



**ESTUDO DE CASO: ANÁLISE DA VARIABILIDADE ASSOCIADA À DURAÇÃO  
DAS CIRURGIAS MASTOLÓGICAS DO HOSPITAL DE CÂNCER III DO INSTITUTO  
NACIONAL DE CÂNCER (INCA)**

Alcides Algarra Grande

Projeto de Graduação apresentado ao Curso de Engenharia de Produção Escola Politécnica, Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do grau de Engenheiro de Produção.

Orientadores: Flavia Maria Pinto Ferreira Landim, D.Sc.

Luiz Antonio Meirelles, D.Sc.

Rio de Janeiro

Julho de 2013

**ESTUDO DE CASO: ANÁLISE DA VARIABILIDADE ASSOCIADA À DURAÇÃO  
DAS CIRURGIAS MASTOLÓGICAS DO HOSPITAL DE CÂNCER III DO INSTITUTO  
NACIONAL DE CÂNCER (INCA)**

Alcides Algarra Grande

PROJETO DE GRADUAÇÃO SUBMETIDO AO CORPO DOCENTE DO CURSO  
DE ENGENHARIA DE PRODUÇÃO DA ESCOLA POLITÉCNICA DA  
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS  
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE  
ENGENHEIRO DE PRODUÇÃO

Examinado por:

---

Prof. Samuel Jurkiewicz, D.Sc. (Examinador)

---

Prof. Edilson Fernandes de Arruda, D.Sc. (Examinador)

---

Prof. Leonardo Luiz Lima Navarro, M.Sc. (Examinador)

---

Profa. Flávia Maria Pinto Ferreira Landim, D.Sc. (Orientadora)

---

Prof. Luiz Antônio Meirelles, D.Sc. (Co-orientador)

Rio de Janeiro

Julho 2013

Algarra Grande, Alcides

Estudo de caso: Análise da variabilidade associada à duração das cirurgias mastológicas do hospital de câncer III do Instituto Nacional de Câncer (INCA)/ Alcides Algarra Grande.

– Rio de Janeiro: UFRJ/ Escola Politécnica, 2013.

X, p.141 il.; 29,7 cm.

Orientadores: Flavia Maria Pinto Ferrira Landim, D.Sc.

Luiz Antônio Meirelles, D.Sc.

Projeto de Graduação – UFRJ/ Escola Politécnica/

Curso de Engenharia de Produção, 2013.

Referencias Bibliográficas: p. 138-141.

1. Cirurgias 2. Fatores de Variabilidade. 3. Predição.

4. Programação. 5. Estudo de Caso

I. Flavia Maria Pinto Ferreira Landim, Luiz Antônio Meirelles.

II. Universidade Federal do Rio de Janeiro, Escola Politécnica,

Curso de Engenharia De Produção. III. Título.

## AGRADECIMENTOS

*Agradeço primeiramente a minha família, por entender a minha vontade de viver a experiência brasileira.*

*Em segundo lugar a Laura, sem ela muitas coisas na minha vida não seriam possíveis.*

*Agradeço também a Rafael, Isabela e Belen, por virar minha família desde o primeiro dia que cheguei no Rio.*

*Agradeço aos meus orientadores, a professora Flavia e o professor Meirelles, por ajudar-me em todo momento. E muito gratificante e motivador encontrar professores verdadeiramente preocupados pelo aluno.*

*Agradeço também ao doutor Rodrigo e ao Cleber, por ter-me ajudado em tudo o que precisava.*

*E, por último, agradeço a “excelentíssima” Eliana Labra, a chilena de rua Alice por sua inestimável ajuda em muitos sentidos, humanos e intelectuais, virada minha nova avó “putativa”.*

Resumo do Projeto de Graduação apresentado à Escola Politécnica/ UFRJ como parte dos requisitos necessários para a obtenção do grau de Engenheiro de Produção.

Estudo de caso: Análise da variabilidade associada à duração das cirurgias mastológicas do hospital de câncer III do Instituto Nacional de Câncer (INCA)

Alcides Algarra Grande

Julho/2013

Orientadores: Flavia Maria Pinto Ferreira Landim, D.Sc.

Luiz Antonio Meirelles, D.Sc.

Curso: Engenharia de Produção

Neste projeto, foram estudadas as fontes de variabilidades associada com a duração das principais cirurgias da área de mastologia do Hospital de Câncer III do INCA (Instituto Nacional do Câncer), a fim de reduzir a incerteza associada com a demora dos procedimentos, e fazer programações (mapas cirúrgicos) que possuam uma maior correspondência com a sua execução posterior.

Foram investigados fatores associados ao paciente (IMC, hipertensão, diabetes), cirurgião (sexo, idade, trimestre, composição da equipe cirúrgica e hora do dia que começa a cirurgia) e ao próprio procedimento (Biopsia Linfonodo Sentinela, ampliação de margens e Linfodectomia Axilar), a fim de estabelecer uma equação de regressão linear múltipla, que permita predizer com mais precisão a duração de cirurgias futuras. Para esse fim obteve-se uma amostra aleatória de prontuários médicos dos anos de 2010, 2011 e 2012.

Alguns dos fatores incluídos se revelaram significativos na análise exploratória de dados, testes de hipóteses e na seleção de variáveis da equação de regressão (*Stepwise Regression*). No entanto, a quantidade de variabilidade explicada resultante foi moderada.

Palavras-chave: Cirurgias, Fatores de Variabilidade, Predição, Programação, Estudo de Caso

Abstract of Undergraduate Project presented to POLI/UFRJ as a partial fulfillment of the requirements for the degree of Engineer.

Case Study: Analysis of the variability associated with the duration of mastological surgeries of the Hospital de Câncer III of INCA (Instituto Nacional de Câncer)

Alcides Algarra Grande

July 2013

Advisors: Flavia Maria Pinto Ferreira Landim, D.Sc.

Luiz Antônio Meirelles, D.Sc.

Course: Production Engineering

The high volume of resources associated with the operating theatre of a hospital, makes the surgical centre a space of real importance in its organization. In this project we have determined the degree of variability associated with the duration of major surgeries of the mastology area of the Hospital de Câncer III of INCA (Instituto Nacional de Câncer), in order to reduce the uncertainty associated with the delay of the surgeries, and to make schedules (surgical maps) that possess greater robustness with later execution.

We studied factors associated with the patient (IMC, hypertension, diabetes), surgeon (gender, age, trimester surgical team composition and time of day that begins the surgery) and the procedure itself (Linfonodo Biopsy Sentinel, expansion of margins and axillary lymphadenectomy) in order to establish a linear regression equation to predict more accurately the duration of future surgeries. For this we have obtained a random sample of medical files of the years 2010, 2011 and 2012.

Some of the factors included rebelled as significant in exploratory data analysis, in tests of hypotheses and in selecting variables in the regression equation (*Stepwise Selection*). However, the amount of variability explained is moderate.

Key-words: Surgeries, Variability Factors, Prediction, Scheduling, Case Study

## Sumário

|   |    |
|---|----|
| CAPÍTULO 1: INTRODUÇÃO .....  | 15 |
| CAPÍTULO 2: MARCO TEÓRICO.....  | 20 |
| 2.1 Cirurgia .....  | 20 |
| 2.2 Taxa de utilização da sala de cirurgia .....  | 21 |
| 2.3 A incerteza.....  | 24 |
| 2.4 Programação das salas de cirurgia ( <i>Scheduling</i> ) .....                                 | 25 |
| 2.4.2 Metodologias de Scheduling sob incerteza.....   | 29 |
| 2.4.3 Fases para a programação das operações de um centro cirúrgico.....                          | 33 |
| 2.4.4 Ferramentas de pesquisa operacional aplicadas ao planejamento da sala de cirurgia.<br>..... | 36 |
| CAPÍTULO 3: FERRAMENTAS DE ANÁLISE ESTATÍSTICA NO ESTUDO DA<br>VARIABILIDADE.....                 | 37 |
| 3.1 Caracterização dos dados. Exigência de normalidade. ....                                      | 37 |
| 3.2 Comparações de Médias .....   | 41 |
| 3.3 Regressão Linear .....  | 44 |
| 3.3.1 Regressão Linear Simples .....  | 44 |
| 3.3.2 Análise dos resíduos.....   | 50 |
| 3.3.3 Previsão de novas observações.....  | 53 |
| 3.3.4 Regressão Linear Múltipla .....   | 54 |
| 3.3.5 Seleção de variáveis em regressão linear múltipla: <i>Stepwise Regression</i> .....         | 59 |
| CAPÍTULO 4: ESTUDO DE CASO .....  | 61 |
| 4.1 Introdução .....  | 61 |
| 4.1.1 O câncer de mama .....  | 61 |
| 4.1.2 O Instituto Nacional de Câncer (INCA). HCIII.....   | 61 |
| 4.1.3 Procedimentos envolvidos na área de mastologia do HCIII. Descrição geral .....              | 63 |
| 4.1.3.1 Mastectomia .....   | 63 |

|  |     |
|--|-----|
| 4.1.3.2 Segmentectomia .....   | 64  |
| 4.1.4 Fatores de variação estudados.....   | 66  |
| 3.1.4.3 Fatores associados ao procedimento.....  | 73  |
| 3.2 Obtenção da informação.....  | 74  |
| 3.2.1 O sistema de informação do hospital .....  | 74  |
| 3.2.2 O prontuário médico.....   | 75  |
| 3.3 Análise e interpretação dos resultados.....  | 76  |
| 3.3.1 Descrição da metodologia de análise.....   | 76  |
| 3.3.2 Apresentação dos procedimentos considerados e os fatores de variação estudados ..... | 77  |
| 3.3.3 Mastectomia Radical Modificada (MRM) .....   | 81  |
| 3.2.3.1 Análise exploratória dos dados .....   | 83  |
| 3.2.3.2 Interações consideradas no modelo de regressão linear. ....                        | 95  |
| 3.2.3.3 Análise de Regressão e interpretação dos resultados .....                          | 96  |
| 3.3.4 Mastectomia Simples (MS) .....   | 101 |
| 3.3.4.1 Análise exploratória dos dados .....   | 102 |
| 3.3.4.2 Interações consideradas no modelo de regressão linear .....                        | 114 |
| 3.3.4.3 Análise de Regressão e interpretação dos resultados .....                          | 115 |
| 3.3.5 Segmentectomia (SEG).....  | 119 |
| 3.3.5.1 Análise exploratória dos dados .....   | 120 |
| 3.3.5.2 Interações consideradas no modelo de regressão linear. ....                        | 132 |
| 3.3.5.3 Análise de Regressão e interpretação dos resultados .....                          | 133 |
| 3.4 Verificação do modelo ajustado.....  | 136 |
| PARTE 4: CONCLUSÕES E DISCUSSÃO .....  | 141 |
| PARTE 5: BIBLIOGRAFIA .....  | 148 |
| ANEXO I: ESTUDO DE SIMULAÇÃO DAS SALAS DE CIRURGIA .....                                   | 152 |



## LISTA DE FIGURAS

|  |    |
|--|----|
| FIGURA 1. EXEMPLO GRÁFICO DO PROCESSO DE CIRURGIA. EXTRAÍDO DE [6].  | 20 |
| FIGURA 2. ESQUEMA DE TEMPOS DE OPERAÇÃO, TEMPO DE PREPARAÇÃO, SETUP E PERDAS EM UM DIA DE UMA SALA DE CIRURGIA. ELABORAÇÃO PRÓPRIA.  | 22 |
| FIGURA 3. PROGRAMAÇÃO DA SALA MELHORADA. ELABORAÇÃO PRÓPRIA.   | 23 |
| FIGURA 4. CORRESPONDÊNCIA ENTRE A ESTIMAÇÃO E A REALIDADE NA DURAÇÃO DE UM PROCESSO. ELABORAÇÃO PRÓPRIA.   | 25 |
| FIGURA 5. ESQUEMA DE FLOW SHOP CLÁSSICO. ELABORAÇÃO PRÓPRIA.   | 26 |
| FIGURA 6. JOB-SHOP SCHEDULING CLÁSSICO. ELABORAÇÃO PRÓPRIA.  | 26 |
| FIGURA 7. CONSEQUÊNCIAS E IMPLICAÇÕES DOS DIFERENTES CENÁRIOS RESULTANTES DA INCERTEZA DO PROGRAMADOR DAS SALAS RESPEITO ÀS DURAÇÕES DE UM PROCEDIMENTO TIPO. ELABORAÇÃO PRÓPRIA.              | 28 |
| FIGURA 8. SUPERFÍCIE PARETO-ÓPTIMA CONSIDERANDO A DURAÇÃO TOTAL NO HORIZONTE DE PROGRAMAÇÃO DAS OPERAÇÕES ( <i>MAKESPAN</i> ), A ROBUSTEZ DA SOLUÇÃO E A ROBUSTEZ DO MODELO. EXTRAÍDO DE [13]. | 31 |
| FIGURA 9. PAPEL PROBABILÍSTICO NORMAL. EXEMPLO DE AMOSTRA EM QUE A SUPOSIÇÃO DE NORMALIDADE NÃO SE PODE ASSUMIR. ELABORAÇÃO PRÓPRIA.   | 40 |
| FIGURA 10. PAPEL PROBABILÍSTICO NORMAL 2. EXEMPLO DE AMOSTRA EM QUE A SUPOSIÇÃO DE NORMALIDADE SE PODE. ELABORAÇÃO PRÓPRIA.  | 40 |
| FIGURA 11: REPRESENTAÇÃO GRÁFICA DOS COEFICIENTES DE REGRESSÃO LINEAR SIMPLES. EXTRAÍDO DE [19].   | 46 |
| FIGURA 12. TIPOS GENÉRICOS DE GRÁFICOS DE RESÍDUOS VS. VALORES AJUSTADOS. EXTRAÍDO DE [19].  | 51 |
| FIGURA 13. ORGANOGRAMA GERAL DO INSTITUTO NACIONAL DE CÂNCER (INCA). EXTRAÍDO DE [23].   | 62 |
| FIGURA 14. FONTES DE VARIABILIDADE E FATORES ESTUDADOS. ELABORAÇÃO PRÓPRIA.  | 67 |
| FIGURA 15. DISTRIBUIÇÃO DOS INTERVALOS DE IMC EM CADA PROCEDIMENTO.  | 78 |
| FIGURA 16. GRUPOS DE PACIENTES HIPERTENSOS E NÃO HIPERTENSOS.  | 79 |
| FIGURA 17. GRUPOS DE PACIENTES DIABÉTICOS E NÃO DIABÉTICOS.  | 79 |
| FIGURA 18. CIRURGIAS PRATICADAS POR HOMENS E POR MULHERES.   | 80 |
| FIGURA 19. CIRURGIAS COM AS DUAS COMPOSIÇÕES DE EQUIPES CIRÚRGICAS CONSIDERADAS.   | 80 |
| FIGURA 20. GRUPOS DE CIRURGIAS QUE COMEÇAM ANTES E DEPOIS DAS 12:00 H.   | 81 |
| FIGURA 21. HISTOGRAMA E GRÁFICA DE PROBABILIDADE NORMAL DOS DADOS DE DURAÇÃO DAS CIRURGIAS (MRM).  | 82 |
| FIGURA 22. GRÁFICA DE PROBABILIDADE NORMAL DA VARIÁVEL TRANSFORMADA DO TEMPO DE CIRURGIA (MS).   | 83 |
| FIGURA 23. BOX-PLOT DE DURAÇÃO DA CIRURGIA VS. INTERVALO DE IMC (MRM).   | 84 |
| FIGURA 24. BOX-PLOT DE DURAÇÃO DE CIRURGIAS VS. HIPERTENSÃO (MRM).   | 86 |
| FIGURA 25. BOX-PLOT PACIENTES DIABÉTICAS E NÃO DIABÉTICAS (MRM).   | 87 |
| FIGURA 26. BOX-PLOT CIRURGIAS REALIZADAS POR HOMENS E POR MULHERES (MRM).  | 88 |
| FIGURA 27. BOX-PLOT CIRURGIAS REALIZADAS POR CIRURGIÕES QUE NASCERAM ANTES E DEPOIS DE 1960 (MRM).   | 90 |
| FIGURA 28. BOX-PLOT DE CIRURGIAS REALIZADAS EN LOS 4 TRIMESTRES DEL AÑO (MRM).   | 91 |
| FIGURA 29. BOX-PLOT DAS CIRURGIAS EM QUE O CIRURGIÃO <i>STAFF</i> PARTICIPA OU NÃO NA EQUIPE CIRÚRGICA (MRM).  | 93 |

|  |     |
|--|-----|
| FIGURA 30. BOX-PLOT DE CIRURGIAS QUE COMEÇARAM ANTES E DEPOIS DAS 12:00 H (MRM).....                       | 94  |
| FIGURA 31. GRÁFICOS DE RESÍDUOS DA REGRESSÃO LINEAR (MRM).....   | 99  |
| FIGURA 32. HISTOGRAMA E GRÁFICA DE PROBABILIDADE NORMAL DOS DADOS DE DURAÇÃO DAS CIRURGIAS (MS) .....      | 101 |
| FIGURA 33. GRÁFICA DE PROBABILIDADE NORMAL DA VARIÁVEL TRANSFORMADA DO TEMPO DE CIRURGIA (MS).....         | 102 |
| FIGURA 34. BOX-PLOT DE DURAÇÃO DA CIRURGIA VS. INTERVALO DE IMC (MS).....                                  | 103 |
| FIGURA 35. BOX-PLOT DE DURAÇÃO DAS CIRURGIAS VS. HIPERTENSÃO (MS).....                                     | 104 |
| FIGURA 36. BOX-PLOT PACIENTES DIABÉTICAS E NÃO DIABÉTICAS (MS).....  | 105 |
| FIGURA 37. BOX-PLOT CIRURGIAS REALIZADAS POR HOMENS E POR MULHERES (MS).....                               | 107 |
| FIGURA 38. CIRURGIAS REALIZADAS POR CIRURGIÕES QUE NASCERAM ANTES E DEPOIS DE 1960 (MS) .....              | 108 |
| FIGURA 39. BOX-PLOT DE CIRURGIAS REALIZADAS NOS 4 TRIMESTRES DO ANO (MS).....                              | 109 |
| FIGURA 40. BOX-PLOT DAS CIRURGIAS EM QUE O CIRURGIÃO STAFF PARTICIPA OU NÃO NA EQUIPE CIRÚRGICA (MS).....  | 110 |
| FIGURA 41. BOX-PLOT DE CIRURGIAS QUE COMEÇARAM ANTES E DEPOIS DAS 12:00 H (MS) .....                       | 112 |
| FIGURA 42. BOX-PLOT COM AS CIRURGIAS EM QUE SE REALIZOU OU NÃO A BIOPSIA (MS) .....                        | 113 |
| FIGURA 43. GRÁFICOS DE RESÍDUOS DA REGRESSÃO LINEAR (MS).....  | 118 |
| FIGURA 44. HISTOGRAMA E GRÁFICA DE PROBABILIDADE NORMAL DOS DADOS DE DURAÇÃO DAS CIRURGIAS (SEG) .....     | 119 |
| FIGURA 45. . GRÁFICA DE PROBABILIDADE NORMAL DA VARIÁVEL TRANSFORMADA DO TEMPO DE CIRURGIA (SEG).....      | 120 |
| FIGURA 46. BOX-PLOT DE DURAÇÃO DA CIRURGIA VS. INTERVALO DE IMC (SEG) .....                                | 121 |
| FIGURA 47. BOX-PLOT DE DURAÇÃO DAS CIRURGIAS VS. HIPERTENSÃO (SEG).....                                    | 122 |
| FIGURA 48. BOX-PLOT PACIENTES DIABÉTICAS E NÃO DIABÉTICAS (SEG).....                                       | 123 |
| FIGURA 49. BOX-PLOT CIRURGIAS REALIZADAS POR HOMENS E POR MULHERES (SEG).....                              | 124 |
| FIGURA 50. CIRURGIAS REALIZADAS POR CIRURGIÕES QUE NASCERAM ANTES E DEPOIS DE 1960 (SEG) .....             | 126 |
| FIGURA 51. BOX-PLOT DE CIRURGIAS REALIZADAS NOS 4 TRIMESTRES DO ANO (SEG).....                             | 127 |
| FIGURA 52. BOX-PLOT DAS CIRURGIAS EM QUE O CIRURGIÃO STAFF PARTICIPA OU NÃO NA EQUIPE CIRÚRGICA (SEG)..... | 128 |
| FIGURA 53. BOX-PLOT DE CIRURGIAS QUE COMEÇARAM ANTES E DEPOIS DAS 12:00 H (SEG) .....                      | 130 |
| FIGURA 54. BOX-PLOT DE CIRURGIAS REALIZADAS DE ACORDO À VARIAÇÃO DO PROCEDIMENTO (SEG).....                | 131 |
| FIGURA 55. GRÁFICOS DE RESÍDUOS DA REGRESSÃO LINEAR (SEG).....   | 136 |
| FIGURA 56. CIRURGIAS REALIZADAS POR ÁREA CIRÚRGICA DO HOSPITAL.....  | 137 |
| FIGURA 57. CIRURGIAS REALIZADAS NA ÁREA DE MASTOLOGIA .....  | 138 |
| FIGURA 58. COMPARAÇÃO VISUAL DOS ERROS DE ESTIMAÇÃO .....  | 140 |
| FIGURA 59. SIMULAÇÃO DA DURAÇÃO DA CIRURGIA 1. ....  | 145 |
| FIGURA 60. SIMULAÇÃO DA DURAÇÃO DA CIRURGIA 2. ....  | 146 |

## LISTA DE TABELAS

|   |    |
|---|----|
| TABELA 1. RESUMO DOS NÍVEIS DE PLANEJAMENTO, OBJETIVOS DO PLANEJAMENTO E EXEMPLOS DE CRITÉRIOS DE OTIMIZAÇÃO. ....  | 35 |
| TABELA 2. RESUMO DE MODELOS DE OTIMIZAÇÃO, NATUREZA DOS MODELOS E ABORDAGEM DA SOLUÇÃO. DADOS EXTRAÍDOS DE [5]. ....  | 36 |
| TABELA 3. TAMANHO, MÉDIA E DESVIO PADRÃO AMOSTRAIS DOS PROCEDIMENTOS ANALISADOS. ....   | 77 |
| TABELA 4. IMC MÉDIO PARA CADA PROCEDIMENTO. ....  | 78 |
| TABELA 5. TRANSFORMAÇÃO DE BOX-COX PARA OS DADOS (MRM).....   | 82 |
| TABELA 6. QUANTIDADE DE CIRURGIAS POR INTERVALO DE IMC (MRM).....   | 83 |
| TABELA 7. MÉDIA DE DURAÇÃO DA CIRURGIA PARA CADA INTERVALO DO IMC (MRM). ....   | 84 |
| TABELA 8. INTERVALOS DE CONFIANÇA SIMULTÂNEOS DE TUKEY PARA A DIFERENÇA DE MÉDIAS DOS INTERVALOS DE IMC (MRM).....  | 85 |
| TABELA 9. QUANTIDADE OBSERVAÇÕES DE PACIENTES HIPERTENSAS E NÃO HIPERTENSAS (MRM) .....   | 85 |
| TABELA 10. MÉDIA DA DURAÇÃO DAS CIRURGIAS DE PACIENTES HIPERTENSAS E NÃO HIPERTENSAS (MRM).....   | 86 |
| TABELA 11. RESULTADO DO INTERVALO DE 95% DE CONFIANÇA DA DIFERENÇA DE MÉDIAS ENTRE PACIENTES HIPERTENSAS E NÃO HIPERTENSAS (MRM).....   | 86 |
| TABELA 12. PACIENTES DIABÉTICAS E NÃO DIABÉTICAS (MRM).....   | 87 |
| TABELA 13. MÉDIAS DE DURAÇÃO DAS PACIENTES DIABÉTICAS E NÃO DIABÉTICAS (MRM) .....  | 87 |
| TABELA 14. CIRURGIAS REALIZADAS POR HOMENS E POR MULHERES (MRM).....  | 88 |
| TABELA 15. MÉDIA DAS DURAÇÕES DE CIRURGIAS REALIZADAS POR HOMENS E MULHERES (MRM).....  | 89 |
| TABELA 16. RESULTADO DO INTERVALO DE 95% DE CONFIANÇA DA DIFERENÇA DE MÉDIAS DE DURAÇÕES DE CIRURGIAS REALIZADAS POR HOMENS E POR MULHERES (MRM).....   | 89 |
| TABELA 17. CIRURGIAS PRATICADAS POR CIRURGIÕES QUE NASCERAM ANTES E DEPOIS DE 1960 (MRM) .....  | 89 |
| TABELA 18. MÉDIAS DAS CIRURGIAS REALIZADAS POR CIRURGIÕES QUE NASCERAM ANTES E DEPOIS DE 1960 (MRM) .....   | 90 |
| TABELA 19. RESULTADO DO INTERVALO DE 95% DE CONFIANÇA DA DIFERENÇA DE MÉDIAS ENTRE AS CIRURGIAS REALIZADAS POR CIRURGIÕES QUE NASCERAM ANTES E DEPOIS DE 1960 (MRM) .....                                 | 90 |
| TABELA 20. CIRURGIAS REALIZADAS NOS DIFERENTES TRIMESTRES DO ANO (MRM) .....  | 91 |
| TABELA 21. MÉDIA DE DURAÇÃO DAS CIRURGIAS REALIZADAS NOS 4 TRIMESTRES DO ANO (MRM).....   | 91 |
| TABELA 22. RESULTADO DOS INTERVALOS DE CONFIANÇA SIMULTÂNEOS DE TUKEY (95% DE CONFIANÇA PARA A FAMÍLIA) DE DIFERENÇAS DAS MÉDIAS DAS DURAÇÕES DE CIRURGIAS REALIZADAS EM CADA TRIMESTRE DO ANO (MRM)..... | 92 |
| TABELA 23. CIRURGIAS REALIZADAS EM QUE O CIRURGIÃO <i>STAFF</i> PARTICIPA OU NÃO NA EQUIPE CIRÚRGICA (MRM). ....  | 92 |
| TABELA 24. MÉDIAS DAS DURAÇÕES DAS CIRURGIAS EM QUE O CIRURGIÃO <i>STAFF</i> PARTICIPOU OU NÃO NA EQUIPE CIRÚRGICA (MRM).....   | 93 |
| TABELA 25. RESULTADO DO INTERVALO DE 95% DE CONFIANÇA DA DIFERENÇA DAS MÉDIAS DE CIRURGIAS EM QUE O CIRURGIÃO <i>STAFF</i> PARTICIPOU OU NÃO NA EQUIPE CIRÚRGICA (MRM).....                               | 93 |
| TABELA 26. CIRURGIAS QUE COMEÇARAM ANTES E DEPOIS DAS 12:00 H (MRM) .....   | 94 |

