      <fct> 1, 1, 2, 2, 3, 3, 4, 4, 5, 5, 6, 6, 7, 7, 8, 8, 9...
# $ operador  <fct> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1...
# $ repeticao <int> 1, 2, 1, 2, 1, 2, 1, 2, 1, 2, 1, 2, 1, 2, 1, 2, 1...
# $ valor     <int> 21, 20, 24, 23, 20, 21, 27, 27, 19, 18, 23, 21, 2...
```

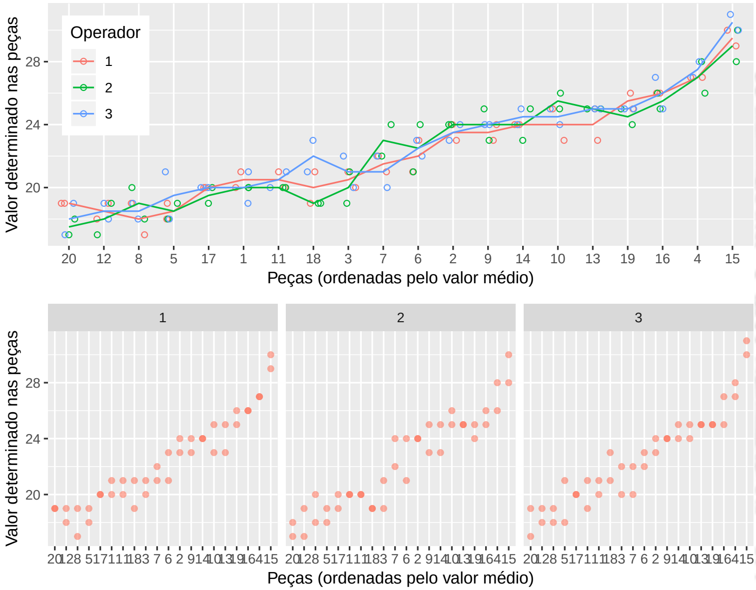
```r
# Quantidade de repetições pode operador x item.
xtabs(~operador + item, data = tb)
```

```
#         item
# operador 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
#        1 2 2 2 2 2 2 2 2 2  2  2  2  2  2  2  2  2  2  2  2
#        2 2 2 2 2 2 2 2 2 2  2  2  2  2  2  2  2  2  2  2  2
#        3 2 2 2 2 2 2 2 2 2  2  2  2  2  2  2  2  2  2  2  2
```

```r
gg1 <-
ggplot(data = tb) +
    aes(x = reorder(item, valor), y = valor,
        color = operador, group = operador) +
    geom_jitter(pch = 1, height = 0, width = 0.25) +
    stat_summary(geom = "line", fun.y = mean) +
    labs(x = "Peças (ordenadas pelo valor médio)",
         y = "Valor determinado nas peças",
         color = "Operador") +
    theme(legend.justification = c(0, 1),
          legend.position = c(0.02, 0.95))

gg2 <-
ggplot(data = tb) +
    aes(x = reorder(item, valor), y = valor) +
    facet_wrap(facets = ~operador) +
    geom_point(alpha = 0.5, color = "tomato") +
    labs(x = "Peças (ordenadas pelo valor médio)",
         y = "Valor determinado nas peças")

# gridExtra::grid.arrange(gg1, gg2, ncol = 1)
```


Operador
1
2
3
Valor determinado nas peças
20
24
28
20 12 8 5 17 1 11 18 3 7 6 2 9 14 10 13 19 16 4 15
Peças (ordenadas pelo valor médio)
1
2
3
Valor determinado nas peças
Peças (ordenadas pelo valor médio)

```r
# Ajusta o modelo aos dados.
m0 <- lm(valor ~ item + operador + item:operador,
         data = tb)

# Exibe o quadro de anova.
anova(m0)
```

```
# Analysis of Variance Table
#
# Response: valor
#               Df  Sum Sq Mean Sq F value Pr(>F)
# item          19 1185.43  62.391 62.9151 <2e-16 ***
# operador       2    2.62   1.308  1.3193 0.2750
# item:operador 38   27.05   0.712  0.7178 0.8614
# Residuals     60   59.50   0.992
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Estimação dos componentes de variância.
anv <- anova(m0)
qm <- anv[, "Mean Sq"]
names(qm) <- rownames(anv)