|   |     |
|---|-----|
| TABELA 27. MÉDIA DAS DURAÇÕES DE CIRURGIAS QUE COMEÇARAM ANTES E DEPOIS DAS 12:00 H (MRM) .....   | 94  |
| TABELA 28. RESULTADO DO INTERVALO DE 95% DE CONFIANÇA DA DIFERENÇA DE MÉDIAS DAS DURAÇÕES DE CIRURGIAS QUE COMEÇARAM ANTES E DEPOIS DAS 12:00 H (MRM) .....   | 94  |
| TABELA 29. VARIÁVEIS DE INTERAÇÃO CONSIDERADAS NO MODELO DE REGRESSAO LINEAR (MRM) .....  | 95  |
| TABELA 30. NATUREZA DAS VARIÁVEIS ORIGINAIS E DE INTERAÇÃO CONSIDERADAS NO MODELO DE REGRESSÃO LINEAR (MRM).....  | 96  |
| TABELA 31. TRANSFORMAÇÃO DE BOX-COX PARA OS DADOS (MS) .....  | 101 |
| TABELA 32. QUANTIDADE DE CIRURGIAS POR INTERVALO DE IMC (MS) .....  | 102 |
| TABELA 33. MÉDIA DE DURAÇÃO DE CIRURGIA PARA CADA INTERVALO DO IMC (MS) .....   | 103 |
| TABELA 34. INTERVALOS DE CONFIANÇA SIMULTÂNEOS DE TUKEY (95% PARA A FAMÍLIA) DE DIFERENÇA DE MÉDIAS DOS INTERVALOS DE IMC (MS) .....  | 103 |
| TABELA 35. PACIENTES HIPERTENSAS E NÃO HIPERTENSAS (MS).....  | 104 |
| TABELA 36. MÉDIA DAS DURAÇÕES DE CIRURGIA DE PACIENTES HIPERTENSAS E NÃO HIPERTENSAS (MS).....  | 104 |
| TABELA 37. RESULTADO DO INTERVALO DE 95% DE CONFIANÇA DA DIFERENÇA DE MÉDIAS DE DURAÇÃO DE CIRURGIAS DE PACIENTES HIPERTENSAS E NÃO HIPERTENSAS (MS). .....   | 105 |
| TABELA 38. PACIENTES DIABÉTICAS E NÃO DIABÉTICAS (MS) .....   | 105 |
| TABELA 39. MÉDIAS DE DURAÇÃO DAS PACIENTES DIABÉTICAS E NÃO DIABÉTICAS (MS) .....   | 106 |
| TABELA 40. RESULTADO DO INTERVALO DE 95% DE CONFIANÇA DE DIFERENÇA DAS MÉDIAS DE DURAÇÃO DE CIRURGIAS DE PACIENTES DIABÉTICAS E NÃO DIABÉTICAS (MS).....  | 106 |
| TABELA 41. CIRURGIAS REALIZADAS POR HOMENS E POR MULHERES (MS) .....  | 106 |
| TABELA 42. MÉDIA DAS DURAÇÕES DE CIRURGIAS REALIZADAS POR HOMENS E MULHERES (MS).....   | 107 |
| TABELA 43. CIRURGIAS PRATICADAS POR CIRURGIÕES QUE NASCERAM ANTES E DEPOIS DE 1960 (MS).....  | 107 |
| TABELA 44. MÉDIAS DAS CIRURGIAS REALIZADAS POR CIRURGIÕES QUE NASCERAM ANTES E DEPOIS DE 1960 (MS).....   | 108 |
| TABELA 45. RESULTADO DO INTERVALO DE 95% DE CONFIANÇA DA DIFERENÇA DE MÉDIAS ENTRE AS CIRURGIAS REALIZADAS POR CIRURGIÕES QUE NASCERAM ANTES E DEPOIS DE 1960 (MS) .....                                  | 108 |
| TABELA 46. CIRURGIAS REALIZADAS NOS DIFERENTES TRIMESTRES DO ANO (MS).....  | 109 |
| TABELA 47. MÉDIA DE DURAÇÃO DAS CIRURGIAS REALIZADAS NOS QUATRO TRIMESTRES DO ANO (MS) .....  | 109 |
| TABELA 48. RESULTADO DOS INTERVALOS DE CONFIANÇA SIMULTÂNEOS DE TUKEY (95% DE CONFIANÇA PARA A FAMÍLIA) DE DIFERENÇAS DAS MÉDIAS DAS DURAÇÕES DE CIRURGIAS REALIZADAS EM CADA TRIMESTRE DO ANO (MS) ..... | 110 |
| TABELA 49. CIRURGIAS EM QUE O CIRURGIÃO STAFF PARTICIPA OU NÃO NA EQUIPE CIRÚRGICA (MS). .....  | 110 |
| TABELA 50. MÉDIAS DAS DURAÇÕES DAS CIRURGIAS EM QUE O CIRURGIÃO STAFF PARTICIPOU OU NÃO NA EQUIPE CIRÚRGICA (MS) .....  | 111 |
| TABELA 51. RESULTADO DO INTERVALO DE 95% DE CONFIANÇA DA DIFERENÇA DAS MÉDIAS DE CIRURGIAS EM QUE O CIRURGIÃO STAFF PARTICIPOU OU NÃO NA EQUIPE CIRÚRGICA (MS) .....                                      | 111 |
| TABELA 52. CIRURGIAS QUE COMEÇARAM ANTES E DEPOIS DAS 12:00 H (MS) .....  | 111 |
| TABELA 53. MÉDIAS DE DURAÇÃO DAS CIRURGIAS QUE COMEÇARAM ANTES E DEPOIS DAS 12:00 H (MS).....   | 112 |
| TABELA 54. RESULTADO DO INTERVALO DE 95% DE CONFIANÇA DA DIFERENÇA DE MÉDIAS DAS DURAÇÕES DE CIRURGIAS QUE COMEÇARAM ANTES E DEPOIS DAS 12:00 H (MS).....   | 112 |

|  |     |
|--|-----|
| TABELA 55. CIRURGIAS EM QUE SE REALIZOU OU NÃO A BIOPSIA LINFONODO SENTINELA (MS).....   | 113 |
| TABELA 56. MÉDIA DA DURAÇÃO DAS CIRURGIAS EM QUE SE REALIZOU OU NÃO A BIOPSIA LINFONODO SENTINELA (MS) .   | 114 |
| TABELA 57. RESULTADO DO INTERVALO DE 95% DE CONFIANÇA DA DIFERENÇA DE MÉDIAS DA DURAÇÃO DAS CIRURGIAS EM QUE SE REALIZOU OU NÃO A BIOPSIA LINFONODO SENTINELA (MS).....                                    | 114 |
| TABELA 58. INTERAÇÕES CONSIDERADAS NO MODELO DE REGRESSAO LINEAR (MS).....   | 114 |
| TABELA 59. NATUREZA DAS VARIÁVEIS ORIGINAIS E DE INTERAÇÃO CONSIDERADAS NO MODELO DE REGRESSÃO LINEAR (MS).....  | 115 |
| TABELA 60. QUANTIDADE DE CIRURGIAS POR INTERVALO DE IMC (SEG) .....  | 120 |
| TABELA 61. MÉDIA DE DURAÇÃO DE CIRURGIA PARA CADA INTERVALO DO IMC (SEG) .....   | 121 |
| TABELA 62. INTERVALOS DE CONFIANÇA SIMULTÂNEOS DE TUKEY (95% PARA A FAMÍLIA) DE DIFERENÇA DE MÉDIAS DOS INTERVALOS DE IMC (SEG) .....  | 121 |
| TABELA 63. PACIENTES HIPERTENSAS E NÃO HIPERTENSAS (SEG).....  | 122 |
| TABELA 64. MÉDIA DAS DURAÇÕES DE CIRURGIA DE PACIENTES HIPERTENSAS E NÃO HIPERTENSAS (SEG).....  | 122 |
| TABELA 65. PACIENTES DIABÉTICAS E NÃO DIABÉTICAS (SEG) .....   | 123 |
| TABELA 66. MÉDIAS DE DURAÇÃO DAS PACIENTES DIABÉTICAS E NÃO DIABÉTICAS (SEG) .....   | 123 |
| TABELA 67. CIRURGIAS REALIZADAS POR HOMENS E POR MULHERES (SEG) .....  | 124 |
| TABELA 68. MÉDIA DAS DURAÇÕES DE CIRURGIAS REALIZADAS POR HOMENS E MULHERES (SEG).....   | 125 |
| TABELA 69. RESULTADO DO INTERVALO DE 95% DE CONFIANÇA DA DIFERENÇA DAS MÉDIAS DE DURAÇÃO DE CIRURGIAS REALIZADAS POR CIRURGIÕES MULHERES E HOMENS (SEG) .....  | 125 |
| TABELA 70. CIRURGIAS PRATICADAS POR CIRURGIÕES QUE NASCERAM ANTES E DEPOIS DE 1960 (SEG).....  | 125 |
| TABELA 71. MÉDIAS DAS CIRURGIAS REALIZADAS POR CIRURGIÕES QUE NASCERAM ANTES E DEPOIS DE 1960 (SEG) .....  | 126 |
| TABELA 72. RESULTADO DO INTERVALO DE 95% DE CONFIANÇA DA DIFERENÇA DE MÉDIAS ENTRE AS CIRURGIAS REALIZADAS POR CIRURGIÕES QUE NASCERAM ANTES E DEPOIS DE 1960 (SEG) .....                                  | 126 |
| TABELA 73. CIRURGIAS REALIZADAS NOS DIFERENTES TRIMESTRES DO ANO (SEG).....  | 127 |
| TABELA 74. MÉDIA DE DURAÇÃO DAS CIRURGIAS REALIZADAS NOS QUATRO TRIMESTRES DO ANO (SEG) .....  | 127 |
| TABELA 75. RESULTADO DOS INTERVALOS DE CONFIANÇA SIMULTÂNEOS DE TUKEY (95% DE CONFIANÇA PARA A FAMÍLIA) DE DIFERENÇAS DAS MÉDIAS DAS DURAÇÕES DE CIRURGIAS REALIZADAS EM CADA TRIMESTRE DO ANO (SEG) ..... | 128 |
| TABELA 76. CIRURGIAS EM QUE O CIRURGIÃO STAFF PARTICIPA OU NÃO NA EQUIPE CIRÚRGICA (SEG) .....   | 128 |
| TABELA 77. MÉDIAS DAS DURAÇÕES DAS CIRURGIAS EM QUE O CIRURGIÃO STAFF PARTICIPOU OU NÃO NA EQUIPE CIRÚRGICA (SEG) .....  | 129 |
| TABELA 78. RESULTADO DO INTERVALO DE 95% DE CONFIANÇA DA DIFERENÇA DAS MÉDIAS DE CIRURGIAS EM QUE O CIRURGIÃO STAFF PARTICIPOU OU NÃO NA EQUIPE CIRÚRGICA (SEG) .....                                      | 129 |
| TABELA 79. CIRURGIAS QUE COMEÇARAM ANTES E DEPOIS DAS 12:00 H (SEG) .....  | 129 |
| TABELA 80. MÉDIAS DE DURAÇÃO DAS CIRURGIAS QUE COMEÇARAM ANTES E DEPOIS DAS 12:00 H (SEG).....   | 130 |
| TABELA 81. RESULTADO DO INTERVALO DE 95% DE CONFIANÇA DA DIFERENÇA DE MÉDIAS DAS DURAÇÕES DE CIRURGIAS QUE COMEÇARAM ANTES E DEPOIS DAS 12:00 H (SEG).....   | 130 |
| TABELA 82. QUANTIDADE DE CIRURGIAS REALIZADAS DE ACORDO COM A VARIAÇÃO DO PROCEDIMENTO (SEG) .....   | 131 |
| TABELA 83. MÉDIAS DE DURAÇÃO DAS CIRURGIAS DE ACORDO A CADA VARIAÇÃO DO PROCEDIMENTO (SEG).....  | 131 |

|  |     |
|--|-----|
| TABELA 84. RESULTADO DOS INTERVALOS DE CONFIANÇA SIMULTÂNEOS DE TUKEY (95% PARA A FAMÍLIA) DAS DIFERENÇAS DE MÉDIAS DE DURAÇÃO DA CIRURGIA DE ACORDO AS VARIAÇÕES DO PROCEDIMENTO (SEG)..... | 132 |
| TABELA 85. INTERAÇÕES CONSIDERADAS NO MODELO DE REGRESSAO LINEAR (SEG).....  | 132 |
| TABELA 86. NATUREZA DAS VARIÁVEIS ORIGINAIS E DE INTERAÇÃO CONSIDERADAS NO MODELO DE REGRESSÃO LINEAR (SEG) .....  | 132 |
| TABELA 87. LEMBRETE DAS EQUAÇÕES DE REGRESSÃO ESTIMADAS PARA CADA PROCEDIMENTO .....   | 136 |
| TABELA 88. RESULTADO DE APLICAR A EQUAÇÃO DE REGRESSÃO LINEAR ÀS INTERVENÇÕES MRM .....  | 138 |
| TABELA 89. RESULTADO DE APLICAR A EQUAÇÃO DE REGRESSÃO LINEAR ÀS INTERVENÇÕES MS .....   | 139 |
| TABELA 90. RESULTADO DE APLICAR A EQUAÇÃO DE REGRESSÃO LINEAR ÀS INTERVENÇÕES SEG .....  | 139 |
| TABELA 91. ERRO MÉDIO DE ESTIMAÇÃO DA PREVISÃO PARA OS TRÊS PROCEDIMENTOS.....   | 139 |

# CAPÍTULO 1: INTRODUÇÃO

---

Os processos cirúrgicos são cruciais dentro de um hospital. Não só do ponto de vista médico e social, mas também econômico: eles representam 70% das receitas e 40% dos custos [1]. Mais especificamente, a sala de cirurgia representa entre 10% e 15% dos custos hospitalares totais [2], levando em conta que cerca de 60% dos pacientes internados em hospitais têm algum tipo de relação com a área de cirurgia [3]. Isso se deve, entre outras coisas, ao envolvimento de pessoal qualificado nas atividades realizadas no interior das salas, ou ao uso intensivo e renovação de instrumentos e máquinas de alta tecnologia, com custos de aquisição e manutenção elevados.

Além disso, o que acontece em uma sala de operações determina a demanda de muitos recursos adicionais dentro do hospital, como a necessidade de leitos para recuperação de pacientes, enfermeiros, etc. [4]. Assim, do estudo da gestão dos recursos envolvidos na sala de cirurgia podem ser obtidos muitos benefícios potenciais no desempenho de um hospital, a partir de diversos pontos de vista.

## O projeto ENGATTI

O projeto ENGATTI (**Engenheiros em Grupo Autônomo para Transformar, Trocar e Integrar**) foi uma iniciativa surgida dos alunos de engenharia André Meirelles e Leonardo Luiz Lima Navarro no ano 2011.

O projeto foi motivado desde a percepção (com a qual, desde a minha experiência do ensino que recebi na Espanha, concordo plenamente) de que os alunos têm dificuldades para ter uma visão holística do que significa ser Engenheiro de Produção, o que leva a uma *crise de identidade* desta área da engenharia.

O principal objetivo dele é integrar os trabalhos de diferentes disciplinas do curso de Engenharia de Produção da UFRJ em uma instituição pública (incidindo aqui no seu caráter e motivação social), visando introduzir aos alunos à labor de campo (imprescindível no ensino em Engenharia) e fazendo interagir aos diferentes grupos de trabalho formados por estudantes da universidade.

A **minha experiência** dentro do projeto foi sumamente positiva por vários motivos. Em primeiro lugar, teve a oportunidade de enfrentar problemas em uma área, a Saúde, na qual os engenheiros de produção (pelo menos, no meu entorno) não percebem como uma possibilidade na qual atuar, e cuja importância social é inquestionável. Em segundo lugar, o trabalho de campo me permitiu entender que a resolução de problemas complexos dificilmente passa pela aplicação de uma única alternativa (relativa a uma disciplina de Engenharia de Produção), mas com a confluência de diferentes perspectivas (associadas a diferentes metodologias) desde as quais se pode observar qualquer sistema governado por pessoas. Finalmente, foi extremamente proveitoso para eu participar de algumas seções de discussão de trabalhos diferentes ao meu, mesmo que nas apresentações que periodicamente foram feitas a integrantes do hospital. Devo ressaltar nesse ponto a aprovação para a realização de testes para a confecção de células flexíveis dos recursos de hospital envolvidos até o primeiro tratamento do câncer de mama. Eu assisti à exposição do projeto à direção do HCIII, sendo uma experiência muito útil de como apresentar **com sucesso** soluções de certa complexidade a pessoas que não necessariamente estão familiarizadas com a Engenharia de Produção.



## **Justificativa**

Nos trabalhos apresentados em novembro de 2012 nas reuniões do Projeto ENGATTI entre a equipe do Departamento de Engenharia de Produção da UFRJ e os médicos do Hospital do Câncer III do INCA, foi detectado um elevado grau de variabilidade associada com a duração das cirurgias relacionadas com o tratamento de câncer da mama.

O estudo dos fatores que afetam a variabilidade dos procedimentos envolvidos na área de mastologia (remoção de tumor mamário) levará a estimativas mais precisas do tempo cirúrgico. Isto resultará na possibilidade de realizar programações das intervenções (através dos mapas cirúrgicos) que permitam aumentar a taxa de utilização das salas cirúrgicas, bem como a reduzir os riscos de atraso e cancelamento, e de realizar horas extras por parte da equipe médica (cirurgiões, anestesistas, enfermeiros, etc.).

## **Objetivos**

Os objetivos do presente projeto são:

1. Compreensão e descrição qualitativa dos principais procedimentos envolvidos na área de mastologia do HCIII do INCA.
2. Medição e análise da variabilidade (incerteza) associada à duração dos principais procedimentos cirúrgicos.
3. Estudo, análise e quantificação (expressa em termos de significância estatística) das diferentes fontes de variação detectadas.
4. Construção de um modelo de regressão linear que permita (através dos fatores de variação que resultem significativos) uma melhor predição da duração das cirurgias.

## **Estruturação do trabalho**

O presente trabalho está estruturado da seguinte forma:

No primeiro capítulo se descreve a **introdução**, a **justificativa** e os **objetivos** do trabalho.

No segundo capítulo, dedicado ao **marco teórico**, se apresentam as características básicas da programação das salas de cirurgia em um hospital. Seu principal objetivo é aproximar as implicações da incerteza associada à duração das cirurgias no processo de programação das salas.

No terceiro capítulo, dedicado à apresentação das **principais ferramentas estatísticas utilizadas no estudo da variabilidade**, se mostram as metodologias que têm sido utilizadas para estudar a duração das diferentes cirurgias analisadas. Seu principal objetivo é estabelecer a base teórica do método desenvolvido no estudo de caso.

No quarto capítulo, dedicado ao **Estudo de Caso**, foi analisada a duração dos procedimentos Mastectomia Radical Modificada (MRM), Mastectomia Simples (MS) e Segmentectomia (SEG). Seu principal objetivo é estabelecer uma equação de previsão, através da Regressão Linear Múltipla, para estimar com maior precisão a duração das intervenções.

No quinto capítulo, dedicado às **conclusões**, se apresentam os principais resultados e se realiza uma avaliação crítica deles, bem como propondo trabalhos futuros na linha de projeto.

Finalmente, o sexto capítulo inclui as **referências bibliográficas**.

# CAPÍTULO 2: MARCO TEÓRICO

Nesse capítulo se apresentam as características básicas da programação das salas de cirurgia em um hospital. Seu principal objetivo é aproximar as implicações da incerteza associada à duração das cirurgias no processo de programação das salas.

## 2.1 Cirurgia

Uma intervenção cirúrgica pode ser definida como uma ou mais ações operacionais com a mesma finalidade terapêutica ou de diagnóstico, feita (s) por cirurgião(s) na sala de cirurgia, na mesma sessão, com anestesia geral, regional ou local, com ou sem presença de anestesista [5].

Além disso, uma cirurgia pode ser eletiva ou de emergência. As cirurgias eletivas são aquelas realizadas a pacientes que não estão em risco iminente de complicações para a sua saúde. A cirurgia de urgência, por outro lado, é efetuada quando é comprometida a função do órgão ou a vida do paciente.

O curso de um paciente que se submete a uma cirurgia dentro de um hospital envolve muitas etapas que variam de acordo com cada hospital e unidade cirúrgica. No entanto, uma boa aproximação pode ser vista no gráfico seguinte:

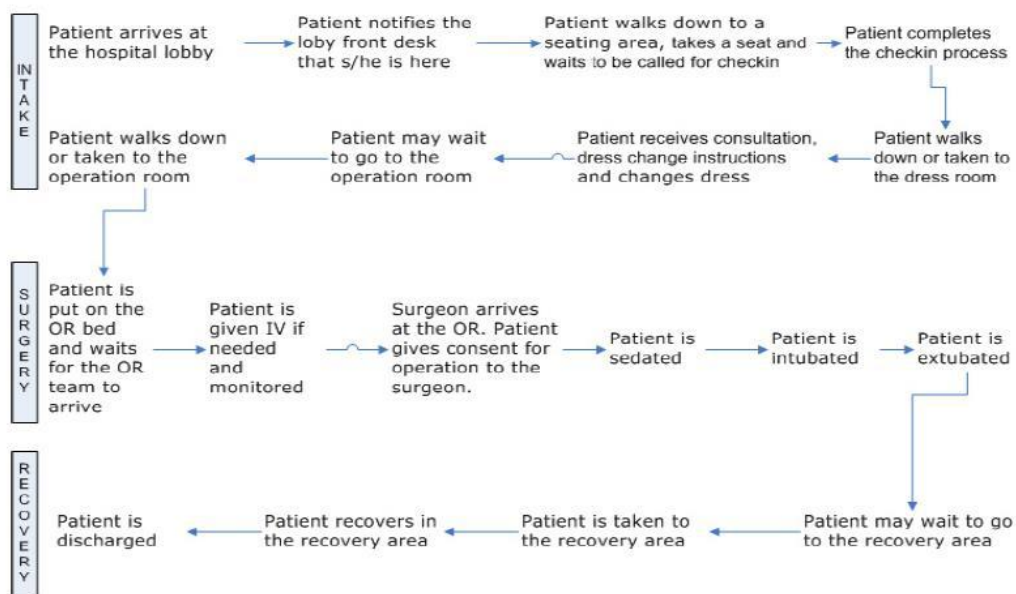


Figura 1. Exemplo gráfico do processo de cirurgia. Extraído de [6].

Na fase de entrada (INTAKE) o paciente é dirigido para a sala de cirurgia. Na fase de cirurgia (SURGERY) é praticado o procedimento cirúrgico planeado. Finalmente, na fase de recuperação (RECOVERY) é submetido a um tratamento de reabilitação pós-operatória.

Da figura 1, apesar de ser esquemática, já se pode deduzir que há muitos recursos envolvidos no processo, tais como cirurgiões, anestesistas, enfermeiros de apoio à cirurgia, equipamentos especializados, leitos na área de recuperação, enfermeiros que cuidam de pacientes já operados, etc.

## 2.2 Taxa de utilização da sala de cirurgia

Uma das medidas de desempenho mais importantes de uma sala de operações, em termos de planejamento, é a sua taxa de utilização. Pode ser definida como segue:

$$taxa\ utilização[\%] = \frac{t_{util}}{t_{OR}} = \left( \frac{t_{op} + t_{setup} + t_{prep}}{t_{OR}} \right) * 100$$

em que:

- $t_{OR}$ : tempo em que a sala de cirurgia permanece aberta (8h, 12h, etc.)
- $t_{op1}$ : tempo em que o paciente está dentro da sala de cirurgia (para uma eventual operação)
- $t_{setup}$ : tempo de limpeza e preparação da sala de cirurgia entre operações.
- $t_{prep}$ : tempo necessário no princípio ou final de uma jornada cirúrgica para acondicionar a sala.

É uma magnitude que reflete claramente a qualidade do processo de programação das salas, e a partir da qual, e conhecendo em profundidade o procedimento específico em estudo, se podem dimensionar oportunidades de melhoria e avalia-las em uma eventual implementação ou simulação.

Para ilustrar as perdas de tempo útil que se produzem dentro de uma sala de cirurgia em um dia, se mostra um exemplo na figura 2.

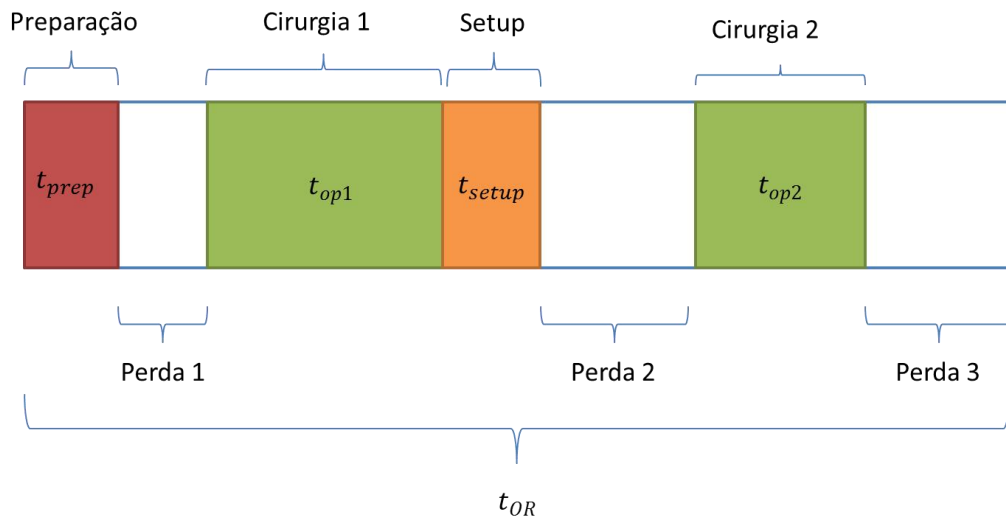


Figura 2. Esquema de tempos de operação, tempo de preparação, setup e perdas em um dia de uma sala de cirurgia. Elaboração própria.

Da figura 2 pode-se definir o cálculo das perdas totais como,

$$t_{inatividade} = t_{perda1} + t_{perda2} + t_{perda3}$$

Também, pode-se expressar o tempo útil de utilização da sala como,

$$t_{util} = t_{OR} - t_{inatividade}$$

e a taxa de utilização se pode expressar em função do tempo de inatividade dela como,

$$taxa\ utilizacao\ [\%] = \left( \frac{t_{OR} - t_{inatividade}}{t_{OR}} \right) * 100 = \left( 1 - \frac{t_{inatividade}}{t_{OR}} \right) * 100$$

É claro, portanto, que maior é o tempo de inatividade da sala, menor será a taxa de utilização.

A figura 3 mostra uma situação hipotética em que, produto de uma boa programação das salas, se reduziu o tempo de inatividade da sala de cirurgia em uma jornada cirúrgica.

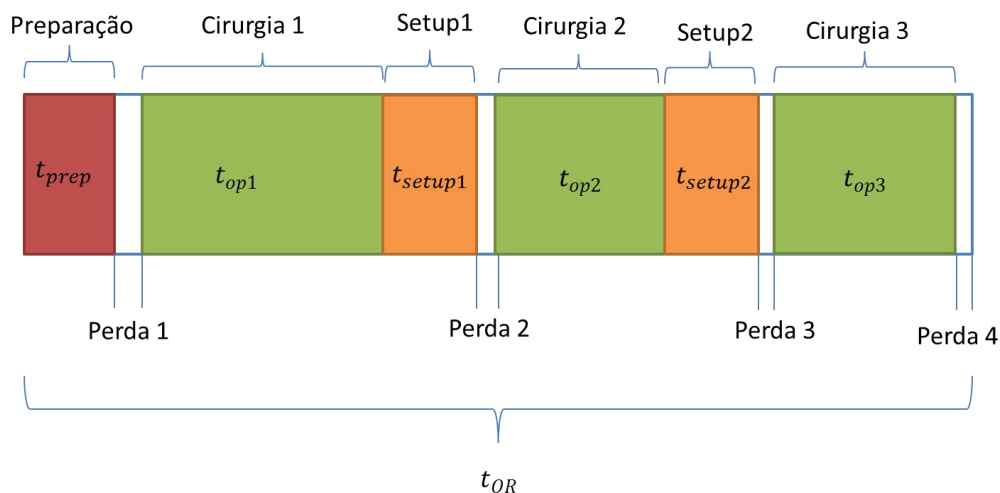


Figura 3. Programação da sala melhorada. Elaboração própria.

Observa-se que, com respeito ao exemplo da figura 2, o fato de ter minimizado as perdas e, portanto, o tempo de inatividade da sala de cirurgia, gera a possibilidade de incluir uma cirurgia mais no mesmo tempo diário disponível da sala de cirurgia.

Podem-se observar aqui os benefícios potenciais de atingir uma taxa de utilização elevada na sala cirúrgica de um hospital: uma maior taxa de utilização implica em um maior número de cirurgias em um mesmo espaço de tempo (por exemplo, um ano), o que significa que um maior número de pacientes vá poder ser operado. Além disso, o fato de realizar um maior número de operações com os mesmos recursos associados ao hospital vai originar um maior grau de sustentabilidade económica dele.

A melhoria, em consequência, da taxa de utilização das salas de cirurgia traz consigo importantes benefícios, e deve ser um dos objetivos prioritários de uma boa programação das salas de cirurgias. Porém, como se verá nos capítulos a seguir, atingir um aumento da taxa de utilização envolve, pelas características específicas das operações cirúrgicas, riscos e complicações que incorporam complexidade às análises enfocadas nesse sentido.

A *American Hospital Association* afirma que uma utilização aceitável de qualquer sala de cirurgia deve ser igual ou superior a 75%, e atingir taxas acima de 80% requer o uso de sistemas de apoio excelentes [7].

As causas mais importantes de baixa utilização são os cancelamentos ou atrasos no início da cirurgia [8], e as mais importantes razões para o cancelamento são [9]:

- Coordenação de cada especialidade cirúrgica.
- Prolongamento de cirurgias prévias

Em relação ao prolongamento de cirurgias anteriores, ele bem condicionado por um fenômeno que será objeto de estudo neste trabalho: **a incerteza em prever a duração da cirurgia**. A incapacidade de prever com exatidão o tempo de duração de cada procedimento influencia decisivamente a precisão e a eficiência do processo de programação da sala de cirurgia (*scheduling*), o que resulta em uma redução da taxa de utilização.

### **2.3 A incerteza**

A incerteza pode ser definida como a diferença entre o volume de informação necessária para realizar uma tarefa, e a quantidade de informação que se possuía antes de realizá-la [10].

Na vida real, existem muitas fontes de incerteza que dificultam o planejamento eficaz de um processo e, na maioria dos casos, são impossíveis de remover.

Nomeadamente, os processos que regem a área da medicina muitas vezes carregam consigo um grau de incerteza inerente particularmente elevado [11], sendo uma das causas que explicam as poucas experiências de sucesso, em comparação com outros setores como a indústria, de implementação de modelos de pesquisa operacional nos processos de programação.

A figura 4 ilustra um exemplo genérico dos problemas de estimar a duração de um processo padrão, do qual se possui certo grau de incerteza.



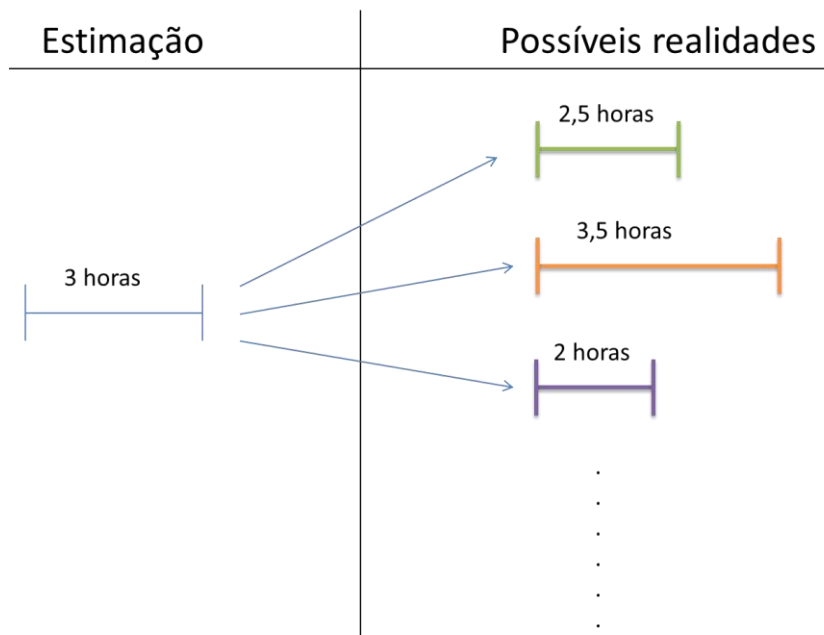


Figura 4. Correspondência entre a estimaco e a realidade na duraco de um processo. Elaboraco prpria.

Note-se que um elevado nvel de incerteza produz uma s estimaco da duraco do processo para uma realidade que  muito **varivel**. At mais, o fato de no conhecer os motivos dessa variabilidade, limita as possibilidades de prever com preciso (estimaco) a duraco real do processo. Este cenrio  crtico para os objetivos da programaco, considerando que seu escopo principal , genericamente, estabelecer duraces tentativas das cirurgias nas salas de cirurgia do hospital, para construir os horrios do mapa cirrgico.

## 2.4 Programaco das salas de cirurgia (*Scheduling*)

O termo *scheduling* pode ser definido como o processo de tomada de deciso que visa responder s questes de **como**, **onde** e **quando** realizar as operaes (no sentido amplo do termo) em um determinado centro de produo.

Mais especificamente, o objetivo do *scheduling*  a programaco de uma srie de trabalhos (*Jobs*) em mquinas (*Machines*), que esto sujeitas a um conjunto de restries, de modo que uma funo objetivo deve ser otimizada [12].

Basicamente existem dois tipos de *scheduling*: *Flow Shop Scheduling* e *Job Shop Scheduling*.

Os problemas *Flow Shop Scheduling* podem ser descritos como um conjunto de  $n$  trabalhos que devem ser associados a um conjunto de  $m$  máquinas.

Neste caso, não há restrições de precedência entre as operações dos diferentes trabalhos. As operações não podem ser interrompidas, e cada máquina pode processar apenas uma operação de cada vez. De modo que se precisam encontrar as seqüências de trabalhos nas máquinas que minimizam o tempo total de realização de operações (*makespan*).

A figura 5 mostra o processo de produção em Flow Shop.

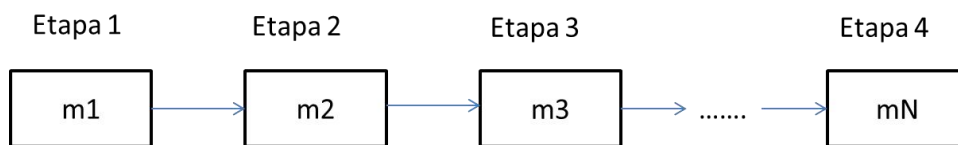


Figura 5. Esquema de Flow Shop clássico. Elaboração própria.

Percebe-se que todos os trabalhos têm a mesma rota tecnológica. Na sua extensão *Flow Shop flexível* são consideradas máquinas diferentes em paralelo em uma ou mais etapas.

No caso do *Job-shop Scheduling*, as rotas tecnológicas de cada trabalho variam, e eventualmente as operações se repetem em uma ou mais máquinas. A figura 6 mostra um diagrama básico esquemático de *Jop-Shop Scheduling*.

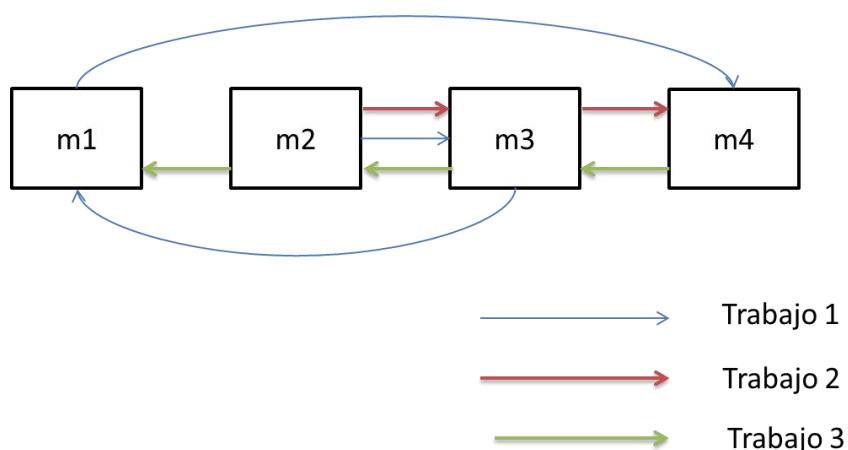


Figura 6. Job-Shop Scheduling clássico. Elaboração própria.

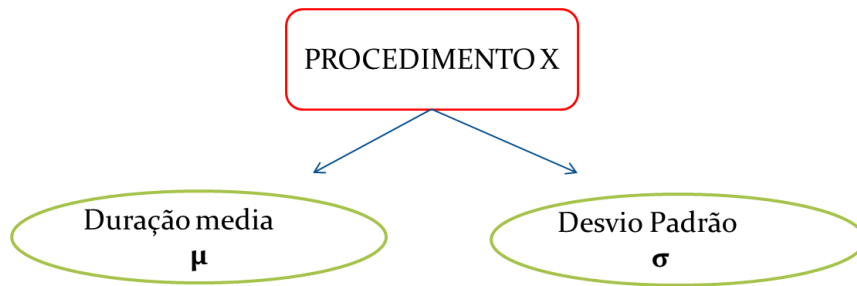
Note-se que cada trabalho possui a sua própria rota. Como no *Flow Shop*, se podem incluir máquinas em paralelo em cada etapa, e nesse caso se fala de *Job-Shop flexível*. O *Job-Shop Scheduling* é um problema NP-hard; ou seja, sua resolução por métodos exatos não é possível em tempo polinomial, por isso é frequentemente abordado desde perspectivas heurísticas.

#### **2.4.1 A incerteza no processo de planejamento das operações**

Tradicionalmente, na indústria, o *Scheduling* de um centro de produção foi formado a partir de um ponto de vista determinístico. É comum ignorar a variabilidade inerente associada a qualquer processo físico, pois a grande complexidade que incorporam os modelos estocásticos de planejamento e otimização das operações não justificam, em termos de benefícios obtidos, a sua construção, análise, resolução e implementação.

No entanto, um elemento de diferenciação clara entre as salas de cirurgia de um hospital e os centros de produção industrial habituais, é que as operações têm uma forte componente humana e uma duração média relativamente elevada (geralmente, entre 1 e 5 horas). O limite tecnológico existente na maioria das cirurgias de certa complexidade, que faz com que seja difícil (ou mesmo impossível) a sua automatização completa, bem como o número de decisões que deve tomar um eventual cirurgião, ou as características diferentes entre pacientes, impõem a existência de uma variabilidade intrínseca a uma cirurgia padrão, o que traz um alto nível de incerteza quanto à sua duração prevista.

A figura 7 mostra as consequências e implicações da programação deficiente de uma cirurgia típica, produto da incerteza que se tem da duração real do procedimento cirúrgico.



| Cenário | Duração Programada | Duração real | Consequência                         | Implicações gerais  |
|---------|--------------------|--------------|--------------------------------------|---|
| 1       | 1 hora             | 1:30 horas   | Seguinte operação retardada 30 mins. | <ul style="list-style-type: none"> <li>- Horas extra pessoal médico.</li> <li>- Risco de cancelamento da seguinte cirurgia</li> </ul>   |
| 2       | 2 horas            | 1:30 horas   | Ociosidade 30 mins.                  | <ul style="list-style-type: none"> <li>- Baixa taxa de utilização.</li> <li>- Menor número de pacientes operados.</li> </ul>  |
| 3       | 1:30 horas         | 1:30 horas   | -                                    | <ul style="list-style-type: none"> <li>- Taxa de utilização da sala elevada.</li> <li>- Elevado número de pacientes operados..</li> <li>- Não necessidade de horas extra.</li> <li>- Baixo risco de cancelamento</li> </ul> |

Figura 7. Consequências e implicações dos diferentes cenários resultantes da incerteza do programador das salas respeito às durações de um procedimento tipo. Elaboração própria.

Um programador que possui elevada incerteza à respeito da duração real das cirurgias gerará programações (*schedulings*) situadas nos cenários 1 e 2; o que implicará riscos de cancelamento das cirurgias e necessidade por parte do pessoal médico de realizar horas extraordinárias (cenário 1), ou taxas de utilização baixas e baixo número de pacientes operados (cenário 2).

Por outro lado, um eventual programador que conheça (tem baixa incerteza) a duração real das cirurgias com antecedência à sua execução, poderá gerar *schedulings* situados no cenário 3, o que implicará uma necessidade baixa de realizar horas extra do pessoal médico, baixo risco de cancelamento, altas taxas de utilização da sala de cirurgia e, portanto, elevado número de pacientes operados.

Assim, uma possível solução (e escopo do presente projeto) para o programador é adquirir um maior conhecimento dos **fatores que explicam a variabilidade** associada à duração das cirurgias, o que permitirá gerar programações baseadas em durações previstas dos procedimentos cirúrgicos que terão maior correspondência com a

realidade da sua execução posterior, o que comportará todos os benefícios descritos no cenário 3.

#### **2.4.2 Metodologias de Scheduling sob incerteza**

Dentro do processo de Scheduling se podem classificar dois tipos de planejamento, dependendo de quando ele é feito. O *Preventive Scheduling* é aquele que determina o tempo de início e duração das atividades **antes** da sua execução, e atua como um mecanismo de planejamento puro com base nos requisitos e critérios de estruturação de um centro de produção. Por outro lado, o *Reactive Scheduling* é construído como uma ferramenta para a revisão da programação inicial, que visa reagir, mais ou menos eficazmente, à realização de eventos inesperados **durante** a execução das atividades [13].

A seguir, apresenta-se uma breve descrição sobre os dois tipos de planejamento.

##### **Reactive Scheduling**

No campo industrial, o *Reactive Scheduling* lida com operar sob dois tipos de incerteza: a relacionada com a realização de eventos inesperados decorrentes de fatores que modificam os tempos de processamento das atividades, e aquelas relacionadas a cancelamentos e / ou flutuações na demanda.

O principal objetivo é atualizar a programação original e reagir imediatamente, como discutido acima, aos eventos inesperados que não são tidos em conta na fase de programação inicial. O *Reactive Scheduling* opera no momento certo quando as atividades são realizadas.

Outro nome dado a estas técnicas para reagir à incerteza no campo do planejamento de operações é **Reescheduling**.

## **Preventive Scheduling**

### **(i) Programação Estocástica**

Na programação estocástica, o scheduling inicial (que na programação determinística não leva em conta as futuras variações do modelo original) considera as possíveis incertezas do modelo como variáveis estocásticas.

Neste caso, o objetivo é otimizar um modelo de programação das operações com respeito a um critério de desempenho, como o tempo total de duração ou *makespan*, levando em conta variáveis aleatórias.

Embora a programação inicial se desenhe contemplando variáveis aleatórias, é um problema que exige uma fase de programação reativa (*two-stage* ou *multi-stage*) devido à realização de eventos inesperados. Mesmo que inicialmente tenham sido considerados em uma primeira programação, pela sua natureza estocástica, é necessária uma revisão *real-time* perante comportamentos operacionais imprevistos.

De modo que se estrutura otimizando o sequenciamento das atividades *sob o cenário mais provável*, mas que, pela própria definição de probabilidade e incerteza, exige atenção e planos de ação após o planejamento inicial.

Esta é a abordagem mais comum na literatura de pesquisa operacional e inclui técnicas como a programação linear inteira mista com variáveis estocásticas.

### **(ii) Otimização Robusta**

Desde a perspectiva da modelação de processos, a robustez pode ser definida como a capacidade de um modelo de permanecer estável a perturbações estocásticas.

O *scheduling* robusto baseia-se em construir um modelo preventivo tentando minimizar o *gap* entre o planejamento inicial e a realização efetiva das operações.

Neste caso, se pondera o fato de ter de fazer as mínimas alterações posteriores em um eventual *scheduling* de caráter reativo; ou seja, se aumenta a probabilidade de correspondência entre o planejamento original e a execução das operações.

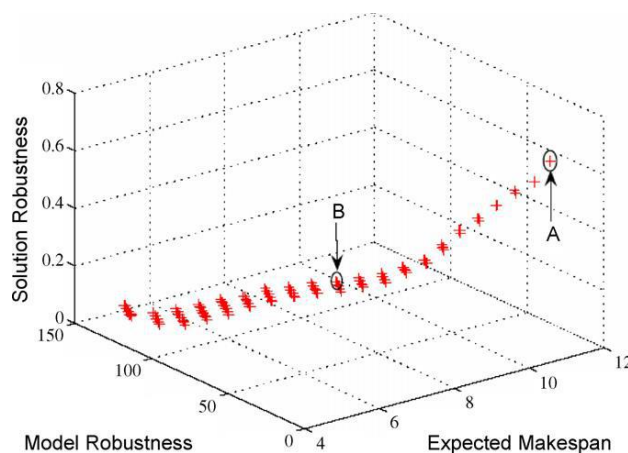
Do ponto de vista prático, exemplos de ações que aumentem a robustez da programação da sala de cirurgia são a adição no final do horizonte de planejamento de *buffers* temporários onde se operam aos pacientes que não se puderem submeter à cirurgia (cancelamento) ou a imputação de grupos de pacientes com cirurgias associadas com alta variabilidade a uma única sala de cirurgia (para controlar as cirurgias eletivas com baixa incerteza em uma sala de cirurgia diferente).

No entanto, deve notar-se que existe um contraste claro entre robustez e eficiência da sala de cirurgia (entendida como o grau de utilização, em tempo e recursos). Introduzir muitos espaços vazios de tempo na expectativa de qualquer cancelamento ou extensão imprevista de cirurgias (produto da incerteza inerente que o programador possui sobre os diferentes procedimentos) produz um risco de perda excessiva de capacidade da sala de cirurgia, com a perda de eficiência subsequente.

Por outro lado, se não se programam suficientes intervalos de tempo em previsão de cancelamentos ou outra perturbação estocástica das cirurgias, se pode incorrer em cancelamentos permanentes (acrescentando a lista de espera para a sala de cirurgia) e / ou a necessidade de realizar horas extraordinárias pela equipe médica [3].

Se fala aqui de um problema de otimização multicritério, no qual existem inclusive experiências de modelagem da região Pareto-ótima em termos de robustez da solução (a capacidade solução para permanecer próxima do ótimo em diferentes cenários), duração total de cirurgias (*makespan*) e robustez do modelo (capacidade de um modelo para ser flexível a diferentes cenários).

A seguir se pode observar um gráfico extraído de [13]



**Figura 8. Superfície Pareto-ótima considerando a duração total no horizonte de programação das operações (*makespan*), a robustez da solução e a robustez do modelo. Extraído de [13].**

Note-se que à medida que aumenta o tempo total da programação planeada (makespan) é mais provável que o modelo se comporte conforme previamente planeado (“model robustness”) e, além disso, no caso de ocorrerem perturbações na execução das operações (e o *scheduling* reativo tivesse de reagir modificando a planificação inicial), as soluções a adotar poderiam resultar próximas ao ótimo (“solution robustness”).

Um exemplo de multi-objetivo, neste caso, pode ser formulado como aquele em que a solução passa por otimizar a alocação das cirurgias de modo **a minimizar o risco de trabalho suplementar, o cancelamento não ocorrer, enquanto que a eficiência da sala de cirurgia é máxima**. As soluções propostas da pesquisa operacional muitas vezes passam através da construção de modelos de programação estocástica e, devido à complexidade do cálculo de tais modelos, a resolução é abordada através da aplicação de heurísticas ou metaheurísticas [14].

Finalmente, é preciso dizer que é neste ponto onde se pode apreciar mais claramente a influência prejudicial que causa incerteza (manifestada através da variabilidade não explicada) que se possui do tempo de duração das diferentes cirurgias. Quanto maior a variabilidade associada à variável aleatória "tempo de cirurgia," maior o intervalo de tempo necessário para absorver possíveis perturbações estocásticas do modelo de otimização robusta e, por conseguinte, maior perda de capacidade e taxa de utilização da sala de operações.

### **(iii) Lógica Fuzzy**

A lógica Fuzzy (em Inglês, *Fuzzy Logic*) é baseada na captura de informação para a tomada de decisões e, através de heurísticas de programação, estabelecer conjuntos difusos que contem elementos parciais; ou seja, um elemento pertence a um conjunto difuso com um *grau* de verdade.

Respeito da programação estocástica, o uso da lógica Fuzzy (a qual pode ser utilizada para criar mecanismos de aprendizagem, tais como as redes neurais) tem a característica de não necessitar de informação prévia sobre os dados de entrada (distribuição de probabilidade, variabilidade, etc.).



#### (iv) Análise de Sensibilidade

Considerando a elevada complicação que envolve a resolução de modelos MILP (ou outros de ainda maior complexidade computacional) com variáveis aleatórias, o uso rigoroso e sistemático de uma ferramenta de sensibilidade pode ser muito útil. A Análise de Sensibilidade é uma metodologia poderosa quando se quer aumentar a robustez de um modelo. É possível identificar os elementos importantes de um modelo alterando o rango das variáveis envolvidas nele. O objetivo é estudar seu desempenho quando submetido a variações controladas, procurando identificar quais são as características que o tornam robusto; ou seja, estudar desde que perspectiva o modelo é *sensível* a alterações às quais, quase certamente, se vai ver submetido quando ocorram eventos inesperados durante sua execução real.

#### 2.4.3 Fases para a programação das operações de um centro cirúrgico

Basicamente, se podem definir três níveis de planejamento do centro cirúrgico de um hospital, dependendo do horizonte de tempo a que se referem. Esta estrutura é conhecida como a **decomposição hierárquica** [5].

Em um primeiro plano, conhecido como *case mix planning*, o tempo disponível na sala de operações é distribuído entre os diferentes cirurgiões, levando em conta o dimensionamento da capacidade das salas de cirurgia. Contextualiza-se em um espaço de tempo grande (geralmente o horizonte é de um ano), e é considerado um nível de planejamento **estratégico** das salas de operações.

Nesta fase, o principal problema é decidir quantos pacientes vão ser operados no curso do ano (pode ser um horizonte mais longo), baseando-se principalmente em dois critérios: a satisfação de níveis de benefício prefixados ou a maximização dos benefícios do hospital [15], dependendo do porte de cada organização (hospital público, privado, etc.). Nesta etapa, se determina a receita esperada que vá receber o hospital, e se estabelecem os recursos que devem ser consumidos para atender a demanda de pacientes a satisfazer.

Crítérios de otimização no *Case Mix Planning* são a maximização dos benefícios ou a maximização do número de pacientes para operar com um benefício pré-determinado a ser alcançado (fixado pelo grau de sustentabilidade da instituição).

Em geral, o tempo disponível nas salas de cirurgia se associa a um particular cirurgião, ou equipa cirúrgica, em forma de **bloco**. Um bloco pode ser definido como a

menor unidade de tempo de uma sala de operações que pode ser ligada a um cirurgião particular, como pode ser de 4 horas, 7 horas, etc. [4].

Este arranjo tem as seguintes vantagens [16]

- O cirurgião sabe com semanas ou meses de antecedência o dia da semana que deve operar para que possa conciliar o momento da cirurgia com outros compromissos médicos (como consultas médicas, etc.).
- Sabe quantas salas de cirurgia lhe foram atribuídos e pode agendar com suficiente antecedência aos pacientes e ao pessoal médico necessário para a cirurgia (enfermeiros, residentes, etc.).
- Do ponto de vista do hospital, a organização em blocos permite alocar a força de trabalho dos cirurgiões e recursos associados de forma equitativa durante a semana ou até horizontes temporais mais longos.
- A Distribuição em blocos também é útil para alocar aos pacientes mais uniformemente na UTI.

No âmbito **tático** do planejamento das cirurgias se encontra a fase do *Master Surgery Planning* ou *Master Surgery Schedule* (equivalente, em termos industriais, ao plano mestre de produção). Nessa fase se determina quando e em que quantidade as equipes cirúrgicas terão disponibilizadas as salas de cirurgia para realizar as operações atribuídas. O objetivo geral é fazer corresponder os blocos com a demanda do paciente, como foi estabelecido no *Case Mix Planning*. Durante esta etapa são explicitadas as salas disponíveis para realizar as operações, o número de horas que estarão abertas, associando cirurgiões ou grupos de cirurgiões a cada uma delas, mesmo estabelecendo critérios de prioridade de grupos de pacientes com características homogêneas [14].

Crerios de otimização nesta fase podem ser a minimização da demanda pico de camas ou a homogeneização da distribuição de horas de trabalho.

A terceira fase, de natureza **operacional**, é chamada *Elective Case Scheduling*. Nela se planifica cada operação a partir de uma base semanal e diária. Baseia-se no sequenciamento dos pacientes que vão ser submetidos a cirurgias eletivas, e são designados cada um dos blocos de cirurgiões. Em geral é resolvido em duas etapas: a primeira delas (*Advanced Scheduling*) refere-se à alocação dos pacientes em uma sala de cirurgia, enquanto na segunda (*Allocation Scheduling*) são sequenciadas as operações dentro da sala para otimizar a sua eficiência, levando em conta as limitações dos recursos associados com cada operação cirúrgica [2].

Critérios de otimização nesta fase podem ser a maximização da taxa de utilização da sala de cirurgia ou a minimização das horas suplementares a serem realizadas pelo pessoal médico.

A tabela 1 resume os níveis e objetivos de planejamento, e exemplos de otimização.

**Tabela 1. Resumo dos níveis de planejamento, objetivos do planejamento e exemplos de critérios de otimização.**

| Nível de planejamento              | Case Mix Planning  | Master Surgery Planning   | Elective Case Scheduling  |
|------------------------------------|--|---|---|
| <b>Horizonte temporal (aprox.)</b> | 1 ano  | 1 mês   | 1 semana/dia  |
| <b>Objetivos</b>                   | Estabelecer o número de pacientes para operar e distribuir o tempo sala de cirurgia entre cirurgiões (ou grupos) | Atribuir os blocos de cirurgias para atender a demanda de pacientes determinada no <i>Case Mix Planning</i> .   | Sequenciar aos pacientes dentro de cada bloco cirúrgico, atribuindo-os a uma sala de cirurgia.  |
| <b>Critérios de Otimização</b>     | <ul style="list-style-type: none"> <li>Máximo benefício.</li> <li>Máximo número de pacientes.</li> </ul>         | <ul style="list-style-type: none"> <li>homogeneização da utilização dos recursos materiais (leitos, maquinaria, etc.)</li> <li>homogeneização da força de trabalho no horizonte temporal estabelecido.</li> </ul> | <ul style="list-style-type: none"> <li>Maximização da taxa de utilização das salas de cirurgia.</li> <li>Maximização das horas extraordinárias a realizar por parte do pessoal médico.</li> </ul> |

#### 2.4.4 Ferramentas de pesquisa operacional aplicadas ao planejamento da sala de cirurgia.

Do ponto de vista da pesquisa operacional, e levando em conta as diferentes fases do planejamento da sala de cirurgia descritas acima, existem diferentes abordagens na construção de modelos de otimização.