# Número de repetições, operadores e itens.
n <- length(unique(tb$repeticao))
o <- nlevels(tb$operador)
p <- nlevels(tb$item)

sigma_e  <- qm["Residuals"]
sigma_po <- (qm["item:operador"] - qm["Residuals"])/n
sigma_o  <- (qm["operador"] - qm["item:operador"])/(p * n)
sigma_p  <- (qm["item"] - qm["item:operador"])/(o * n)

# Componentes de variância.
c(sigma_e, sigma_po, sigma_o, sigma_p)
```

```
#     Residuals item:operador      operador          item
#    0.99166667   -0.13991228    0.01491228   10.27982456
```

```r
# Valores negativos não fazem sentido.
sigma_po <- 0
```

```r
# Componentes da variância do sistema de medida.
sigma_reprod <- sigma_o + sigma_po
sigma_repeti <- sigma_e
setNames(c(sigma_reprod, sigma_repeti),
         c("Var. reprod", "Var. repeti"))
```

```
# Var. reprod Var. repeti
#  0.01491228  0.99166667
```

```r
sigma_medidor <- sigma_reprod + sigma_repeti
setNames(sigma_medidor, "Var. do medidor")
```

```
# Var. do medidor
#        1.006579
```

```r
# Propoção da variância devido ao sistema de medida.
propor <- sigma_medidor/(sigma_e + sigma_po + sigma_o + sigma_p)
setNames(propor, "Prop. de variância")
```

```
# Prop. de variância
#          0.08918509
```

```r
library(lme4)

# Modelo de efeitos aleatórios com estimação REML.
m0 <- lmer(valor ~
               (1 | item) +
               (1 | operador) +
               (1 | item:operador),
           data = tb,
           REML = TRUE)
# summary(m0)

# Extração das variâncias dos efeitos.
sigmas <- c(sapply(VarCorr(m0), "[[", 1),
            Residuals = attr(VarCorr(m0), "sc"))
sigmas
```

```
# item:operador          item      operador     Residuals
#    0.00000000   10.25127067    0.01062925    0.93976767
```

```r
# Cálculo da proporção da variância.
(sigmas["item:operador"] +
 sigmas["operador"] +
 sigmas["Residuals"])/ sum(sigmas)
```

```
# item:operador
#    0.08484424
```

```r
library(SixSigma)

# Todos os resultados acima conseguidos.
anv <- ss.rr(var = valor,
             part = item,
             appr = operador,
             data = tb)
```

```r
# Quadro de ANOVA do modelo completo.
anv$anovaTable
```

```
#                Df Sum Sq Mean Sq F value Pr(>F)
# item           19 1185.4   62.39  87.647 <2e-16
# operador        2    2.6    1.31   1.838  0.173
# item:operador  38   27.0    0.71   0.718  0.861
# Repeatability  60   59.5    0.99
# Total         119 1274.6
```

```r
# Quadro de ANOVA do modelo reduzido.
anv$anovaRed
```

```
#                 Df Sum Sq Mean Sq F value Pr(>F)
# item            19 1185.4   62.39  70.645 <2e-16
# operador         2    2.6    1.31   1.481  0.232
# Repeatability   98   86.6    0.88
# Total          119 1274.6
```

```r
# Quadro com os componentes de variância.
anv$varComp
```

```
#                       VarComp %Contrib
# Total Gage R&R     0.89379252     8.02
#   Repeatability    0.88316327     7.92
#   Reproducibility  0.01062925     0.10
#     operador       0.01062925     0.10
# Part-To-Part      10.25127103    91.98
# Total Variation   11.14506355   100.00
```

Capacidade do sistema de medida

- Uma forma alternativa de quantificar a **performance do medidor** baseia-se na razão dos desvios padrões:

$$\%RR = 100 \times \frac{\sigma_{\text{Medidor}}}{\sigma_{\text{Total}}}.$$

- Uma classificação usual da **performance de sistemas de medida**, baseada em $\%RR$, é dada por:
  - $\%RR \leq 10\%$: sistema capaz.
  - $10\% < \%RR \leq 30\%$: sistema aceitável.
  - $\%RR > 30\%$: sistema incapaz.
- Dessa forma, seja usando as variâncias ou os desvios padrões, quanto menor a variação relativa ao total, melhor a performance do sistema de medida.