A tabela 2 mostra os tipos de modelos utilizados na literatura, sua natureza (estocástica ou determinística) e as ferramentas tomadas na solução.

**Tabela 2. Resumo de modelos de otimização, natureza dos modelos e abordagem da solução.**  
Dados extraídos de [5].

| Fase da Planificação | Case Mix Planning  | Master Surgery Planning   | Elective Case Scheduling   |
|----------------------|--|---|--|
| Modelos              | <ul style="list-style-type: none"> <li>• Programação linear</li> <li>• Programação linear Inteira</li> </ul> | <ul style="list-style-type: none"> <li>• Programação linear Inteira</li> <li>• Programação linear Inteira Mista.</li> <li>• Programação Quadrática Mista.</li> <li>• Programação Inteira Estocástica</li> </ul> | <ul style="list-style-type: none"> <li>• Programação linear Inteira</li> <li>• Programação linear Inteira Mista.</li> <li>• Programação Inteira Estocástica.</li> <li>• Programação dinâmica</li> <li>• Programação Linear Estocástica</li> <li>• Simulação</li> </ul> |
| Natureza             | Determinística   | <ul style="list-style-type: none"> <li>• Determinística</li> <li>• Estocástica</li> </ul>   | <ul style="list-style-type: none"> <li>• Determinística</li> <li>• Estocástica</li> </ul>  |
| Abordagem da Solução | Métodos exatos   | <ul style="list-style-type: none"> <li>• Métodos exatos</li> <li>• Heurísticas</li> <li>• Meta-heurísticas</li> </ul>   | <ul style="list-style-type: none"> <li>• Métodos exatos</li> <li>• Heurísticas</li> <li>• Meta-heurísticas</li> <li>• Simulação</li> </ul>   |

# CAPÍTULO 3: FERRAMENTAS DE ANÁLISE ESTATÍSTICA NO ESTUDO DA VARIABILIDADE

---

Do estudo da incerteza lida à análise estatística. A seguir se apresentam os instrumentos utilizados para a análise da duração das cirurgias.

## **3.1 Caracterização dos dados. Exigência de normalidade.**

Um dos pressupostos básicos de muitos métodos estatísticos convencionais é a normalidade dos dados. Muitas vezes, tais como na análise da variância ou a regressão linear, uma das características inevitáveis que deve atender a amostra de referência é que as observações que a compõem possuem uma distribuição aproximadamente normal.

No entanto, o fato de o tempo de operação ser uma variável aleatória estritamente positiva, implica que existe certa assimetria (bias) na sua distribuição de probabilidade, o que dificulta o seu tratamento como uma variável gaussiana.

Uma solução possível é transformar os dados de modo a que a distribuição da variável de resposta transformada se aproxime a uma distribuição normal.

### **Transformação de Dados**

A literatura contém vários tipos diferentes de transformações de dados [17]. Em particular, é apresentado um método amplamente utilizado: a transformação Box-Cox. Pela sua simplicidade, eficiência e facilidade de interpretação da transformação das variáveis originais, é uma opção muito atrativa para o tratamento inicial de dados empíricos.

A transformação de Box-Cox é a transformação mais amplamente utilizada em análises estatísticas. Ela consiste de uma transformação potência da variável resposta tal que,

$$y(\lambda) = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \text{si } \lambda \neq 0 \\ \ln(\lambda), & \text{si } \lambda = 0 \end{cases}$$

com  $y$  estritamente positiva.

Como transformações lineares não afetam à análise da variância (metodologia chave na análise estatística), na prática, a transformação mostrada acima pode ser simplificada de modo que [18],

$$y(\lambda) = \begin{cases} y^\lambda - 1, & \text{si } \lambda \neq 0 \\ \ln(\lambda), & \text{si } \lambda = 0 \end{cases}$$

Esta transformação determina o valor do parâmetro  $\lambda$  que minimiza o erro quadrado médio da variável transformada com respeito à distribuição normal teórica.

Em geral, são consideradas as variações de  $\lambda$  dentro do intervalo  $[-2, 2]$ .

A tabela seguinte mostra as diferentes transformações indicadas pelo parâmetro  $\lambda$ .

**Tabela 3. Transformações de Box-Cox**

|      | Transformação $y(\lambda)$        |
|------|-----------------------------------|
| 2    | $y(\lambda) = y^2$                |
| 1    | $y(\lambda) = y$                  |
| 0,5  | $y(\lambda) = \sqrt{y}$           |
| 0    | $y(\lambda) = \ln(y)$             |
| -0,5 | $y(\lambda) = \frac{1}{\sqrt{y}}$ |
| -1   | $y(\lambda) = \frac{1}{y}$        |
| -2   | $y(\lambda) = \frac{1}{y^2}$      |

### **Bondade de ajuste ("Goodness of fit")**

Para avaliar o grau em que os dados se ajustam às distribuições consideradas na análise de inferência, se utilizam um conjunto de ferramentas destinadas a analisar a qualidade do ajuste do modelo.

### **Gráfico de probabilidades**

Um instrumento visual muito utilizado é o gráfico de probabilidades. Baseia-se na linearização da função de densidade acumulada de uma determinada distribuição através da transformação da escala dos seus eixos de coordenadas (por exemplo, a escala logarítmica, no caso da log-normal).

Constrói-se uma linha reta que representa o comportamento da distribuição cumulativa teórica dos dados. Mais tarde, as observações da amostra obtida são representadas. Se os dados empíricos se aproximam razoavelmente à reta da distribuição teórica, significa que a distribuição empírica se ajusta à distribuição teórica, na ausência de uma confirmação quantitativa da hipótese, na distribuição considerada, e é uma boa candidata a levar em conta. Por outro lado, se os dados da amostra não são sobrepostos na reta da distribuição hipotética, é um indicativo de falta de ajuste, e se deve escolher outra distribuição de probabilidade.

A seguir são mostradas duas figuras. A primeira delas consiste na representação de dados de uma amostra hipotética em gráfico de probabilidades normal, no qual o pressuposto de normalidade da amostra não parece razoável. No segundo caso, o pressuposto de normalidade parece ser justificável.

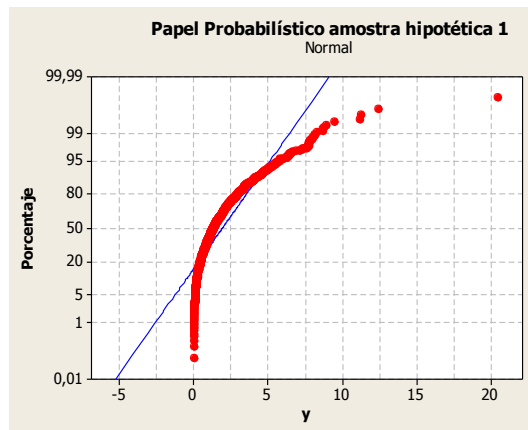


Figura 9. Papel probabilístico Normal. Exemplo de amostra em que a suposição de normalidade não se pode assumir. Elaboração própria.

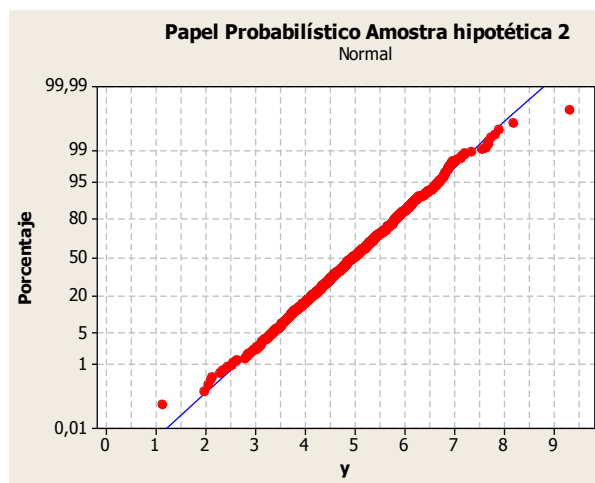


Figura 10. Papel probabilístico normal 2. Exemplo de amostra em que a suposição de normalidade se pode assumir. Elaboração própria.

Observe-se como no segundo caso os dados da amostra se ajustam bem na linha teórica da distribuição normal acumulada.

### Teste de Anderson-Darling

O teste de Anderson-Darling (AD) forma parte da família de testes não paramétricos de aderência, o que significa que não são baseados na suposição de que os dados procedem de uma determinada distribuição de probabilidade.



Este teste baseia-se na estimativa da distância mínima para ajustar um modelo matemático aos dados empíricos de uma amostra, de modo que respeito à distribuição de referência teórica deseja-se testar,

$$\left\{ \begin{array}{l} H_0: \text{Os dados proveem de uma distribuição específica} \\ H_1: \text{Os dados não proveem de uma distribuição específica} \end{array} \right.$$

O procedimento, no caso da distribuição normal, começa ajustando os dados empíricos para a função de distribuição da normal, padronizando os valores da amostra  $X_1, X_2, \dots, X_n$  tal que,

$$Y_i = \frac{X_i - \bar{X}}{\sigma}$$

Posteriormente, são classificados em ordem crescente os valores que formam a variável padronizada gerando a distribuição acumulada empírica como,

$$F(Y_{(j)}) \rightarrow \{Y_{(1)} < Y_{(2)} < \dots < Y_{(k)}\}$$

A fórmula da estatística do teste AD que é utilizada para verificar se os dados da amostra vêm de uma distribuição normal é,

$$A^2 = -n - S$$

com

$$S = \sum_{k=1}^n \frac{2k-1}{n} \{ \ln[F(Y_k)] + \ln[1 - F(Y_{n+1-k})] \}$$

A estatística de teste pode então ser comparada com os valores críticos da distribuição teórica (nesse caso, a normal padrão).

Este teste (utilizado pelo soft *Minitab*) difere do conhecido teste de Kolmogorov-Smirnoff em que é mais sensível nas caudas da distribuição.

### 3.2 Comparações de Médias

Para avaliar se há diferenças significativas entre dois ou mais grupos (níveis) de um fator a ser analisado (por exemplo, diferentes intervalos de IMC aos que um paciente pode pertencer) pode-se fazer uso dos contrastes de médias.

O objetivo é determinar se há evidências de que existem diferenças estatisticamente significativas entre as características de uma população. Se assim for, seria considerado “a priori” um fator de influência na variável resposta que motiva o estudo.

### **Teste $t$ de duas amostras (diferença entre duas médias)**

Quando se trata de comparar duas amostras (ou dois níveis de um fator) o teste que se utiliza é o teste  $t$  de diferença entre médias.

Um estimador não viciado da variância conjunta da população é dado por:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

sendo  $S_1^2$  e  $S_2^2$  as variâncias das duas amostras com tamanhos  $n_1$  e  $n_2$ .

Defina-se  $\bar{Y}_1$  e  $\bar{Y}_2$  como as médias amostrais; então a estatística,

$$T = \frac{\bar{Y}_1 - \bar{Y}_2 - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

tem uma distribuição  $t$  com  $n_1 + n_2 - 2$  graus de liberdade, sob a hipótese de igualdade de médias.

Supondo o teste bilateral, deve ser testado,

$$\left\{ \begin{array}{l} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{array} \right.$$

pelo que sob  $H_0$ , sob a hipótese que as médias são iguais (e, portanto, sua diferença é nula) a estatística de teste,

$$T_0 = \frac{\bar{Y}_1 - \bar{Y}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

segue uma distribuição  $t$  com  $n_1 + n_2 - 2$  graus de liberdade, sob  $H_0$ .

Fixando um nível de significância  $\alpha$  (tipicamente,  $\alpha = 0,95$ ), a hipótese nula é rejeitada no caso em que,

$$|T_0| \geq t_{(1-\alpha/2), n_1+n_2-2}$$

Outra forma de expressar a evidência de que os dois grupos comportam-se diferentemente, é mostrar o resultado dos limites do intervalo de confiança da diferença das médias.

Define-se o intervalo de confiança (1- $\alpha$ )% para a diferença de médias como

$$IC(\mu_1 - \mu_2): \bar{Y}_1 - \bar{Y}_2 \pm t_{(1-\alpha/2), n_1+n_2-2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

De tal modo que, se o “0” está dentro do intervalo, pode-se concluir que não existem diferenças significativas entre as médias das duas populações em estudo.

### **Teste de Tukey de intervalos de confiança simultâneos (comparações múltiplas de médias)**

Uma adaptação do teste  $t$  de comparação de duas médias para o caso em que se deseja estudar a diferença entre as médias de mais de dois níveis de um fator ou amostras de uma população, é o teste de Tukey.

Deve-se notar que, quando se realiza um teste de hipótese no teste  $t$  de diferença de duas médias, se estabelecem afirmações na base de certa significância estatística  $\alpha$ , (probabilidade de rejeitar a hipótese nula quando ela é falsa) para uma comparação de duas médias.

No entanto, quando se compara mais de duas características de cada vez de uma população, aumenta a probabilidade global de cometer um erro tipo 1 (probabilidade de rejeitar a hipótese nula sendo esta certa) por ter realizado o teste de comparação simples várias vezes.

Uma maneira de resolver o problema consiste em considerar um valor de  $\alpha$  para a *família* completa de testes. Isto quer dizer, que a significância global depois de realizar todas as comparações entre as médias seja a inicialmente prefixada (no presente trabalho,  $\alpha = 0,95$ ).

A estatística de teste de Tukey se define como (no caso de amostras com tamanhos diferentes),

$$q = \frac{\bar{y}_{max} - \bar{y}_{min}}{\sqrt{S_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

Nesse teste, duas médias são determinadas estatisticamente diferentes quando o valor absoluto da diferença excede,

$$T_\alpha = q_\alpha(p, f) \sqrt{S_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}$$

com  $q_\alpha$  o valor da distribuição  $q$  para uma significância global  $\alpha$ ,  $p$  o número de níveis do fator, e  $f$  o número de graus de liberdade associados com a variância amostral conjunta  $S_p^2$ .

O intervalo de confiança da diferença de cada par de médias será, portanto,

$$IC_{(1-\alpha)}[\mu_1 - \mu_2]: \bar{y}_1 - \bar{y}_2 \pm q(p, f) \sqrt{S_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}$$

Tal como no caso do teste  $t$  para a comparação simples de duas médias, no caso em que o "0" esteja dentro dos limites de algum dos intervalos, levaria a afirmação de que esse par de médias é equivalente.

### 3.3 Regressão Linear

A análise de regressão é o estudo da dependência de uma variável de resposta (ou dependente) de uma ou várias variáveis explicativas.

O objetivo é confirmar uma hipótese de pesquisa sobre a influência de um ou mais fatores sob uma variável que queremos estudar, podendo fornecer valores futuros da variável dependente.

#### 3.3.1 Regressão Linear Simples

Quando se trata de medir a relação entre uma variável dependente e uma variável explicativa se fala de regressão linear simples.

Pode-se expressar o valor esperado da equação de regressão como se segue,

$$E(Y_i|X_i) = \beta_0 + \beta_1 x_i$$

em que  $Y$  é a variável dependente,  $X_i$  é a variável explicativa e  $\beta_0$  (intercepto) e  $\beta_1$  (inclinação) são os parâmetros da regressão.

Além disso, deve notar-se que quando se utiliza uma amostra de observações, as variáveis envolvidas na análise de regressão são de natureza estocástica, de modo que, no momento de especificar um modelo estendido à população se deve incluir um termo de erro  $\epsilon_i$  tal que,

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

com  $\epsilon_i \sim N(0, \sigma^2)$  y  $COV(\epsilon_i, \epsilon_j) = 0$ .

Ou seja, os erros associados ao modelo são normalmente distribuídos, com média  $\mu = 0$  e variância  $\sigma^2$ , e não estão correlacionados.

Quanto à interpretação dos coeficientes de regressão simples,  $\beta_1$  indica a alteração média na variável de resposta quando a variável explicativa aumenta em uma unidade, e  $\beta_0$  indica o valor médio da variável de resposta, quando a variável explicativa assume o valor 0.

A tradução visual da equação acima é uma linha reta em torno da qual são distribuídos, mais ou menos correlacionadamente (dependendo do verdadeiro grau de ajuste do modelo para as observações da amostra), os dados empíricos que contém o modelo de regressão.

A figura 11 ilustra um exemplo de modelo de regressão linear simples genérica, onde podem ser vistos os parâmetros de regressão descritos.

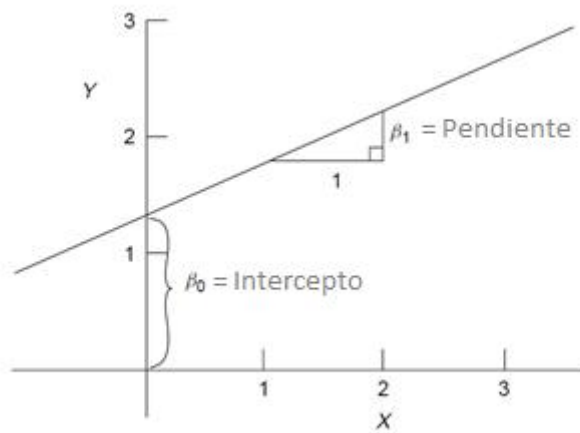


Figura 11: Representação gráfica dos coeficientes de Regressão Linear Simples. Extraído de [19].

### Pressupostos do Modelo de Regressão Linear

Para o resultado da análise de regressão linear ser considerado válido, ele deve atender a uma série de pressupostos que são descritos a seguir:

- **Normalidade.** A suposição de normalidade das observações é muito importante para testes de hipóteses e previsão de valores futuros. Uma vez que a equação de regressão é ajustada, verifica-se esse pressuposto através dos resíduos.
- **Homocedasticidade.** A homocedasticidade (variância constante) das observações é também uma exigência do modelo de regressão linear. A ausência de variância homogênea pode produzir um efeito de inflação da variância dos coeficientes de regressão estimados, de modo que os intervalos de confiança serão largos e os testes de hipóteses rejeitados, levando a conclusões potencialmente erradas.
- **Especificação correta do modelo.** Isto é, a omissão de variáveis relevantes no modelo de regressão, ou a inclusão de variáveis desnecessárias. Deve-se encontrar uma combinação de variáveis preditivas "ideais" para explicar da melhor maneira possível a variável resposta do estudo.

- **Ausência de multicolinearidade.** A Multicolinearidade pode ocorrer apenas em modelos de regressão linear múltipla (explicada em mais detalhe na secção designada para a explicação da multicolinearidade). Este fenómeno acontece nos casos em que existe uma relação linear, exata ou aproximada, entre os variáveis explicativas do modelo. As consequências são uma inflação da variância dos coeficientes estimados da regressão, o que pode conduzir a erros na hipótese de testes de significância estatística dos estimadores de mínimos quadrados.

### Método dos Mínimos Quadrados

Um procedimento amplamente utilizado para estimar os parâmetros  $\beta_i$ 's do modelo de regressão é chamado de método dos mínimos quadrados.

Este método envolve a especificação dos parâmetros da regressão, de modo a minimizar a soma dos quadrados dos resíduos.

Pode-se expressar o termo de erro reescrevendo a equação de regressão como,

$$\epsilon_i = Y_i - \beta_0 + \beta_1 x_i$$

em que  $\beta_0$  e  $\beta_1$  são os parâmetros estimados.

A soma dos quadrados dos resíduos é definida como,

$$S = \sum_{i=1}^n \epsilon_i^2$$

O critério de minimização é,

$$MIN(S) = \sum_{i=1}^n (Y_i - \beta_0 + \beta_1 x_i)^2$$

Para isso, se deriva a expressão acima em função dos parâmetros  $\beta_0$  e  $\beta_1$  e se igualam a 0 as equações resultantes, designadas como **equações normais**,

$$\frac{\partial(S)}{\partial \hat{\beta}_0} = -2 \sum_{i=1}^n (Y_i - \beta_0 + \beta_1 x_i) = 0$$

$$\frac{\partial(S)}{\partial \hat{\beta}_1} = -2 \sum_{i=1}^n (Y_i - \beta_0 + \beta_1 x_i) x_i = 0$$

Resolvendo o sistema de equações referido, se determinam os coeficientes de regressão associados aos estimadores de mínimos quadrados,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x}$$

Em resumo, se devem minimizar os erros do modelo teórico em relação às observações empíricas, definidos os erros como à distância das observações com a linha de regressão estimada.

Sob a suposição de normalidade dos erros, os Estimadores de Mínimos Quadrados obtidos coincidem com os respectivos Estimadores de Máxima Verossimilhança.

### **Bondade de ajuste da equação de regressão**

#### **O coeficiente de determinação R<sup>2</sup>**

O coeficiente de determinação de regressão R<sup>2</sup> é uma medida que corresponde à proporção de variação explicada pelo modelo em relação à variação total, e está entre 0 e 1.

Ele é equivalente ao quadrado do coeficiente de correlação de Pearson.

Para defini-lo matematicamente pode-se fazer uso da partição do conjunto da variabilidade associada com a variável de resposta como,

$$SCT = SCReg + SCRes$$

em que:



- SCT: Soma de quadrados total. É a soma dos quadrados de cada observação em relação à média global. É expressa como:

$$SCT = \sum_{i=1}^n (Y_i - \bar{Y})^2$$

- SCREG: Soma de quadrados da regressão, que é a soma dos quadrados das estimativas da regressão em relação à média global. É definida como:

$$SCReg = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$$

- SCE: soma de quadrados dos resíduos. É a soma das distâncias das observações às correspondentes estimativas da equação de regressão:

$$SCRes = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Finalmente, o coeficiente de determinação  $R^2$  é expresso como:

$$R^2 = \frac{SCReg}{SCT} = 1 - \frac{SCE}{SCT}$$

$$0 \leq R^2 \leq 1$$

A sua interpretação consiste em uma medida da quantidade de variação explicada pelo modelo de regressão em relação ao total da variabilidade associada com a variável de resposta. Em geral, se expressa em percentagem.

### **Intervalos de confiança para os parâmetros da regressão**

Outra medida de qualidade de ajuste do modelo consiste em especificar a precisão das estimativas dos coeficientes de regressão. Deve-se calcular previamente o desvio padrão deles. Os estimadores não viciados da variância do erro  $\sigma^2$  e o desvio padrão do erro  $\sigma$  são definidos como,

$$\hat{\sigma}^2 = \frac{SCRes}{N - 2} = \frac{\sum_{i=1}^n \epsilon_i^2}{N - 2}$$

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \sqrt{\frac{\sum_{i=1}^n \epsilon_i^2}{N - 2}}$$

O que equivale ao erro quadrático médio utilizado na análise de variância.

Pode-se definir o desvio padrão das estimativas da regressão como,

$$SE[\hat{\beta}_0] = \sqrt{VAR[\hat{\beta}_0]} = \sqrt{\sigma^2 \frac{\sum_{i=1}^n x_i^2}{N \sum_{i=1}^n (x_i - \bar{x})^2}} = \sigma \sqrt{\frac{\sum_{i=1}^n x_i^2}{N \sum_{i=1}^n (x_i - \bar{x})^2}}$$

$$SE[\hat{\beta}_1] = \sqrt{VAR[\hat{\beta}_1]} = \sqrt{\frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} = \frac{\sigma}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

No caso de  $\hat{\beta}_1$ , se tem a quantidade pivotal,

$$PIV = \frac{\hat{\beta}_1 - \beta_1}{SE[\hat{\beta}_1]}$$

que segue uma distribuição  $t$  de student conhecida (não dependente do parâmetro desconhecido  $\beta_1$  que se quer estimar), a qual coisa permite estabelecer, operando sob a expressão acima, um intervalo de confiança  $(1-\alpha)\%$  de tal modo que,

$$IC_{(1-\alpha)} [\hat{\beta}_1]: \hat{\beta}_1 \pm t_{\alpha/2} SE[\hat{\beta}_1]$$

Do mesmo modo, para o caso do parâmetro de intercepto se obtém,

$$IC_{(1-\alpha)} [\hat{\beta}_0]: \hat{\beta}_0 \pm t_{\alpha/2} SE[\hat{\beta}_0]$$

### 3.3.2 Análise dos resíduos

A fim de verificar os pressupostos básicos do modelo, se pode fazer uso do gráfico de resíduos, de modo que é possível estudar a violação de qualquer dos principais pressupostos (normalidade, homocedasticidade).

- **Gráfico de probabilidade normal dos resíduos.**

Como explicado na secção que descreve o gráfico de probabilidades, se podem representar os resíduos padronizados (ponderados pelo seu desvio padrão) em um gráfico de probabilidades e estudar visualmente a correlação dos resíduos em relação à linha teórica que representa o comportamento da distribuição normal. Em caso de existir desvios sensíveis respeito à normalidade hipotética, o pressuposto de normalidade não pôde ser verificado, de maneira que poderia estar comprometida a validade dos resultados e as análises realizadas.

- **Gráfico dos resíduos em relação ao conjunto de valores.**

Os resíduos são representados ao respeito dos valores ajustados da equação de regressão, para poder observar se o pressuposto de variância homogênea e / ou a correta especificação de modelo é aceitável. A figura a seguir, extraída de [19], mostra os diferentes tipos de gráficos de resíduos desse tipo.

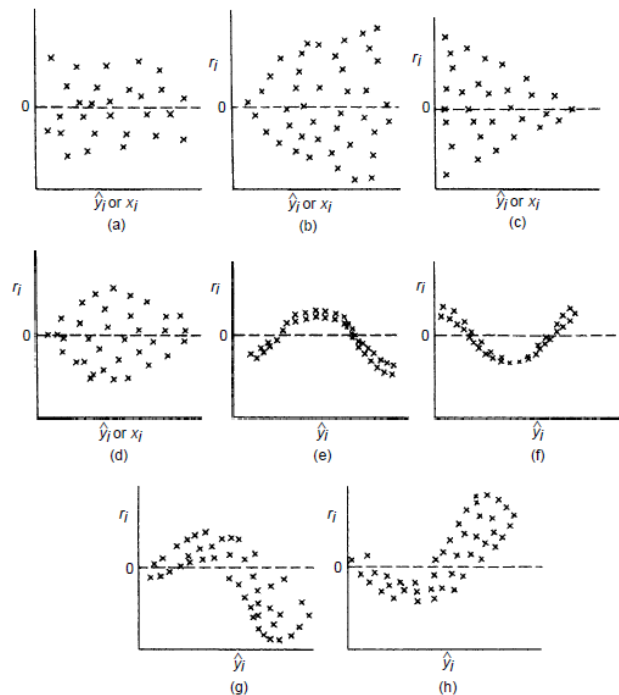


Figura 12. Tipos genéricos de gráficos de resíduos vs. valores ajustados. Extraído de [19].

No primeiro caso, a), se pode observar pela dispersão dos resíduos que a variância de erro pode ser considerada homogênea. As gráficas b), c) e d) indicariam uma variação heterogênea em resíduos (tipicamente, em forma de funil). Os casos e) e f), indicam um modelo de regressão incorretamente especificada, e os casos g) e h) mostram uma violação de ambos pressupostos.

### **Análise da variância (ANOVA)**

A partição indicada da variabilidade total associada com a variável dependente na análise de regressão permite introduzir o conceito de análise de variância (ANOVA, do inglês *Analysis of Variance*).

Os vários elementos descritos na partição das fontes de variabilidade são resumidos na tabela 4.

Tabela 4. Estrutura da tabela ANOVA.

| Fonte de variabilidade | SQ<br>(Somadas de Quadrados) | D.F. | QM<br>(Quadrados Médios) | F           | p-valor |
|------------------------|------------------------------|------|--------------------------|-------------|---------|
| Regressão              | SSReg                        | 1    | SSReg/1                  | MSReg/MSRes |         |
| Resíduos               | SSRes                        | n-2  | SSRes/(n-2)              |             |         |
| Total                  | SST                          | n-1  |                          |             |         |

A primeira coluna indica as fontes de variabilidade da variável em estudo. A segunda coluna refere-se à soma dos quadrados associados com cada fonte de variabilidade. A terceira coluna representa os graus de liberdade associados a cada partição da variabilidade. A quarta coluna contém os quadrados médios constituídos pelo quociente da soma dos quadrados entre os graus de liberdade de cada fonte de variação. A quinta coluna contém o quociente do quadrado médio da regressão e o quadrado médio dos erros, o que define a estatística  $F$ . A última coluna refere-se ao p-valor da regressão.

Quando  $\beta_1 = 0$ , o estatístico,

$$F = \frac{QMReg}{QMRes} = \frac{SSReg/d.f.}{SSRes/d.f.}$$

comporta-se como uma distribuição  $F$  com 1 e  $n - 2$  graus de liberdade.

Então, na realidade, está a se testar,

$$\left\{ \begin{array}{l} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{array} \right.$$

Pelo que se avalia a relevância da variabilidade associada com a regressão. No caso em que esta seja elevada, a estatística  $F$  aumenta, atingindo a região de rejeição do teste.

A fim de reportar a significância estatística, se expressa o p-valor da regressão, entendido como o valor máximo de  $\alpha$  a partir do qual, valores maiores de distribuição  $F$ , sempre conduzem à rejeição da hipótese nula.

### 3.3.3 Previsão de novas observações

Considerando o fato de que há certo grau de incerteza quanto às estimativas dos parâmetros de regressão (manifestada na existência de uma porção da variabilidade da variável dependente não explicada, adicionada ao erro), significa que ao fazer projeções futuras deve-se quantificar o grau de incerteza associado com a estimativa da variável dependente  $\hat{Y}$ ; ou seja, ao *resultado* da aplicação da equação de regressão para novas observações, a fim de prever o comportamento da variável dependente, se deve adicionar uma medida que indique o grau de incerteza que se tem do valor obtido. Se pode, por conseguinte, construir um intervalo de  $(1-\alpha)\%$  da estimativa da variável dependente.

Existem dois tipos de predição na regressão linear:

- Previsão para a média de Y
- Previsão para um valor pontual de Y

No caso da previsão média, o objetivo é estimar  $E[Y|X = x_0]$ . Isto é, refere-se que no caso em que foram realizadas muitas previsões considerando um valor concreto da variável explicativa (e, portanto, pode ser assumido que a soma dos erros será 0), *em média*, o valor da variável dependente seria o especificado.

Por contraste, na predição de um valor pontual, não pode presumir-se que a variável dependente terá o valor médio, de modo que se deve incluir um termo de erro à estimativa  $\hat{Y}$ .

No caso da **previsão média**, o estimador não viciado da esperança da variável resposta é dado por,

$$\hat{Y} = E[Y_i|X = x_0] = \beta_0 + \beta_1 x_0$$

Assim que a distribuição do estimador pode ser definida como,

$$\hat{Y}_{media} \sim Normal \left( \beta_1 + \beta_2 X_i, \sigma^2 \left[ \frac{1}{n} + \frac{(X_0 - \bar{x})^2}{\sum_{i=1}^n x_i^2} \right] \right)$$

E se utiliza a quantidade pivotal,

$$PIV = \frac{\hat{Y} - (\beta_0 + \beta_1 x_0)}{SE[\hat{Y}]}$$

que segue uma distribuição *t* de Student com  $n - 2$  graus de liberdade. O cálculo do desvio padrão da estimativa da variável de resposta se efetua através do cálculo da raiz quadrada do estimador da variância, explicado acima.

Assim, podem-se estabelecer os limites do intervalo de confiança do  $(1-\alpha)\%$  para a previsão média como,

$$IC_{(1-\alpha)}[\hat{Y}_{media}]: \beta_0 + \beta_1 x_0 \pm t_{\alpha/2} SE[\hat{Y}_{media}]$$

Na **previsão pontual**, a distribuição do estimador de resposta é dada por,

$$\hat{Y}_{puntual} \sim Normal\left(\beta_0 + \beta_1 x_0, \sigma^2 \left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n x_i^2}\right]\right)$$

Observe-se, como explicado, que se acrescentou uma unidade para o termo que multiplica ao estimador da variância do erro.

O intervalo de confiança para a estimativa da previsão pontual, portanto, é

$$IC_{(1-\alpha)}[\hat{Y}_{puntual}]: \beta_0 + \beta_1 x_0 \pm t_{\alpha/2} SE[\hat{Y}_{puntual}]$$

É claro, portanto, que como a variância associada com a estimativa de previsão pontual é mais elevada do que a previsão média, deve notar-se que,

$$IC_{(1-\alpha)}[\hat{Y}_{puntual}] > IC_{(1-\alpha)}[\hat{Y}_{media}]$$

### 3.3.4 Regressão Linear Múltipla

Na regressão linear múltipla é generalizado o modelo anterior, de modo a incluir varias variáveis explicativas na equação.

Assim, o modelo de regressão linear múltipla, pode ser expresso como,

$$Y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \epsilon_i$$

com  $k$  variáveis explicativas .

$\beta_i$ 's são os coeficientes de regressão,  $x_i$  as variáveis explicativas e  $\epsilon$  é o erro aleatório.

Pode entender-se o modelo anterior como um conjunto de equações tais que,

$$Y_1 = \beta_0 + \beta_1 x_{11} + \dots + \beta_k x_{k1} + \epsilon_1$$

$$Y_2 = \beta_0 + \beta_1 x_{12} + \dots + \beta_k x_{k2} + \epsilon_2$$

⋮

$$Y_n = \beta_0 + \beta_1 x_{1n} + \dots + \beta_k x_{kn} + \epsilon_n$$

Na notação matricial tem-se,

$$Y = X\beta + \epsilon$$

com Y um vector (nX1) de respostas, X uma matriz [n x (k+1)],  $\beta$  um vector [(k+1) x 1] dos coeficientes de regressão e  $\epsilon$  um vector (n x 1) contendo todos os erros aleatórios.

Assim, a equação matricial resultaria,

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \dots & x_{k1} \\ 1 & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{k1} \end{bmatrix} \times \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

com  $\epsilon_i \sim N(0, \sigma^2)$  e  $COV(\epsilon_i, \epsilon_j) = 0$ .

### **Mínimos Quadrados na Regressão Linear Múltipla**

O procedimento é análogo ao da Regressão Linear Simples, mas considerando que estamos agora com variáveis matriciais.

Isolando o vetor do erro e elevando-o ao quadrado, da álgebra linear tem-se,

$$\sum_{i=1}^n \epsilon^2 = \epsilon' \times \epsilon = (Y - X\beta)' \times (Y - X\beta)$$

em que o superíndice “ ‘ ” indica que é a matriz transposta (as filas de uma matriz são equivalentes às colunas da sua transposta).

Operando, derivando e igualando a 0, se obtém as estimativas dos coeficientes de regressão:

$$\hat{\beta} = (X'X)^{-1} \times (X'Y)$$

## Bondade de ajuste da regressão múltipla: o coeficiente de determinação $R^2$ -ajustado

Como medida de bondade de ajuste da equação de regressão linear múltipla se utiliza o coeficiente de determinação ajustado  $R^2$ -ajustado.

Esta adaptação do ajuste é usada porque quando aumenta o número de variáveis explicativas em uma eventual equação de regressão, o coeficiente de  $R^2$  aumenta artificialmente.

Neste caso, se leva em conta o número de variáveis incluídas no modelo de tal forma que,

$$R^2 - ajust = 1 - (1 - R^2) \frac{n - 1}{n - k}$$

em que  $n$  é o tamanho da amostra e  $k$  o número de variáveis incluídas no modelo de regressão.

Note-se que para grandes amostras o denominador  $(n - k)$  não varia muito pelo que o coeficiente ajustado se aproximará ao coeficiente simples.

É um método que, basicamente, é usado para comparar diferentes equações com diferente número de variáveis explicativas.

Como no caso do coeficiente de determinação não ajustado, se expressa geralmente em porcentagem.

## A multicolinearidade

Um dos efeitos indesejáveis na regressão linear múltipla é a **multicolinearidade**. Ela ocorre quando duas ou mais variáveis explicativas são correlacionadas umas com as outras. Existem dois tipos de colinearidade entre variáveis:

- Colinearidade exata.
- Colinearidade aproximada

Na **colinearidade exata** a correlação entre duas variáveis é perfeita.



Para realizar a estimativa dos coeficientes de regressão  $\beta_i$ , deve ser tido em conta o cálculo da inversa da matriz  $(X'X)$  de tal modo que,

$$(X'X)^{-1} = \frac{1}{\det(X'X)} \text{adj}(X'X)$$

No caso da colinearidade perfeita, a matriz  $(X'X)$  é *singular* o que significa que,

$$\det(X'X) = 0$$

e o estimador de mínimos quadrados é indeterminado, por não poder ser obtido

$$\hat{\beta} = (X'X)^{-1} \times (X'Y)$$

A colinearidade perfeita ocorre em poucos casos, e sua detecção é rápida.

No que respeita à **colinearidade aproximada**, ocorre em circunstâncias em que a relação linear entre duas ou mais variáveis é quase perfeita. Nestes casos, a matriz  $(X'X)$  não é singular, mas seu determinante é próximo ao 0.

A consequência é um aumento da variância associada com as estimativas dos coeficientes de regressão, cujo cálculo se pode definir como,

$$\text{VAR}(\hat{\beta}) = \hat{\sigma}^2 (X'X)^{-1} = \hat{\sigma}^2 \frac{\text{Adj}(X'X)}{\det(X'X)}$$

Isto implica que, nas estatísticas  $t$  de significância individual da regressão definidas como,

$$t = \frac{\hat{\beta}_i}{\sqrt{\text{VAR}(\hat{\beta}_i)}} \sim t_{n-k}$$

tendem a rejeitar a hipótese nula, chegando a conclusões potencialmente errôneas com que as variáveis colineares não são importantes para explicar a variável resposta.

Para detectar e resolver o problema da multicolinearidade se podem seguir os seguintes passos.

- a) Estudo preliminar da correlação das variáveis.

Antes de seleccionar as variáveis incluídas no modelo de regressão linear múltipla se pode determinar sua correlação (por meio do coeficiente de correlação de Pearson), a

fim de prever o seu potencial de influência nos efeitos de colinearidade sobre a variável de resposta.

b) Cálculo dos Fatores de Inflação da Variância (FIV).

Os FIV podem ser definidos como,

$$FIV(\beta_i) = \frac{1}{1 - R_i^2}$$

Sendo  $R_i^2$ , o coeficiente de determinação da variável  $X_i$  em relação às outras variáveis explicativas. O FIV evidencia a contribuição à colinearidade do modelo de regressão por parte da variável explicativa analisada.

Pode-se considerar que uma variável contribui decisivamente para o efeito indesejado de multicollinearidade quando  $FIV > 5$  [20].

c) Reespecificação do modelo de regressão

Em casos em que há alta correlação entre as variáveis pode-se omitir uma delas de forma de não perder grau de explicação e previsão do modelo. Da mesma forma se podem rescrever as variáveis colineares em uma nova variável (como a diferença das dois, por exemplo) e inclui-la no modelo de regressão linear.

### **Extensão do modelo. Dummies e interações**

A consideração de variáveis *Dummy* surge a partir da necessidade de incluir informação qualitativa na análise de regressão, simples ou múltipla.

Por exemplo, no estudo da variabilidade do tempo de cirurgia se pode estar interessado em considerar o género do cirurgião responsável como um fator que influencia a duração das intervenções. Nesse caso, se pode criar uma variável binária que recolha a informação das observações históricas de cirurgias passadas sobre o sexo do cirurgião.

O procedimento geral para a construção destas variáveis é atribuir dois valores possíveis [0,1], cada um dos quais representa a realização ou não do evento considerado (exemplo: 1 = o cirurgião é do sexo feminino, 0 = o cirurgião é do sexo masculino).

Tal como no caso das variáveis quantitativas habituais, a significância destas variáveis será determinada pela quantidade de variabilidade explicada por elas. Isto é, quando a variável binária leva um ou outro dos seus dois valores possíveis, a variável de resposta tem uma alteração que pode ser considerada em uma possível equação de regressão final, através da inclusão da variável Dummy como variável explicativa.

Outra possível extensão do modelo de regressão que pode eventualmente ser útil é a incorporação na equação de regressão das variáveis de interação.

### 3.3.5 Seleção de variáveis em regressão linear múltipla: *Stepwise Regression*

Ao contrário dos modelos de regressão linear simples, na análise de regressão linear múltipla encontra-se o problema de decidir qual das combinações de variáveis explicativas é a que melhor explica a variável resposta. Um método amplamente utilizado é chamado de **Stepwise Regression**. Tem a vantagem sobre os seus homólogos unidirecionais (*Forward Regression* e *Backward Regression*), que quando uma nova variável é incluída em qualquer passo do método, não permanece definitivamente na equação até alcançar o modelo final. Em vez disso, ao estabelecer critérios de entrada e saída das variáveis envolvidas na equação de regressão, é avaliado em cada passo do processo de construção da equação final se deve rejeitar-se alguma das variáveis incluídas anteriormente.

As etapas do algoritmo de seleção de variáveis na Stepwise Regression são mostradas a seguir:

- **Passo 1:** Estabelecer um critério de entrada  $t_{IN}$  e um critério de saída  $t_{OUT}$ .
- **Passo 2:** Calcular uma equação de regressão para cada uma das variáveis explicativas consideradas na análise.

Nesse ponto, é avaliado para cada uma das potenciais variáveis explicativas se  $t_k \geq t_{IN}$ . Dentre os casos em que a significância individual da variável seja maior que o critério de entrada estabelecido, é escolhida aquela variável com maior  $t_k$ .

- **Passo 3:** Selecionar uma variável adicional e incorporá-la ao modelo definido no passo 2. Repete-se o procedimento para todas as demais variáveis a incluir.

Volta-se a escolher a variável com maior significância estatística no global da regressão.

- **Passo 4:** Estudar se a(s) variável(is) incluídas com anterioridade ao último passo se encontram dentro da região de rejeição definida no critério  $t_{OUT}$ , de modo que se  $t_{k-1} < t_{out}$  as variáveis incluídas anteriormente são susceptíveis de ser eliminadas do modelo (as quais, por efeitos de confusão, podem ter modificado sua significância). De maneira que, dentre as que cumplam o critério de rejeição,  $\min(t_{k-1})$  é removida da equação.
  
- **Passo 5:** Repetir os passos 3 e 4 até que nenhuma variável não é mais suscetível a entrar ou sair da equação de regressão linear.

# CAPÍTULO 4: ESTUDO DE CASO

---

## 4.1 Introdução

### 4.1.1 O câncer de mama

No Brasil o câncer de mama é o tipo de câncer mais habitual entre as mulheres (o câncer de mama em homens, apesar de existir, significa unicamente 1% dos casos), e representa 22% dos casos novos de câncer cada ano [21].

Entre os fatores de risco mais importantes de desenvolver câncer de mama se encontram o peso, a alimentação, a falta de atividade física, o consumo de álcool e o tabagismo.

Esses riscos explicam em parte que precisamente a região mais desenvolvida do país, a região sudeste, é na que ocorre a maior incidência, onde a taxa bruta de casos é de 64,54/100mil mulheres. Este fato evidencia que é nas regiões industrializadas onde acontece maior número de casos de câncer novos, devido ao consumo mais inadequado de alimentos, o aumento de peso decorrente e o estresse.

### 4.1.2 O Instituto Nacional de Câncer (INCA). HCIII.

O Instituto Nacional de Câncer (INCA) é uma instituição ligada ao Ministério da Saúde, cujo escopo principal é a assistência oncológica, promoção de pesquisa orientada ao estudo do câncer e o desenvolvimento de ações estratégicas para o enfrentamento do câncer no Brasil [22].

O organograma geral da organização se pode observar na seguinte figura.

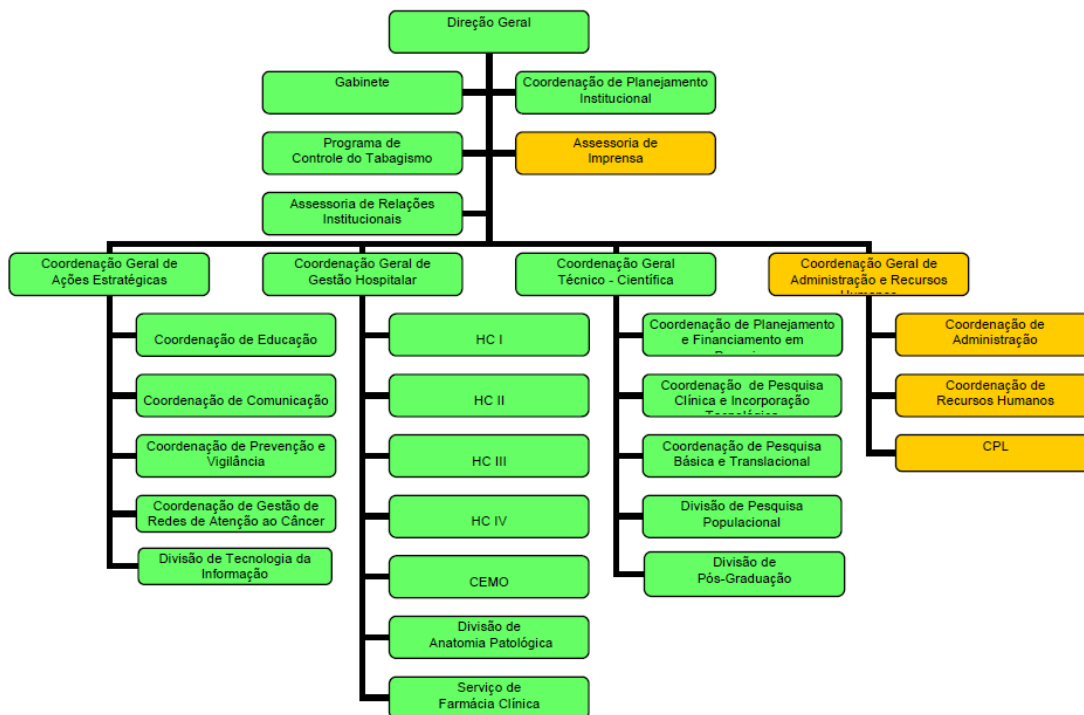


Figura 13. Organograma geral do Instituto Nacional de Câncer (INCA). Extraído de [23].

No que respeita à sua capacidade e organização assistencial, o INCA possui 4 hospitais divididos por área cirúrgica, disponibilizando em total 391 leitos, e distribuídos entre os diferentes hospitais como:

- Hospital de Câncer I, atendimento geral à maioria de subespecialidades: 188 leitos
- Hospital de Câncer II, ginecologia oncológica e tecido ósseo conectivo: 83 leitos.
- Hospital de Câncer III, câncer de mama: 52 leitos.
- Hospital de câncer IV, cuidados paliativos a pacientes com cânceres avançados sem possibilidade de tratamento curativo.

No que respeita ao câncer de mama, à missão institucional específica do HCIII consiste em **participar de políticas públicas de atenção ao câncer de mama através de ações de assistência, ensino e pesquisa** [24].

Tradicionalmente, as intervenções realizadas dentro no centro cirúrgico do HCIII unicamente eram relacionadas com a área de mastologia (remoção de tumor mamário). Porém, no ano 2011 foi transferida a unidade de cirurgia plástica (reconstrução da mama após cirurgia mastológica) devido a que foi detectada capacidade sobranete (leitos) no hospital [25].

### **4.1.3 Procedimentos envolvidos na área de mastologia do HCIII. Descrição geral**

Existem dois tipos básicos de procedimentos relacionados com a área da cirurgia mastológica: A Mastectomia e a Segmentectomia [25].

A Mastectomia é uma operação na qual toda a mama é removida. Por outro lado, a Segmentectomia inclui apenas a remoção do tecido canceroso no peito. A decisão de submeter-se a um ou outro procedimento depende de cada paciente, e obviamente, enriquecida com as recomendações do seu médico.

#### **4.1.3.1 Mastectomia**

Em geral, as mulheres com câncer de mama optam se submeter a uma operação de Mastectomia nos seguintes casos:

- Se o tumor é muito volumoso e após uma Segmentectomia há poucas chances de preservar o tecido mamário.
- Se não querem se submeter a sessões de radioterapia, necessárias após a operação de Segmentectomia.
- Pela ansiedade e o medo à recorrência do câncer, mais provável em casos de procedimentos de Segmentectomia.

Historicamente, a Mastectomia foi combinada com a linfodectomia axilar (esvaziado da maior parte dos gânglios linfáticos axilares). Na última década, no entanto, os avanços no procedimento cirúrgico permitiram a realização da chamada Biopsia Linfonodo Sentinela, na qual são extraídos alguns linfonodos da axila (tipicamente entre 2 e 3, embora dependa das características do paciente e a cirurgia) e pode ser analisado no decurso da intervenção, se é necessário proceder a um esvaziamento dos gânglios linfáticos axilares remanescentes para eliminar a presença de células cancerosas.

Essa evolução recente derivou em uma modificação da nomenclatura associada à Mastectomia, devendo o cirurgião que escolher entre os seguintes tipos de Mastectomia:

- **Mastectomia Simples:** O cirurgião remove todo o tecido da mama, mas não elimina o tecido muscular debaixo do peito. Isto é combinado com a Biópsia Linfonodo Sentinela, cujo resultado negativo não conduz ao esvaziamento axilar.
- **Mastectomia Radical Modificada:** Esta cirurgia envolve uma Mastectomia Simples, e pelo resultado positivo da BLS, se produz um esvaziamento axilar.

Em geral, a cirurgia consiste das seguintes atividades (para a Mastectomia Simples):

- Esterilização da sala de cirurgia.
- Administração de medicamentos por via intravenosa ao paciente durante a cirurgia.
- Controle da frequência cardíaca pela máquina de eletrocardiograma.
- Colocação de gaze estéril sobre o corpo da mulher para prevenir a infecção.
- Administração de anestesia geral.
- Dependendo do procedimento planeamento, o cirurgião faz a incisão na mama.
- Remoção do tecido mamário interno à incisão, e envio para a sua análise em um laboratório de patologia.
- Encerramento das feridas com suturas ou grampos.
- Acorda-se ao paciente e a sala de cirurgia é liberada, enviando-o para o pós-operatório.

#### 4.1.3.2 Segmentectomia

A Segmentectomia é um procedimento cirúrgico que envolve a remoção de uma lesão maligna (cancerosa), e uma pequena porção do tecido da mama circundante de uma mulher.

O tecido é testado para verificar que não existem células cancerosas adicionais.

Como no caso da Mastectomia, se pode realizar uma Biópsia de Linfonodo Sentinela, a fim de decidir se a proceder ao esvaziamento axilar.

Ao contrário da Mastectomia, as mulheres submetidas a uma Segmentectomia recebem tratamentos de radiação por cerca de seis semanas após a conclusão do



procedimento (que às vezes é um impedimento para decidir se optar por uma Segmentectomia ou uma Mastectomia, em que a radioterapia pós-intervenção não é sempre necessária) para matar as células cancerígenas que podem ter permanecido no peito após a remoção do tumor.

A Segmentectomia, nascida após a Mastectomia foi imposta como o principal procedimento realizado (quando possível) nas últimas décadas, devido a que a cirurgia permite preservar a mama (e sua aparência natural), enquanto que o tumor maligno é removido. É por isso que este procedimento é chamado de *breast-conserving procedure*.

O volume e a massa de mama removidos dependem do volume e a localização do tumor na mama, e são de fato dois fatores importantes considerados por um médico para determinar se a Segmentectomia é um procedimento adequado ou, pelo contrario, deve-se proceder à remoção total da mama.

Em geral, uma paciente é candidata a se submeter a uma Segmentectomia se o câncer for diagnosticado precocemente.

No entanto, há uma série de casos especiais em que não se pode realizar. São descritos a seguir:

- Múltiplos tumores em diferentes localizações da mesma mama. Nestes casos, a deformação da mama, resultante da remoção do tecido canceroso, levaria à decisão de remover completamente o seno.
- Sessões de radiação após um procedimento de Segmentectomia anterior.
- Alto volume do tumor. Quando o tumor se espalhou para uma vasta região da mama é aconselhável remover completamente a mama.

Neste caso, os seios pequenos sofrem de um risco aumentado de veto quando a possibilidade de se proceder a uma Segmentectomia. No entanto, às vezes se pode optar por fazer sessões prévias de quimioterapia, o que pode reduzir o volume do câncer a dimensões que permitem a conservação da mama através de Segmentectomia.

- tumores agressivos, tais como aqueles que crescem com grande velocidade ou que danam órgãos importantes do paciente.

- Pacientes que estão grávidas, e que a fase de radioterapia pós-operatória pode afetar o feto.
- Condições especiais que tornam inviável a radioterapia.

O processo completo do procedimento é similar à Mastectomia, exceto que neste caso, como já mencionado, apenas uma parte é removida da mama. **O desafio para o cirurgião é remover todo o tumor e uma seção ao lado dele, tentando preservar o máximo de tecido mamário possível.**

#### 4.1.4 Fatores de variação estudados

Em uma cirurgia existem inúmeros fatores que podem afetar o tempo de duração da mesma. A fim de limitar as análises, foi estudada uma porção deles dependendo da sua origem:

- Fatores associados à paciente
- Fatores associados ao cirurgião
- Fatores associados com o procedimento

A cirurgia pode ser realizada de várias maneiras (procedimento), mesmo que os pacientes tenham características diversas (paciente), e pode ser praticada por diferentes médicos (cirurgião).

O gráfico a seguir mostra os vários fatores estudados e a sua origem.

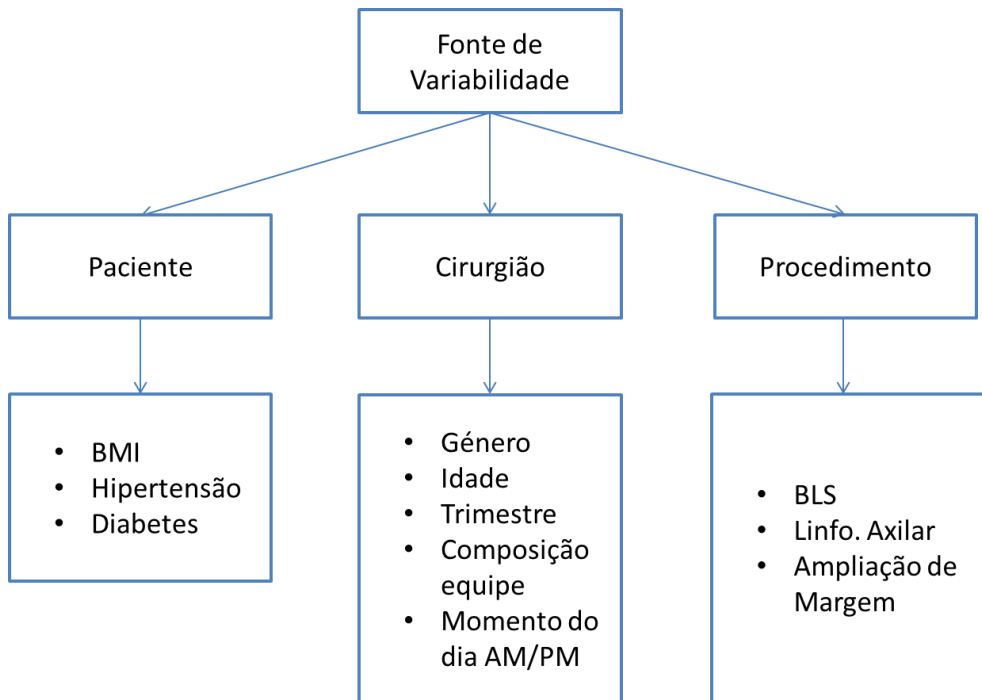


Figura 14. Fontes de variabilidade e fatores estudados. Elaboração própria.

### (i) Fatores associados ao paciente

Através de várias reuniões com os cirurgiões do hospital foi recomendado o estudo de três fatores que poderiam influenciar a duração da cirurgia:

- Índice de Massa Corporal (IMC)
- Hipertensão.
- Diabetes.

### Índice de Massa Corporal (IMC)

O índice de massa corporal é um dos indicadores do estado nutricional de um indivíduo. É calculado da seguinte forma:

$$IMC = \frac{Peso \left[ \frac{kg}{m^2} \right]}{altura^2 \left[ \frac{m^2}{m^2} \right]}$$

Foi popularizado nos anos 70 com o objetivo de representar quantitativamente o nível relativo de gordura no corpo de qualquer pessoa.

Embora existam estudos que mostram suas limitações como uma medida eficaz de gordura corporal [26], tais como a existência de uma pequena correlação entre a altura e o indicador, ou proporcionalidade corporal assimétrica entre sujeitos, é amplamente utilizado na área da saúde clínica.

Existe uma convenção médica sobre a qualificação que se realiza de um indivíduo, em termos de IMC, dependendo se ele pertence a qualquer um dos seguintes intervalos:

- IMC <20: Baixo Peso
- $20 \leq \text{IMC} < 25$ : peso normal ou eutrofia
- $25 \leq \text{IMC} < 30$ : Sobrepeso
- IMC  $\geq 30$ : Obesidade

Dentro do processo de cirurgia de um hospital, um dos testes que se realizam para determinar a condição física do paciente são a pesagem e medição da altura dele, que são as duas variáveis necessárias para o cálculo do IMC, como indicado na expressão acima.

Estes exames são realizados entre um ou dois dias antes da cirurgia, de modo que são considerados confiáveis como representação do verdadeiro IMC do paciente durante a cirurgia.

## **Hipertensão**

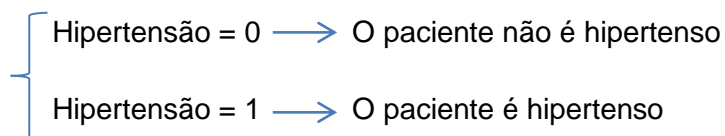
A pressão arterial é expressa por duas etapas, dependendo da fase do coração para bombear sangue.

A pressão arterial em relação à fase de ejeção do sangue para as artérias é conhecida como hipertensão arterial sistólica. Quando o coração relaxa e a pressão diminui, a medição refere-se à hipertensão arterial diastólica. É geralmente expressa em uma nomenclatura dupla, onde o primeiro número indica a pressão arterial sistólica e o segundo indica a pressão arterial diastólica. Por exemplo, uma medida de 110/80 mm Hg, consiste de uma tensão arterial sistólica de 110 mmHg e uma pressão diastólica de 80 mm Hg (mm Hg, milímetros de mercúrio, 1 mm Hg = 1 atm)

Existem três diferentes diagnósticos de hipertensão dependendo do grau da mesma [27]:

- Estado 1 de hipertensão: A pressão é medida na clínica é igual ou superior a 140/90 mmHg e inferior a 160/100 mmHg.
- Estado 2 de hipertensão: pressão medida clínica é igual ou superior a 160/100 mmHg e a pressão sistólica inferior a 180 mmHg.
- Hipertensão grave: pressão sistólica medida na clínica é igual ou superior a 180 mmHg.

Para os fins deste projeto, e como os mesmos cirurgiões do hospital indicam no relatório de cirurgia, só será considerado **se o paciente é hipertenso ou não**. De modo que, nas sucessivas análises da variância, a variável *hipertensão* será de caráter binário tal que,



## **Diabetes**

A diabetes é uma doença em que o organismo não produz insulina, ou não a utiliza adequadamente. A insulina é um hormônio necessário para converter açúcar, amidos e outros alimentos em energia necessária para realizar atividades diárias.

É uma doença que afeta cerca de 22 milhões de pessoas no Brasil, representando 12% de sua população [28]. Caracteriza-se por um aumento anormal da glicose ou açúcar no sangue.

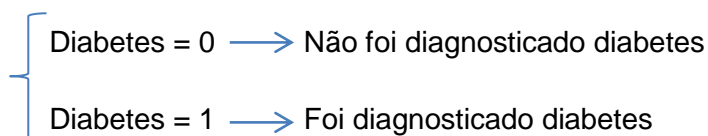
Existem vários tipos de diabetes [29].

- *Diabetes tipo 1 ou diabetes juvenil*: Neste tipo de diabetes, a quantidade de insulina produzida pelo corpo não é suficiente, de modo que os níveis de glicose no sangue aumentam. No caso da diabetes do tipo 1, o pâncreas deixa de produzir insulina.

- *Diabetes Tipo 2*: Esta é a forma mais comum de diabetes. Caracteriza-se por uma resistência corporal à insulina, resultando no aumento da produção dela pelo pâncreas, e ao longo dos anos em um esgotamento final das células que produzem a hormona. Por esta razão, é mais comum (embora nas últimas décadas é também manifestada nos jovens) que esta doença surja na terceira idade.

- *Diabetes gestacional*: É diagnosticada a mulheres que não tiveram anteriormente diabetes, as quais durante a gravidez sofrem um aumento anormal do açúcar no sangue.

Tal como no caso da hipertensão, na análise deste trabalho será considerada, apesar da classificação anterior meramente ilustrativa, dois cenários possíveis relativos à informação qualitativa que possui a variável binária *Diabetes*:



## **(ii) Fatores associados ao cirurgião**

Sendo uma atividade em grande parte manual e executada por pessoas, é inevitável não considerar os cirurgiões como uma possível fonte de variação no tempo da cirurgia. De fato, alguns autores sugerem que é a principal fonte de variação [30].

Para seu estudo deve ser classificado em termos do tipo de cirurgião que realiza o procedimento. Uma proposta de classificação é indicada em [31] e mostra uma categorização da equipe cirúrgica nos seguintes fatores:

- *Género do cirurgião*. Determina se existem diferenças na duração da cirurgia, dependendo do cirurgião ser mulher ou homem.
- *Idade do cirurgião*, sob a suposição inicial de que os cirurgiões mais velhos são mais rápidos para praticar cirurgias.
- *Cirurgias realizadas anteriormente*: Colete a influência do número de cirurgias realizadas antes da intervenção, assumindo que a curva de experiência de um eventual cirurgião indica que a maior número de intervenções praticadas, mais curtas as durações dos procedimentos sucessivos realizados.

- *Equipe cirúrgica*: Este fator inclui o fato de se participou ou não um ou mais residentes em cirurgia, considerando um possível efeito sobre a duração das cirurgias devido a aspectos de treinamento dos médicos em formação.

- *Momento do início da cirurgia*: Este fator poderia indicar duas possíveis causas da variabilidade. Em primeiro lugar, existem pessoas dependendo da hora do dia têm mais ou menos excitabilidade cerebral, de modo que são mais ou menos rápidas para realizar as mesmas atividades. Por outro lado, o fato de ter praticadas outras cirurgias no mesmo dia da intervenção, e acumulando várias horas de trabalho durante a jornada, pode conduzir ao esgotamento físico e mental, o que leva a um aumento do tempo necessário para executar as mesmas tarefas.

Há várias diferenças na abordagem de classificação acima que devem ser considerados no presente projeto.

É possível avaliar a influência do sexo, da idade do cirurgião, e o tempo de dia da cirurgia é efetuada.

Pelo contrário, o fato de que na equipe cirúrgica participem ou não residentes não fornece nenhuma informação nova para a análise, uma vez que as cirurgias realizadas no HCIII **sempre envolvem um residente**. A residência na área de mastologia consiste de dois anos de duração. Um primeiro ano em se assiste às cirurgias, mas não se participa ativamente nelas (a definição do cirurgião é R1), e uma segunda fase de um ano de duração em que o cirurgião residente implementa, sob a supervisão de cirurgiões experimentados *Staff* do hospital, o aprendido durante seu primeiro ano de residência (nesta fase o estado cirurgião é R2).

Da análise dos prontuários é apreciado um fenômeno que pode ser tomado em consideração na equipa cirúrgica. Há uma percentagem de casos em que no relatório de cirurgia se inclui ao cirurgião responsável *staff* na equipe cirúrgica. Quando isso ocorre, a ordem é muitas vezes o mesmo, e na seção de “equipe cirúrgica” primeiro é colocado o cirurgião R2 residente, segundo o cirurgião *staff*, e em terceiro lugar o R1 residente. Esta evidência, e assumindo que haja quórum entre os cirurgiões para completar as diferentes seções do relatório, pode indicar que existem formas diversas de participação ativa do cirurgião *staff* na cirurgia: uma como cirurgião responsável e outra como auxiliar.

Assim, o fator *equipe cirúrgica* indicando a participação de residentes na cirurgia foi adaptado de modo que reúne informações sobre a participação do cirurgião *staff* na equipe cirúrgica como assistente ou como cirurgião responsável.

No que diz respeito ao fator relativo ao número de operações realizadas antes da intervenção, é difícil obter dados confiáveis. No entanto, o fato de que a estrutura das equipes sempre contém um cirurgião R2 residente envolvido na cirurgia (seja na forma que é), e tendo em conta que todos os residentes começam a sua fase de treinamento no início de ano, se optou por avaliar a influência da data em que a cirurgia é realizada. O objetivo é comprovar a suposição de que um cirurgião inexperiente vai operar de forma mais rápida no quarto trimestre do que no primeiro trimestre do ano, devido ao número de operações realizadas durante todo o ano, o que pode constituir uma contribuição de experiência que poderia conduzir a uma redução do tempo necessário para executar um procedimento.

De acordo com o exposto, as variáveis a serem consideradas no estudo da contribuição da variabilidade do tempo dos cirurgiões são:

a) Género do cirurgião responsável.

Variável binária com espaço de soluções,

$$\left[ \begin{array}{l} 0, \text{ se o cirurgião é um homem} \\ 1, \text{ se o cirurgião é uma mulher} \end{array} \right.$$

b) Idade do cirurgião responsável

No estudo de regressão será considerada uma variável quantitativa que consistirá do ano de nascimento do cirurgião responsável.

Na análise exploratória dos dados, porém, consistirá de uma variável binária tal que,

$$\left[ \begin{array}{l} 0, \text{ se o cirurgião nasceu antes de 1960} \\ 1, \text{ se o cirurgião nasceu depois de 1960} \end{array} \right.$$

c) Trimestre de realização da cirurgia.

Para a análise individual do fator foi considerada uma variável com quatro níveis tais que,



- 1, se a cirurgia foi realizada no trimestre 1 (1 janeiro - 31 março)
- 2, se a cirurgia foi realizada no trimestre 2 (1 abril - 30 junho)
- 3, se a cirurgia foi realizada no trimestre 3 (1 julho - 30 setembro)
- 4, se a cirurgia foi realizada no trimestre 4 (1 outubro - 31 março)

d) O cirurgião responsável incluído ou não na equipe cirúrgica.

Variável binária que consiste em,

- 0, se o cirurgião *staff* não foi incluído dentro da equipe cirúrgica
- 1, se o cirurgião *staff* foi incluído dentro da equipe cirúrgica

e) Momento do dia em que começa a cirurgia.

Variável binária que coleta as informações relativas ao fator da seguinte forma,

- 0, se a cirurgia foi iniciada antes das 12:00 h
- 1, se a cirurgia foi iniciada após às 12:00 h

### **3.1.4.3 Fatores associados ao procedimento**

Existem diferentes técnicas que são utilizadas para cada procedimento. A seguir se descreve cada uma delas de acordo com o tipo de cirurgia que está sendo praticada.

#### **Mastectomia simples**

No caso da Mastectomia simples, do estudo dos relatórios de cirurgia foi identificado que durante a intervenção pode ser executada ou não a Biopsia Linfonodo Sentinela. Assim, um dos fatores a considerar nesse procedimento consiste de uma variável binária que contém as informações sobre o fato da BLS ter sido ou não realizada tal que,

- 0, se a BLS não foi realizada
- 1, se a a BLS foi realizada

## **Segmentectomia**

O fato de ter observações nas quais a Biopsia Linfonodo Sentinela (explicada no apartado anterior) resulta positiva, implica que se detectou nos relatórios de cirurgia a realização da **linfodectomia axilar**. Esta atividade consiste na drenagem do tecido canceroso da axila.

Além disso, outra atividade que se realiza na Segmentectomia habitualmente é a **ampliação de margens**. Como se explicou na descrição da cirurgia, o desafio do cirurgião consiste em retirar a mínima quantidade de mama possível, porém extraindo completamente o tumor. As vezes acontece que uma porção de tecido canceroso permanece nas margens da incisão feita pelo médico e se deve realizar a sua ampliação. Isto pode agregar também duração a cirurgia.

Considerando as possíveis variações da Segmentectomia obtidas na amostra, para a análise exploratória dos dados se gerou uma variável tal que,

- 1, se realizou a Segmentectomia sem BLS
- 2, se realizou a BLS, porém foi negativa
- 3, se realizou ampliação de margens
- 4, se realizou a linfodectomia axilar

No entanto, na análise de regressão serão consideradas variáveis binárias que recolham a informação de realização ou não de alguma das variações do procedimento.

## **3.2 Obtenção da informação**

### **3.2.1 O sistema de informação do hospital**

Existem certas dificuldades que impedem a extração de dados a partir do sistema de informação hospitalar. Eles são enumerados a seguir:

1. Não existem dados no sistema dos fatores a analisar (paciente, cirurgião).
2. Não há descrição adequada do procedimento realizado durante a cirurgia.

Esta circunstância implica a necessidade de recorrer à fonte de informação primária.

Em geral, há baixa confiabilidade a respeito das informações contidas no sistema de informação do hospital, de modo que desde o início foi decidido extrair as informações dos prontuários médicos.

### 3.2.2 O prontuário médico.

Para a análise dos fatores descritos, e sua influência no momento da cirurgia, obtivemos uma amostra aleatória obtida a partir de prontuários dos últimos três anos.

Um prontuário médico contém a soma das informações de um paciente hipotético durante a sua passagem através de uma instituição médica.

Inclui muitos relatórios, entre os quais se investigaram basicamente 3:

- **Relatório de cirurgia:** Nele figuram os dados básicos do paciente (ID, nome, data da cirurgia, etc.), o cirurgião responsável e a equipe cirúrgica, diagnóstico prévio, procedimento geral praticado e a descrição detalhada das atividades executadas durante a cirurgia.

É preenchido e carimbado por um dos cirurgiões envolvidos na operação.

- **Enfermagem / Centro cirúrgico:** Documento preenchido pela equipe de enfermagem que auxilia o cirurgião. Ele inclui informações sobre o paciente no seu passo através de todo o processo de cirurgia, uma vez que ele está hospitalizado (internamento) até que seja enviado para a fase de recuperação pós-operatória.

O registro inclui dados básicos pessoais (ID, nome, etc.), o diagnóstico de hipertensão arterial, diagnóstico de diabetes e cirurgia realizada, entre outros.

Também é o documento que contém dados sobre a duração das cirurgias.

- **Ambulatório:** Documento mais geral que descreve a evolução do paciente durante vários meses, e até mesmo anos. É cronologicamente estruturado, de modo que no espaço em relação a um ou dois dias antes da operação inclui várias informações sobre o estado físico do paciente, entre as quais se encontra o IMC dele.

## 3.3 Análise e interpretação dos resultados

### 3.3.1 Descrição da metodologia de análise

Inicialmente, os três procedimentos em estudo são mostrados, indicando a duração média da cirurgia e o desvio de cada uma das amostras obtidas.

Além disso, se apresentam gráficos de barras expondo os tamanhos de amostra de cada fator e os seus níveis.

Subsequentemente, para cada procedimento foi seguido um padrão de análise que consiste no seguinte:

#### a. Determinação da transformação de Box-Cox.

Entre as diversas transformações possíveis relativas à transformação de Box-Cox, foi determinada a transformação que derivava em um melhor ajuste da distribuição normal da variável transformada. Para permitir a comparação entre os três métodos, decidiu-se escolher a mesma transformação para todos eles (MRM, MS, SEG) sendo mesmo verificáveis os pressupostos básicos da regressão linear nas equações das três intervenções analisadas.

#### b. Análise exploratória de dados.

Apresentam-se os Box-Plot's das variáveis qualitativas, são mostradas as médias e se realizam testes de hipóteses de diferenças médias de cada nível do fator para a duração das cirurgias.

#### c. Descrição das variáveis consideradas na análise de regressão.

Nesse ponto se apresentam todas as variáveis a serem incluídas na análise posterior de regressão e sua natureza, bem seja binária, quantitativa ou de interação.

#### d. Análise de regressão e interpretação dos resultados.

Mostram-se as variáveis selecionadas através do método de *Stepwise Regression*, se inclui a análise da variância e a significância de cada parâmetro da regressão. Também se apresentam gráficos de resíduos para a verificação dos pressupostos da regressão.

Após, destacam-se as variáveis identificadas como significativas e, com base do sinal do coeficiente de regressão, se faz uma valoração de quais são finalmente os fatores que influenciam a duração das cirurgias.

### 3.3.2 Apresentação dos procedimentos considerados e os fatores de variação estudados

Como explicado na seção relativa à descrição da área de cada cirurgia de mama, os procedimentos envolvidos no estudo são os seguintes:

- **Mastectomia radical modificada (MRM).**
- **Mastectomia simples (MS).**
- **Segmentectomia (SEG)**

A tabela a seguir apresenta o tamanho, a média e desvio-padrão amostrais para cada um deles.

**Tabela 5. Tamanho, média e desvio padrão amostrais dos procedimentos analisados.**

| Procedimento realizado | N   | Média  | Desvio-padrão |
|------------------------|-----|--------|---------------|
| <b>MRM</b>             | 205 | 1:28 h | 0:26 h        |
| <b>MS</b>              | 130 | 1:09 h | 0:23 h        |
| <b>SEG</b>             | 160 | 1:07 h | 0:31 h        |

Como pode ser observado, o desvio da duração dos vários procedimentos é consideravelmente elevado, atingindo na Segmentectomia quase 50% da duração média da intervenção.

Além disso, uma observação importante a fazer é que justamente o procedimento que demora em média menos é o que apresenta maior variabilidade (SEG). A hipótese intuitiva, segundo a qual os procedimentos menos duráveis (em média) são menos complexos e, portanto, implicam menor variabilidade, não é comprovável nesse caso. Assim, as observações da amostra que contém as intervenções de Segmentectomia são, em média, as que possuem menor duração (01:06 h), e, apesar disso, têm um maior desvio padrão (31 minutos) maior.

A seguir se apresentam os tamanhos de amostra de todas as fontes de variação estudadas.

- IMC

A tabela a seguir resume o IMC médio em cada procedimento.

Tabela 6. IMC médio para cada procedimento.

| Procedimento | Média IMC |
|--------------|-----------|
| MRM          | 28,647    |
| MS           | 28,848    |
| SEG          | 29,305    |

Comprova-se como a média do IMC das pacientes que se submeteram a cada um dos três procedimentos está no intervalo de sobrepeso.

A figura a seguir mostra como os intervalos do IMC estão distribuídos entre as pacientes nas quais foram praticadas cada intervenção.

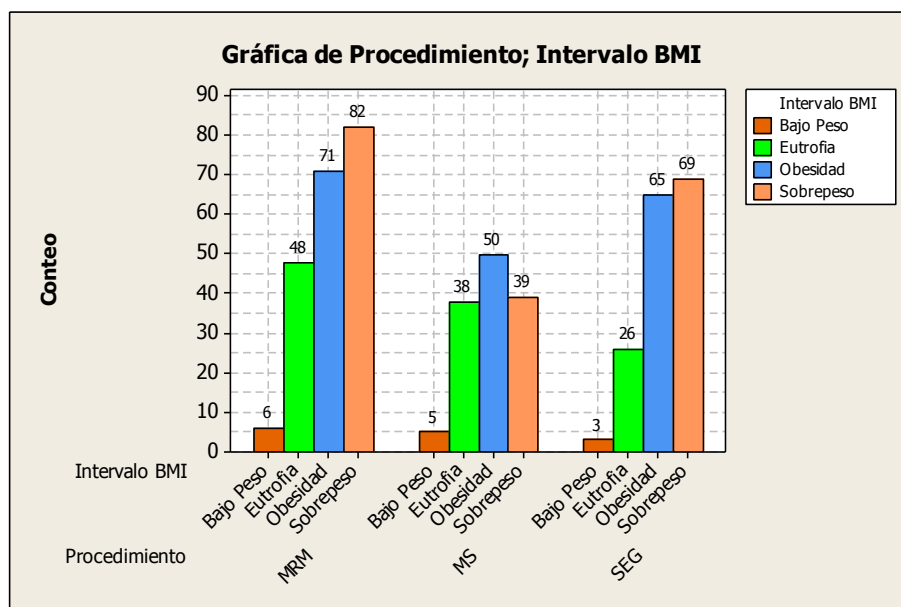


Figura 15. Distribuição dos intervalos de IMC em cada procedimento.

Percebe-se que a distribuição proporcional do estado nutricional de cada procedimento é semelhante em todos eles. Porém, o grupo mais numeroso na MRM e na SEG é o grupo de pacientes com obesidade e na MS é o grupo de pacientes com excesso de peso.

- Hipertensão

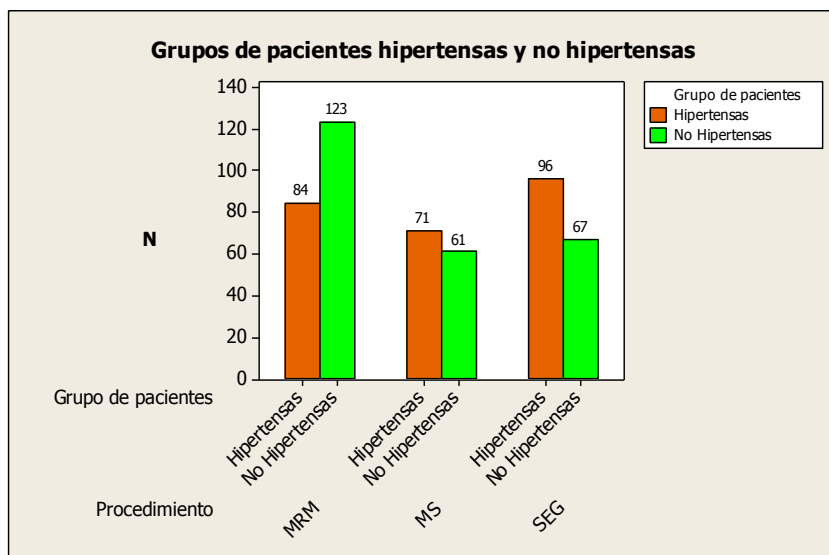


Figura 16. Grupos de pacientes hipertensos e não hipertensos.

Nos três podem ser observados, de um modo geral, que os dois grupos têm uma elevada percentagem de casos. Contudo, o grupo de hipertensos é proporcionalmente maior no caso da MRM, e no caso da MS e a SEC o grupo mais numeroso de pacientes são os hipertensos.

#### - Diabetes

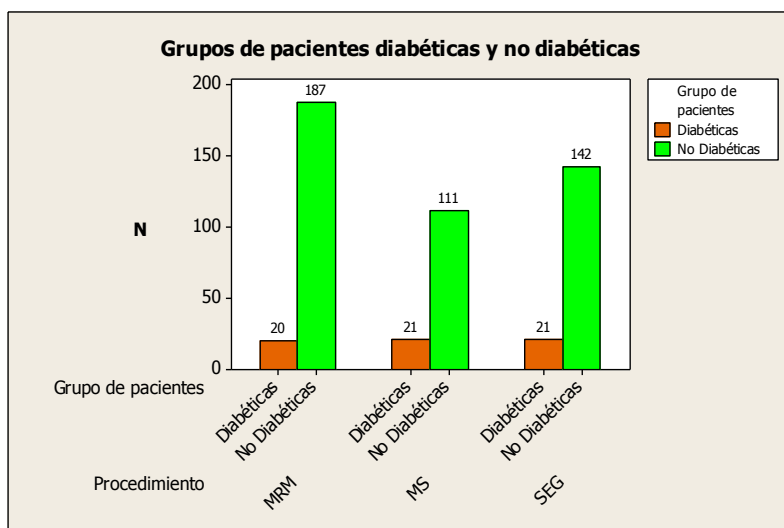


Figura 17. Grupos de pacientes diabéticos e não diabéticos.

Podemos ver como os pacientes diabéticos não representam um grupo muito grande no global de cada procedimento, sendo proporcionalmente maior na Mastectomia Simples, onde aproximadamente um 20% dos pacientes têm diabetes.

- Género do cirurgião responsável

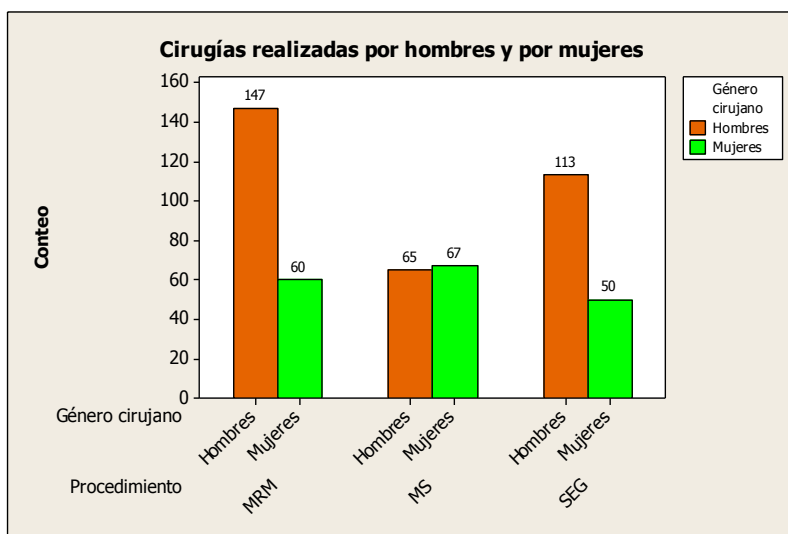


Figura 18. Cirurgias praticadas por homens e por mulheres.

Existe certa disparidade entre a MS, e a MRM e SEG. No caso da segunda, há um número aproximadamente igual de cirurgias executadas por homens e mulheres. Em contraste, nos outros dois procedimentos, a proporção de cirurgias realizadas por mulheres é significativamente mais baixa do que a dos homens.

- Composição da equipe cirúrgica

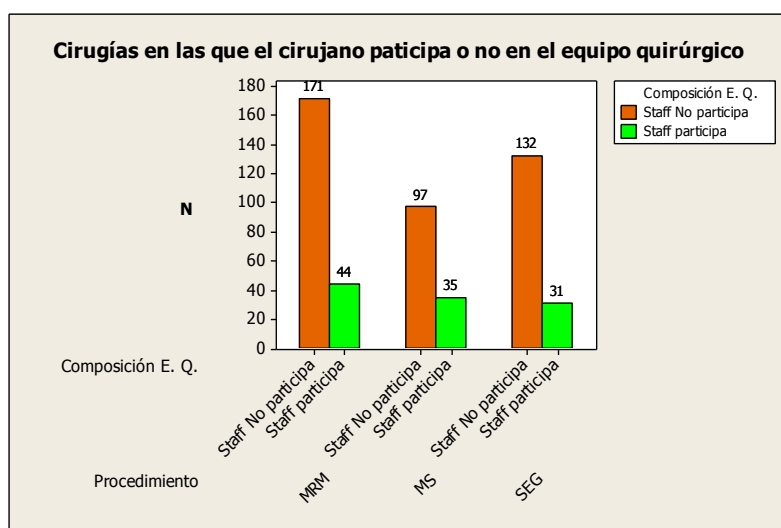


Figura 19. Cirurgias com as duas composições de equipes cirúrgicas consideradas.



Neste caso, a proporção de intervenções em que o cirurgião *staff* participa na equipe cirúrgica é semelhante nos três procedimentos, sem ser de qualquer um deles um grupo grande de observações.

#### - Começo da cirurgia

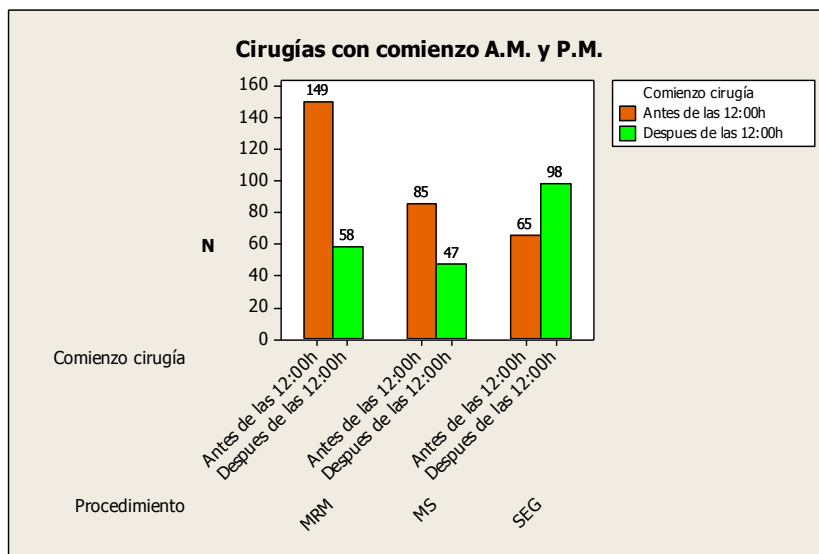
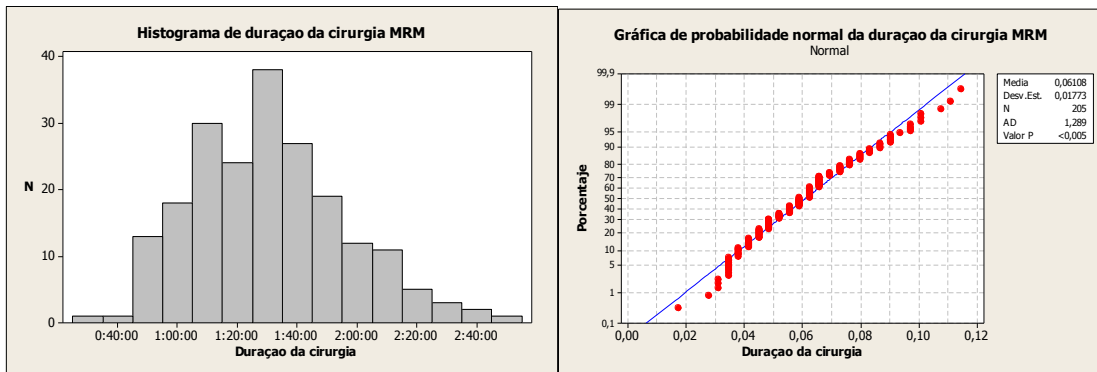


Figura 20. Grupos de cirurgias que começam antes e depois das 12:00 h.

Aqui vemos como no caso da SEG o grupo de intervenções que começaram com anterioridade às 12:00 h é proporcionalmente menor do que na MS e a MRM. Dado que é o procedimento que tem menor duração (em média, 20-30 minutos menos do que a MRM), ele pode ser programado mais tarde de modo que se tem tempo de realiza-lo durante as horas em que está aberta a sala de cirurgia.

### **3.3.3 Mastectomia Radical Modificada (MRM)**

A seguir se mostra um gráfico de probabilidade normal e um histograma dos dados de duração de cirurgia das diferentes observações de Mastectomia Radical Modificada.



**Figura 21. Histograma e gráfica de probabilidade normal dos dados de duração das cirurgias (MRM).**

Pode ser visto que a assimetria à direita não permite tratar os dados originais como se fossem normais. Observamos, também, que o valor-p inferior a 0,05 do teste de AD de normalidade rejeita a hipótese nula de que os dados provêm de uma distribuição Normal.

A tabela a seguir resume os resultados da melhor transformação de Box-Cox (Obtida de *Minitab*).

**Tabela 7. Transformação de Box-Cox para os dados (MRM)**

| Transformação   | Tipologia     | Coefficiente AD normalidade | p-valor |
|-----------------|---------------|-----------------------------|---------|
| $y' = \sqrt{y}$ | Raiz quadrada | 0,683                       | 0,090   |

Observe-se que com a transformação de raiz quadrada da resposta se obtém um p-valor maior ao predeterminado para rejeitar a hipótese nula de que os dados seguem uma distribuição normal.

Abaixo se encontra um gráfico de probabilidade normal dos dados transformados de acordo com a raiz quadrada da resposta.

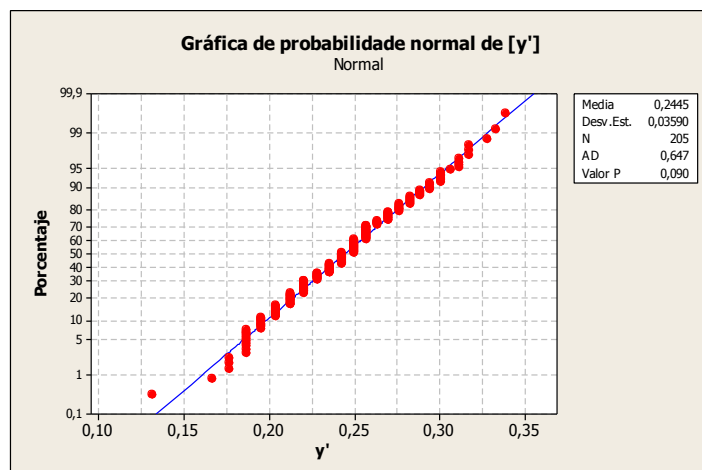


Figura 22. Gráfica de Probabilidade normal da variável transformada do tempo de cirurgia (MS)

Trabalhando-se com a transformação raiz quadrada dos tempos de duração no caso do procedimento SEG, conclui-se que é razoável supor a normalidade dos dados transformados.

### 3.2.3.1 Análise exploratória dos dados

#### Fatores associados ao paciente

- IMC

A tabela a seguir lista as observações que pertencem a cada um dos intervalos do IMC.

Tabela 8. Quantidade de cirurgias por intervalo de IMC (MRM).

| Intervalo IMC | N  |
|---------------|----|
| Baixo Peso    | 5  |
| Eutrofia      | 48 |
| Sobrepeso     | 71 |
| Obesidade     | 81 |

A seguir se apresenta um Box-Plot, já com os dados transformados, no qual se podem ver diferenças nas durações de acordo com a classificação do doente quanto ao IMC.

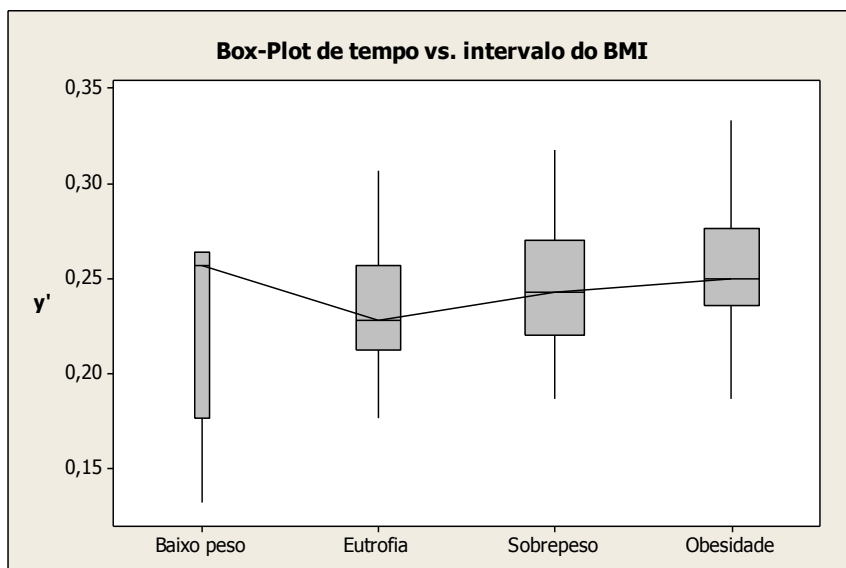


Figura 23. Box-Plot de duração da cirurgia vs. Intervalo de IMC (MRM)

Podemos ver como ele distingue uma diferença: quando o IMC do paciente aumenta, a duração da cirurgia também parece crescer. No caso do baixo peso só se tem 6 observações, o que é insuficiente para fazer qualquer verificação preliminar ou análise estatística.

As tabelas a seguir mostram a média (da variável transformada e a variável original) da duração de cada cirurgia dependendo do intervalo do IMC, e os resultados do teste de Tukey para os intervalos de confiança simultâneos (95% para a família)

Tabela 9. Média de duração da cirurgia para cada intervalo do IMC (MRM).

| Intervalo   | Média (variable normalizada) | Média (variable original) |
|-------------|------------------------------|---------------------------|
| Baixo Peso* | 0,227                        | 1:18 h                    |
| Eutrofia    | 0,232                        | 1:19 h                    |
| Sobrepeso   | 0,245                        | 1:28 h                    |
| Obesidade   | 0,254                        | 1:35 h                    |

**Tabela 10. Intervalos de confiança simultâneos de Tukey para a diferença de médias dos intervalos de IMC (MRM)**

| Diferença de médias         | Limite Inferior | Limite Superior | Contem o 0? |
|-----------------------------|-----------------|-----------------|-------------|
| Obesidade – Baixo Peso      | -0,015          | 0,069           | Sim         |
| Sobrepeso – Baixo Peso      | -0,024          | 0,059           | Sim         |
| Eutrofia – Baixo Peso       | -0,038          | 0,047           | Sim         |
| <b>Obesidade – Eutrofia</b> | <b>0,005</b>    | <b>0,039</b>    | <b>Não</b>  |
| Sobrepeso – Eutrofia        | -0,004          | 0,029           | Sim         |
| Obesidade – Sobrepeso       | -0,006          | 0,024           | Sim         |

Notamos que no teste de Tukey só indica como diferença significativa os tempos médios entre obesos e eutróficos.

Deve ser mencionado que se usou a variável IMC como categórica, o que é útil para ter uma visão geral da tendência e sua influência no momento da cirurgia.

Na análise de regressão é considerada como uma variável quantitativa (de acordo com o seu valor real), bem como todas as interações com ela.

Na análise de regressão se avaliará a sua influência real sobre a duração da cirurgia.

#### - Hipertensão

A tabela a seguir mostra o número de observações que pertencem a cada um dos dois grupos.

**Tabela 11. Quantidade observações de pacientes hipertensas e não hipertensas (MRM)**

| <b>Grupo de pacientes</b> | <b>N</b> |
|---------------------------|----------|
| Não hipertensas [0]       | 121      |
| Hipertensas [1]           | 84       |

E a seguir encontra-se o Box-Plot (variável transformada), que representa o caso em que a paciente é ou não hipertensa [1,0].

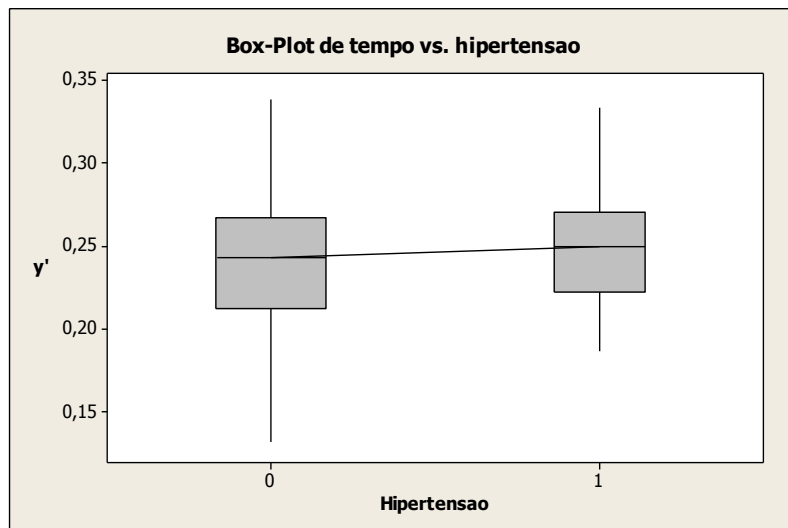


Figura 24. Box-Plot de duração de cirurgias vs. Hipertensão (MRM)

A diferença parece ser pequena, de modo que não parece ser um fator particularmente importante a considerar no estudo de tempo de cirurgia na MRM.

A seguir se mostram as médias nos dois casos.

Tabela 12. Média da duração das cirurgias de pacientes hipertensas e não hipertensas (MRM)

| Hipertensao                 | Média (variável transformada) | Média (variável original) |
|-----------------------------|-------------------------------|---------------------------|
| Paciente nao hipertensa [0] | 0,241                         | 1:26 h                    |
| Paciente hipertensa [1]     | 0,249                         | 1:31 h                    |

E a tabela a seguir mostra os limites do intervalo do 95% de confiança da diferença de médias para o teste  $t$ .

Tabela 13. Resultado do intervalo de 95% de confiança da diferença de médias entre pacientes hipertensas e não hipertensas (MRM)

| diferença                         | Límite inferior | Límite superior | Contiene el 0? |
|-----------------------------------|-----------------|-----------------|----------------|
| No hipertensa [0] –Hipertensa [1] | -0,018          | 0,001           | Sim            |

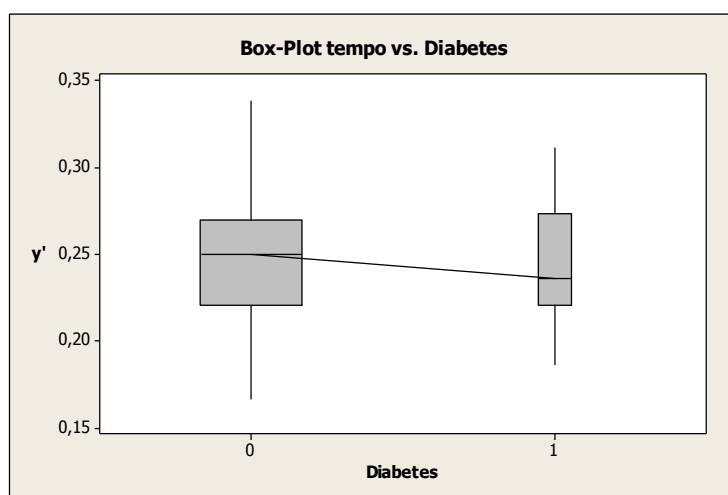
Embora exista uma ligeira diferença significada apreciada no Box-Plot no caso em que o paciente é hipertenso, não parece que seja um fator determinante de influência na duração da cirurgia.

- Diabetes.

Abaixo estão os tamanhos de amostra e o gráfico Box-Plot, que mostra a comparação entre o grupo de pacientes com diabetes [1] e o grupo de pacientes não-diabéticos [0].

**Tabela 14. Pacientes diabéticas e não diabéticas (MRM)**

| Grupo de pacientes | N   |
|--------------------|-----|
| Nao Diabéticas [0] | 185 |
| Diabéticas [1]     | 20  |



**Figura 25. Box-Plot pacientes diabéticas e não diabéticas (MRM)**

Pode-se observar como existe uma ligeira diferença na linha de conexão da mediana, mas pode ser principalmente devido à assimetria da distribuição dos dois grupos.

Na tabela abaixo se apresentam as médias de cada grupo.

**Tabela 15. Médias de duração das pacientes diabéticas e não diabéticas (MRM)**

| Diabetes                   | Média (variável transformada) | Média (variável original) |
|----------------------------|-------------------------------|---------------------------|
| Paciente nao diabética [0] | 0,085                         | 1:28 h                    |
| Paciente diabética [1]     | 0,091                         | 1:28 h                    |

O fato de a paciente ser ou não diabética não parece, portanto, um fator que seja significativo, uma vez que as médias dos dois grupos são iguais. Não se precisa incluir

o resultado do teste  $t$  de diferença, embora a variável seja levada em conta na análise de regressão posterior.

### **Fatores associados ao cirurgião**

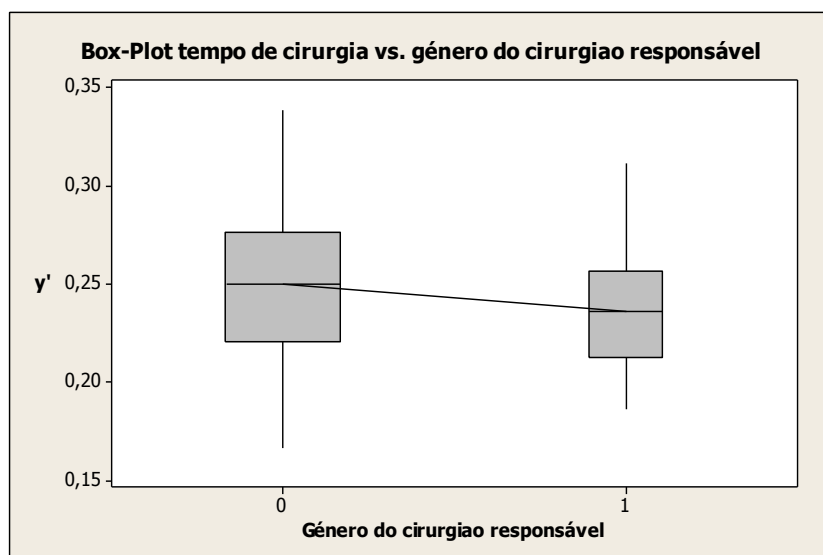
- Gênero do cirurgião.

A seguir se mostra a tabela com os tamanhos amostrais do grupo de cirurgias praticadas por homens e por mulheres.

**Tabela 16. Cirurgias realizadas por homens e por mulheres (MRM)**

| <b>Gênero do cirurgião</b> | <b>N</b> |
|----------------------------|----------|
| Homens [0]                 | 147      |
| Mulheres [1]               | 60       |

No Box-Plot a seguir pode-se observar a tendência na duração das cirurgias nos dois grupos.



**Figura 26. Box-Plot cirurgias realizadas por homens e por mulheres (MRM).**

Percebe-se diferença, o que indica maior velocidade das mulheres (1) na realização das cirurgias do que os homens (0).

A tabela a seguir mostra as médias.



**Tabela 17. Média das durações de cirurgias realizadas por homens e mulheres (MRM)**

| Gênero do cirurgião responsável | Média (variável transformada) | Média (variável original) |
|---------------------------------|-------------------------------|---------------------------|
| Homem [0]                       | 0,249                         | 1:31 h                    |
| Mulher [1]                      | 0,234                         | 1:20 h                    |

E a seguir se mostram os limites do intervalo de 95% de confiança da diferença das médias dos dois grupos.

**Tabela 18. Resultado do intervalo de 95% de confiança da diferença de médias de durações de cirurgias realizadas por homens e por mulheres (MRM).**

| Diferença gênero          | Limite inferior | Limite superior | Contem o 0? |
|---------------------------|-----------------|-----------------|-------------|
| Homens [0] – Mulheres [1] | 0,005           | 0,024           | No          |

Conclui-se que no procedimento de Mastectomia Radical o gênero do cirurgião pode ser um fator ativo. Nesta análise inicial verificou-se que as mulheres são 11 minutos, em média, mais rápidas do que os homens na hora de praticar a cirurgia.

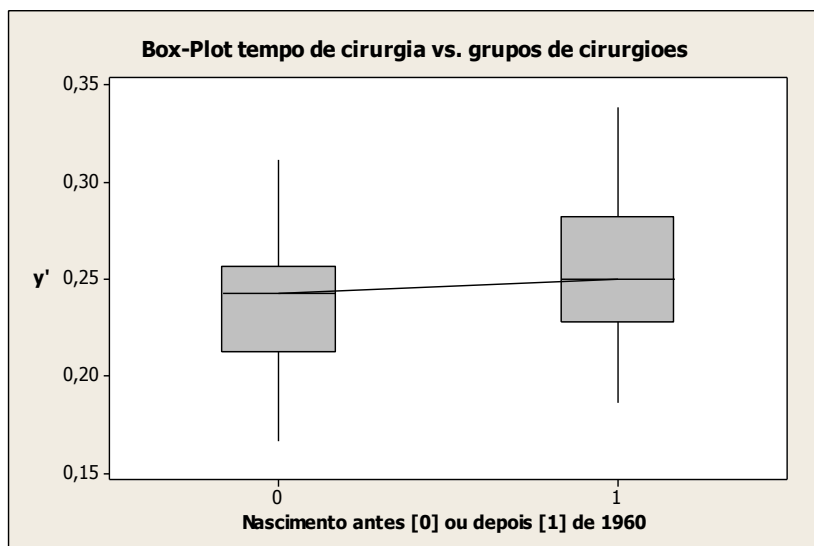
- Ano de nascimento do cirurgião

A tabela a seguir mostra os tamanhos amostrais do grupo de cirurgiões que nasceram depois e antes de 1960.

**Tabela 19. Cirurgias praticadas por cirurgiões que nasceram antes e depois de 1960 (MRM)**

| Grupo de cirurgiões         | N   |
|-----------------------------|-----|
| Nasceram antes de 1960 [0]  | 112 |
| Nasceram depois de 1960 [1] | 93  |

O Box-Plot a seguir mostra a tendência dos dois grupos.



**Figura 27. Box-Plot cirurgias realizadas por cirurgiões que nasceram antes e depois de 1960 (MRM)**

Parece existir uma duração menor da cirurgia quando é praticada por cirurgiões com mais experiência (e que nasceram antes de 1960) do que aqueles com menor idade (e que nasceram depois de 1960).

A seguir se mostram as médias e os limites do intervalo do 95% de confiança da diferença.

**Tabela 20. Médias das cirurgias realizadas por cirurgiões que nasceram antes e depois de 1960 (MRM)**

| Grupo de cirurgiões     | Média variável transformada | Média variável original |
|-------------------------|-----------------------------|-------------------------|
| Nasceram antes do 1960  | 0,235                       | 1:21 h                  |
| Nasceram depois do 1960 | 0,255                       | 1:35 h                  |

**Tabela 21. Resultado do intervalo de 95% de confiança da diferença de médias entre as cirurgias realizadas por cirurgiões que nasceram antes e depois de 1960 (MRM)**

| Diferença de médias              | Limite Inferior | Limite superior | Contem o o? |
|----------------------------------|-----------------|-----------------|-------------|
| Antes 1960 [0] – depois 1960 [1] | -0,030          | -0,010          | Não         |

Como a diferença é significativa, esse é um fator que pode ser importante.

- Trimestre da cirurgia

A seguir estão listados os tamanhos amostrais de cada grupo de cirurgias, de acordo com o trimestre no qual a intervenção foi realizada.

Tabela 22. Cirurgias realizadas nos diferentes trimestres do ano (MRM)

| Trimestre de realização da cirurgia | N  |
|-------------------------------------|----|
| Trimestre 1 [1]                     | 55 |
| Trimestre 2 [2]                     | 48 |
| Trimestre 3 [3]                     | 53 |
| Trimestre 4 [4]                     | 49 |

No seguinte Box-Plot se pode detectar a tendência das observações dependendo do seu trimestre de realização.

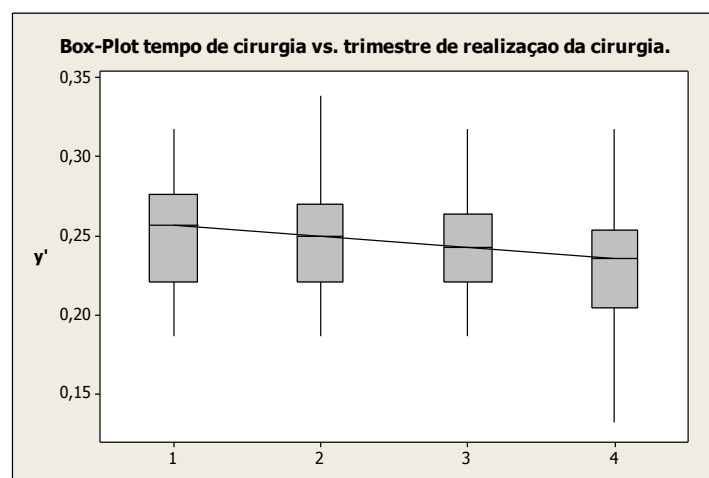


Figura 28. Box-Plot de cirurgias realizadas en los 4 trimestres del año (MRM)

Percebe-se uma diferença, de modo que no trimestre 4 as cirurgias demoram menos.

Tabela 23. Média de duração das cirurgias realizadas nos 4 trimestres do ano (MRM)

| Trimestre | Média (variável transformada) | Média (variável original) |
|-----------|-------------------------------|---------------------------|
| 1         | 0,250                         | 1:32 h                    |
| 2         | 0,249                         | 1:31 h                    |
| 3         | 0,245                         | 1:28 h                    |
| 4         | 0,232                         | 1:20 h                    |

A tabela a seguir mostra o resultado dos intervalos de confiança simultâneos de Tukey.

**Tabela 24. Resultado dos intervalos de confiança simultâneos de Tukey (95% de confiança para a família) de diferenças das médias das durações de cirurgias realizadas em cada trimestre do ano (MRM)**

| Diferença de médias       | Limite Inferior | Limite Superior | Contem o 0? |
|---------------------------|-----------------|-----------------|-------------|
| Trimestre 4 – trimestre 1 | -0,036          | 0+              | Sim         |
| Trimestre 3 – trimestre 1 | -0,023          | 0,012           | Sim         |
| Trimestre 2 – trimestre 1 | 0,019           | 0,017           | Sim         |
| Trimestre 4 – trimestre 2 | -0,035          | 0,002           | Sim         |
| Trimestre 3 – trimestre 2 | -0,022          | 0,014           | Sim         |
| Trimestre 4 – trimestre 3 | -0,030          | 0,005           | Sim         |

Não se detecta uma diferença significativa em nenhum deles. Porém o intervalo de diferença de médias do trimestre 1 e o trimestre 4 é muito próximo a ser significativo.

A análise de regressão determinará se é um fator que se deve levar em conta.

- Composição da equipe cirúrgica.

A continuação se mostram os tamanhos amostrais.

**Tabela 25. Cirurgias realizadas em que o cirurgião *staff* participa ou não na equipe cirúrgica (MRM).**

| <b>Equipe cirúrgica</b>                            | <b>N</b> |
|--|----------|
| <i>Staff</i> não participa na equipe cirúrgica (0) | 164      |
| <i>Staff</i> participa na equipe cirúrgica (1)     | 41       |

E a seguir se mostra o Box-Plot

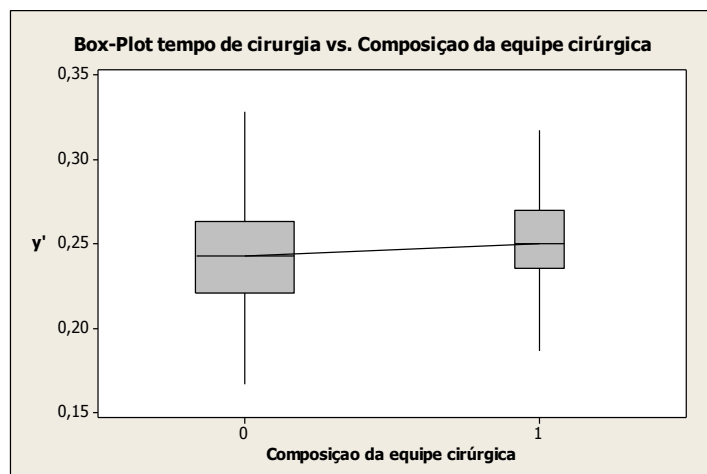


Figura 29. Box-Plot das cirurgias em que o cirurgião *staff* participa ou não na equipe cirúrgica (MRM)

Detecta-se diferença na duração das cirurgias quando o cirurgião *staff* participa da equipe cirúrgica.

A seguir se mostram as médias dos dois grupos e os limites do intervalo de 95% de confiança da diferença de médias.

Tabela 26. Médias das durações das cirurgias em que o cirurgião *staff* participou ou não na equipe cirúrgica (MRM)

| Composição da equipe cirúrgica           | Média (variável transf.) | Média (variável original) |
|--|--------------------------|---------------------------|
| <i>Staff</i> participa não participa [0] | 0,242                    | 1:26 h                    |
| <i>Staff</i> participa [1]               | 0,254                    | 1:35 h                    |

E a seguir se mostra o resultado do intervalo de confiança de diferença de médias.

Tabela 27. Resultado do intervalo de 95% de confiança da diferença das médias de cirurgias em que o cirurgião *staff* participou ou não na equipe cirúrgica (MRM)

| Diferença composição    | Límite inferior | Límite superior | Contiene el 0? |
|-------------------------|-----------------|-----------------|----------------|
| equipe [0] – equipe [1] | -0,025          | 0,001           | Sim            |

Não se detectam diferenças significativas, porém pode ser um fator de interesse.

- Momento do dia em que começa a cirurgia

A tabela a seguir resume os tamanhos amostrais do conjunto de observações em que a cirurgia começou com anterioridade às 12:00 h [0] e o grupo no qual a intervenção começou depois das 12:00 h [1].

Tabela 28. Cirurgias que começaram antes e depois das 12:00 h (MRM)

| Começo da cirurgia                      | N   |
|---|-----|
| Cirurgia começou antes das 12:00 h [0]  | 147 |
| Cirurgia começou depois das 12:00 h [1] | 58  |

O Box-Plot a seguir ilustra a possível diferença em relação ao fato.

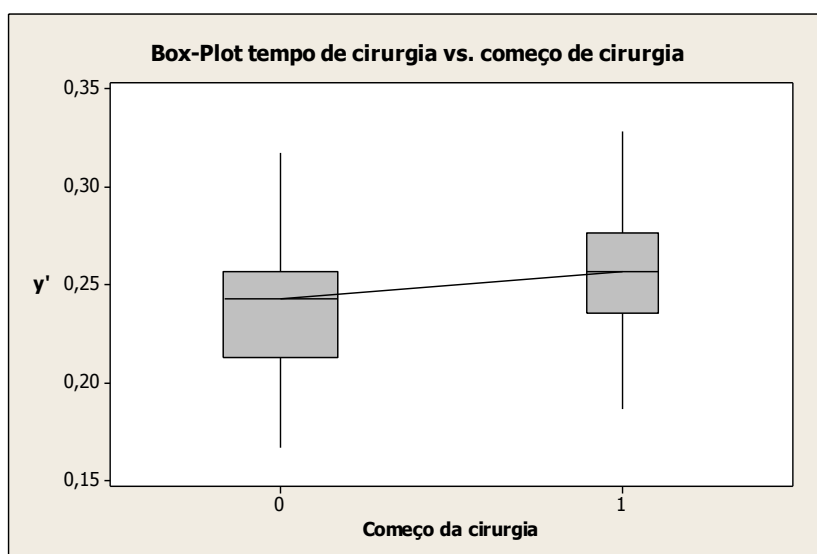


Figura 30. Box-Plot de cirurgias que começaram antes e depois das 12:00 h (MRM)

Pode-se notar um aumento na duração da intervenção em aquelas cirurgias que começaram depois de meio-dia.

As médias são mostradas na tabela a seguir.

Tabela 29. Média das durações de cirurgias que começaram antes e depois das 12:00 h (MRM)

| Começo da cirurgia         | Média (variável transf.) | Média (variável original) |
|----------------------------|--------------------------|---------------------------|
| Antes de las 12:00 h [0]   | -0,052                   | 1:24 h                    |
| Después de las 12:00 h [1] | 0,440                    | 1:37 h                    |

A tabela a seguir mostra os limites do intervalo de confiança de 95%.

Tabela 30. Resultado do intervalo de 95% de confiança da diferença de médias das durações de cirurgias que começaram antes e depois das 12:00 h (MRM)

| Diferença começo cirurgia     | Limite inferior | Limite superior | Contem o 0? |
|-------------------------------|-----------------|-----------------|-------------|
| Começo AM [0] – Começo PM [1] | - 0,018         | - 0,008         | Nao         |

De modo que este pode ser um fator importante a levar em conta na equação de regressão final. Se percebe que as cirurgias que começaram após 12:00 h demoram, em média, 13 minutos mais.

### 3.2.3.2 Interações consideradas no modelo de regressão linear.

A fim de refinar a análise de regressão se consideraram todas as possíveis interações entre os potenciais fatores que influenciam na duração da cirurgia.

A tabela a seguir mostra as variáveis de interação que foram criadas para a análise posterior.

O fator trimestre, de quatro níveis, foi subdividido em quatro variáveis binárias que consistem em realizar a cirurgia ou não em qualquer dos trimestres do ano (T1, T2, T3, T4).

**Tabela 31. Variáveis de Interação consideradas no modelo de Regressão Linear (MRM)**

|                        | IMC | Hiper   | Diab.   | Gênero  | Ano nascim. | TRI1     | TRI2     | TRI3     | TRI4     | Começo cirurgia | Comp equipe cirúrgica |
|------------------------|-----|---------|---------|---------|-------------|----------|----------|----------|----------|-----------------|-----------------------|
| IMC                    |     | IMCxHIP | IMCxDIA | IMCxGEN | IMCxANO     | IMCxTRI1 | IMCxTRI2 | IMCxTRI3 | IMCxTRI4 | IMCxCOMI        | IMCxCOMMP             |
| Hipert.                |     |         | HIPxDIA | HIPxGEN | HIPxANO     | HIPxTRI1 | HIPxTRI2 | HIPxTRI3 | HIPxTRI4 | HIPxCOMI        | HIPxCOMMP             |
| Diabetes               |     |         |         | DIAxGEN | DIAxANO     | DIAxTRI1 | DIAxTRI2 | DIAxTRI3 | DIAxTRI4 | DIAxCOMI        | DIAxCOMMP             |
| Gênero                 |     |         |         |         | GENxANO     | GENxTRI1 | GENxTRI2 | GENxTRI3 | GENxTRI4 | GENxCOMI        | GENxCOMMP             |
| Ano nascimento         |     |         |         |         |             | TRI1xAÑO | TRI2xAÑO | TRI3xAÑO | TRI4xAÑO | ANOxCOMI        | ANOxCOMMP             |
| TRI1                   |     |         |         |         |             |          | -        | -        | -        | TRI1xCOMI       | TRI1xCOMMP            |
| TRI2                   |     |         |         |         |             |          |          | -        | -        | TRI2xCOMI       | TRI2xCOMMP            |
| TRI3                   |     |         |         |         |             |          |          |          | -        | TRI3xCOMI       | TRI3xCOMMP            |
| TRI4                   |     |         |         |         |             |          |          |          |          | TRI4xCOMI       | TRI4xCOMMP            |
| Começo cirurgia        |     |         |         |         |             |          |          |          |          |                 | COMIxCOMMP            |
| Comp. Equipe cirúrgica |     |         |         |         |             |          |          |          |          |                 |                       |

De modo geral, se podem fazer duas interpretações diferentes das interações.

- **Interações de variáveis Dummy:** Estas interações resumem a informação de dois estados simultaneamente. Por exemplo, a variável de interação dummy GENxTRI3 coletaria informações sobre as cirurgias realizada por mulheres durante o terceiro trimestre.

- **Interações de variáveis quantitativas:** Estas ocorrem quando interagindo as variáveis quantitativas (IMC e Ano de nascimento) entre elas ou com as demais variáveis. Assim, por exemplo, uma influência positiva na análise de regressão da variável IMCxCOMP indicaria que quando a equipa cirúrgica é composta por cirurgiões *staff* como auxiliares, o IMC influencia no tempo da cirurgia.

### 3.2.3.3 Análise de Regressão e interpretação dos resultados

Foram, portanto, consideradas inicialmente 60 variáveis explicativas, a natureza das quais é descrita na tabela a seguir.

**Tabela 32. Natureza das variáveis originais e de interação consideradas no modelo de Regressão Linear (MRM)**

| Natureza da variável | Variáveis originais | Variáveis de interação |
|----------------------|---------------------|------------------------|
| Dummy                | 9                   | 30                     |
| Quantitativa         | 2                   | 19                     |

Neste ponto será discutido o grau de explicação geral da variabilidade associada com o tempo de cirurgia, o que irá determinar a incerteza que tem um eventual planejador em relação à duração da Mastectomia Radical Modificada.

### Stepwise Regression

A seguir se mostra a saída de *Minitab* para a Stepwise Regression ( $\alpha_{entrada} = 0,05$ ;  $\alpha_{saida} = 0,05$ ) levando em conta as variáveis nomeadas acima.



60 variáveis, com N = 205

| Paso               | 1       | 2       | 3       | 4        | 5        | 6        |
|--------------------|---------|---------|---------|----------|----------|----------|
| Constante          | 0,1851  | 0,1807  | 0,1849  | 0,1899   | -0,9205  | -0,9091  |
| IMCxANO            | 0,00000 | 0,00000 | 0,00000 | 0,00000  | 0,00000  | 0,00000  |
| Valor T            | 4,35    | 4,42    | 4,46    | 4,51     | 4,14     | 3,57     |
| Valor P            | 0,000   | 0,000   | 0,000   | 0,000    | 0,000    | 0,000    |
| COMIxANO           |         | 0,00001 | 0,00001 | 0,00001  | 0,00001  | 0,00001  |
| Valor T            |         | 3,37    | 3,52    | 3,55     | 2,85     | 3,02     |
| Valor P            |         | 0,001   | 0,001   | 0,000    | 0,005    | 0,003    |
| Trimestre 4        |         |         | -0,0160 | -0,0179  | -0,0182  | -0,0184  |
| Valor T            |         |         | -2,97   | -3,35    | -3,44    | -3,51    |
| Valor P            |         |         | 0,003   | 0,001    | 0,001    | 0,001    |
| IMCxGEN            |         |         |         | -0,00051 | -0,00047 | -0,00052 |
| Valor T            |         |         |         | -2,88    | -2,68    | -2,94    |
| Valor P            |         |         |         | 0,004    | 0,008    | 0,004    |
| ANONASC            |         |         |         |          | 0,00057  | 0,00057  |
| Valor T            |         |         |         |          | 2,22     | 2,22     |
| Valor P            |         |         |         |          | 0,028    | 0,027    |
| HIPxCOMP           |         |         |         |          |          | 0,0173   |
| Valor T            |         |         |         |          |          | 2,07     |
| Valor P            |         |         |         |          |          | 0,040    |
| S                  | 0,0344  | 0,0336  | 0,0329  | 0,0324   | 0,0321   | 0,0318   |
| R-cuad.            | 8,54    | 13,40   | 17,04   | 20,35    | 22,27    | 23,91    |
| R-cuad. (ajustado) | 8,08    | 12,54   | 15,80   | 18,76    | 20,31    | 21,60    |

Pode-se ver que as iterações são realizadas de forma aditiva, e em nenhum momento são removidas variáveis. O resumo das diferentes etapas é o seguinte:

- **Passo 1:** A variável *IMCxANO* é adicionada
- **Passo 2:** A variável *COMIxANO* é adicionada
- **Passo 3:** A variável *Trimestre 4* é adicionada
- **Passo 4:** A variável *IMCxGEN* é adicionada
- **Passo 5:** A variável *ANONASC* é adicionada
- **Passo 6:** A variável *HIPxCOMP* é adicionada

E a seguir se mostra a análise de variância para a regressão.

## Análise da variância

| Fuente         | GL  | SC       | MC       | F     | P  |
|----------------|-----|----------|----------|-------|----|
| Regresión      | 6   | 0,062878 | 0,010480 | 10,37 | 0+ |
| Error residual | 198 | 0,200102 | 0,001011 |       |    |
| Total          | 204 | 0,262980 |          |       |    |

Que resulta significativa.

E a continuação se mostram os diferentes coeficientes de regressão individuais para cada fator, a estatística *t*, o p-valor e os fatores de inflação da variância (VIF). Além disso, se fornecem o coeficiente de determinação da regressão normal e ajustado (explicação total de variabilidade) e o coeficiente de determinação da previsão ajustado.

| Variável     | Coef       | SE Coef    | T     | P     | VIF   |
|--------------|------------|------------|-------|-------|-------|
| Constante    | -0,9091    | 0,4973     | -1,83 | 0,069 |       |
| IMCxANONASC  | 0,00000084 | 0,00000023 | 3,57  | 0,000 | 1,089 |
| COMIxANONASC | 0,00000789 | 0,00000262 | 3,02  | 0,003 | 1,087 |
| Trimestre 4  | -0,018420  | 0,005250   | -3,51 | 0,001 | 1,017 |
| IMCxGEN      | -0,0005196 | 0,0001765  | -2,94 | 0,004 | 1,042 |
| ANONASC      | 0,0005657  | 0,0002547  | 2,22  | 0,027 | 1,120 |
| HIPxCOMP     | 0,017327   | 0,008382   | 2,07  | 0,040 | 1,084 |

S = 0,0317902 R-cuad. = 23,9% R-cuad. (ajustado) = 21,6%

PRESS = 0,215088 R-cuad. (pred) = 18,21%

Percebe-se que os sinais dos coeficientes da regressão concordam com a análise individual dos fatores, de modo que o IMC, o ano de nascimento e a composição da equipe têm sinal positivo; ou seja, com o aumento do valor das variáveis de previsão mencionadas, aumenta a duração da cirurgia. Além disso, o IMC e o “Ano de Nascimento” aparecem como interação. A interpretação é que quando os cirurgiões são menos experimentados (tem menos idade) tem mais contribuição positiva à duração da cirurgia quando o IMC cresce.

Mesmo fenómeno acontece no caso em que as cirurgias começam depois das 12:00 h. A interação *COMIxANO* mostra que, quando os cirurgiões são menos experimentados, as cirurgias que começaram depois das 12:00 h demoram mais.

Percebe-se também que as cirurgias que se fazem no trimestre 4 demoram menos, presumivelmente pela experiência acumulada dos cirurgiões residentes ao longo do ano.

O sinal negativo da variável *IMCxGEN* indica que as mulheres do hospital realizam cirurgias que demoram menos do que quando as realizam os homens, quando o IMC do paciente aumenta. Isso não significa que a maior IMC menor duração da cirurgia. Em realidade, essa variável atua como um “fator de correção”: o fato de a primeira variável detectada como influente na duração da cirurgia tenha sido (em interação com o ano de nascimento) o *IMC*, indica que o IMC em todos os casos é um fator que faz aumentar a duração da cirurgia. Porém, quando as mulheres realizam cirurgias e o *IMC* aumenta de umas a outras, o aumento da demora das cirurgias é menor do que nos homens.

Por último, a variável *HIPxCOMP* indica que quando a equipe cirúrgica é composta por cirurgiões *staff* como auxiliares, o fato da paciente ser hipertensa é um fator que agrega tempo de duração da cirurgia.

Também se pode verificar que os fatores de inflação da variância são próximos a 1, de modo que aqui não há colinearidade sensível entre as variáveis da regressão.

Finalmente, conclui-se que se pode explicar de 21,6% por cento da variabilidade total, e se tem um grau de previsão do 18,21%.

Nos gráficos de resíduos a seguir se podem ser verificados os pressupostos de normalidade e homocedasticidade para a regressão.

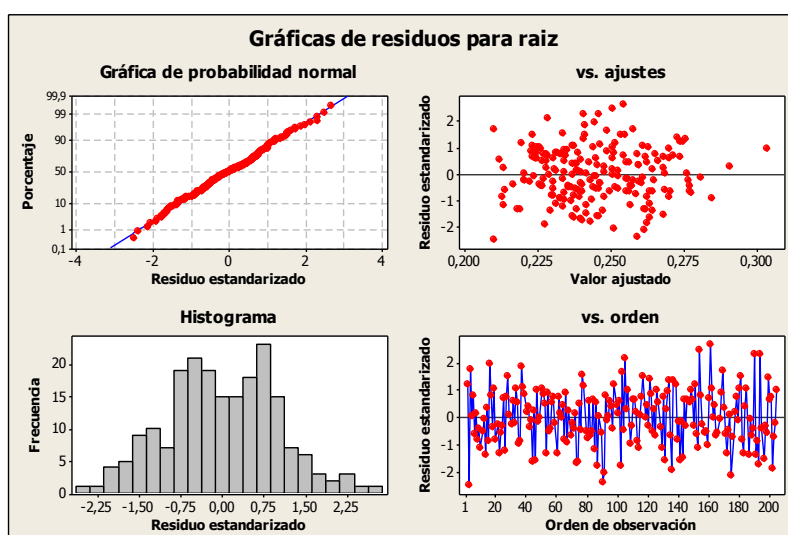


Figura 31. Gráficos de resíduos da Regressão Linear (MRM)

Onde se verificou que o comportamento dos resíduos é aproximadamente normal, e não se percebe uma tendência evidente no gráfico de dispersão de resíduos vs. valores ajustados.

A equação de regressão final é como se segue:

$$Y' = -0,909 + 0,000001 * \mathbf{IMCxANO} + 0,000008 * \mathbf{COMIxANO} - 0,0184 * \mathbf{Trimestre\ 4} - \\ - 0,000520 * \mathbf{IMCxGEN} + 0,000566 * \mathbf{ANONASC} + 0,0173 * \mathbf{HIPxCOMP}$$

### 3.3.4 Mastectomia Simples (MS)

A seguir se mostra o gráfico de probabilidade normal e o histograma dos dados de duração da Mastectomia Simples.

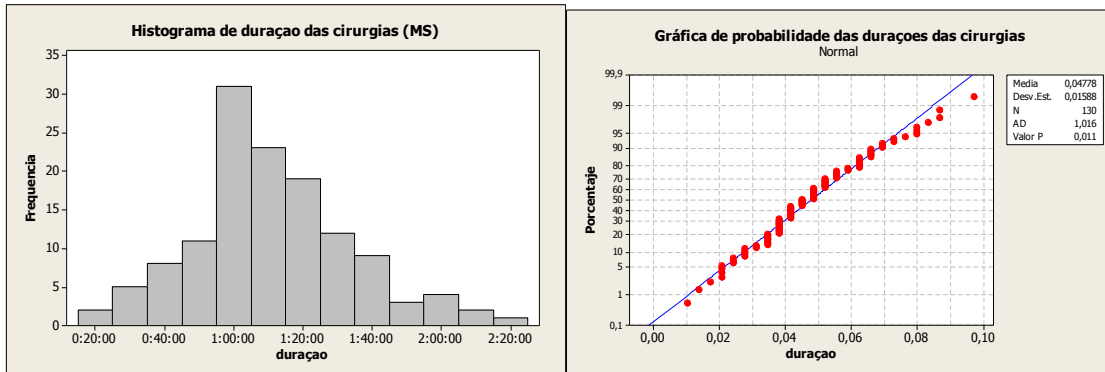


Figura 32. Histograma e gráfica de probabilidade normal dos dados de duração das cirurgias (MS)

Embora, neste caso, a normalidade dos dados originais é maior do que no caso da MRM, note-se que o coeficiente de AD ainda faz rejeitar a hipótese nula no teste de normalidade (com um p-valor de 0,011).

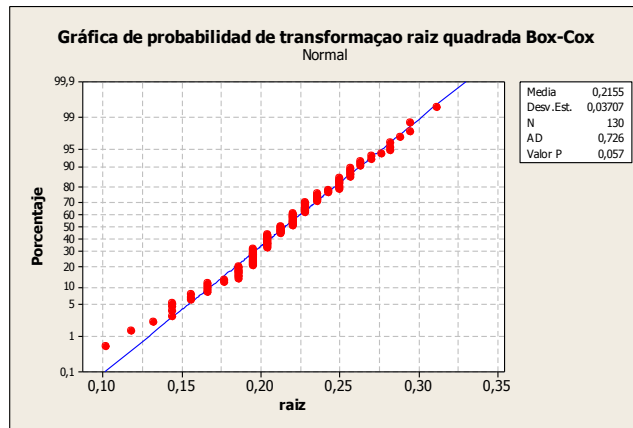
A tabela a seguir resume os resultados da melhor transformação de Box-Cox (Obtida de *Minitab*).

Tabela 33, Transformação de Box-Cox para os dados (MS)

| Transformação   | Tipologia     | Coefficiente AD normalidade | p-valor |
|-----------------|---------------|-----------------------------|---------|
| $y' = \sqrt{y}$ | Raiz quadrada | 0,726                       | 0,057   |

Com a transformação de raiz quadrada da resposta se obtém um p-valor maior ao predeterminado (0,05) para rejeitar a hipótese nula de que os dados seguem uma distribuição normal.

A seguir se mostra o teste de normalidade da variável transformada segundo transformação de Box-Cox de raiz quadrada.



**Figura 33. Gráfica de Probabilidade normal da variável transformada do tempo de cirurgia (MS)**

Trabalhando-se com a transformação raiz quadrada dos tempos de duração no caso do procedimento MS, conclui-se que é razoável supor a normalidade dos dados transformados.

### 3.3.4.1 Análise exploratória dos dados

#### Fatores associados ao paciente

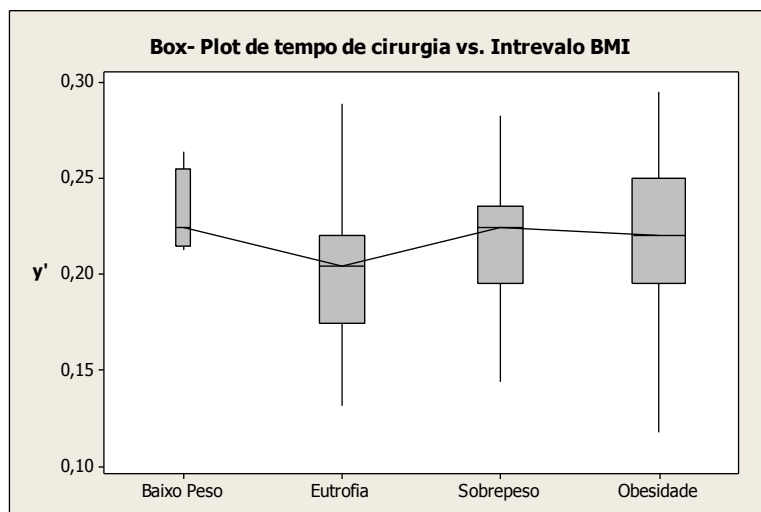
- IMC

A tabela a seguir lista as observações que pertencem a cada um dos intervalos do IMC.

**Tabela 34. Quantidade de cirurgias por intervalo de IMC (MS)**

| <b>Intervalo IMC</b> | <b>N</b> |
|----------------------|----------|
| Baixo Peso           | 4        |
| Eutrofia             | 38       |
| Sobrepeso            | 38       |
| Obesidade            | 50       |

A seguir apresenta-se um Box-Plot, já com os dados normalizados, onde se podem ver diferenças nas durações de acordo com se o doente pertence a um intervalo nutricional ou outro.



**Figura 34. Box-Plot de duração da cirurgia vs. Intervalo de IMC (MS)**

Neste caso, não existe uma diferença muito significativa que apoie a hipótese de que quanto maior IMC do paciente, as cirurgias são mais longas.

As tabelas a seguir mostram a média (da variável transformada e a variável original) da duração de cada cirurgia dependendo do intervalo do IMC, e os resultados do teste de Tukey para os intervalos de confiança simultâneos (95% para a família)

**Tabela 35. Média de duração de cirurgia para cada intervalo do IMC (MS)**

| Intervalo   | Média (variable normalizada) | Média (variable original) |
|-------------|------------------------------|---------------------------|
| Baixo Peso* | 0,231                        | 1:17 h                    |
| Eutrofia    | 0,201                        | 1:00 h                    |
| Sobrepeso   | 0,2208                       | 1:12 h                    |
| Obesidade   | 0,2213                       | 1:13 h                    |

**Tabela 36. Intervalos de confiança simultâneos de Tukey (95% para a família) de diferença de médias dos intervalos de IMC (MS)**

| Diferença de médias         | Limite Inferior | Limite Superior | Contem o 0? |
|-----------------------------|-----------------|-----------------|-------------|
| Obesidade – Baixo Peso      | -0,059          | 0,039           | Sim         |
| Sobrepeso – Baixo Peso      | -0,059          | 0,039           | Sim         |
| Eutrofia – Baixo Peso       | -0,080          | 0,019           | Sim         |
| <b>Obesidade – Eutrofia</b> | <b>0+</b>       | <b>0,041</b>    | <b>Não</b>  |
| Sobrepeso – Eutrofia        | 0,002           | 0,042           | Sim         |
| Obesidade – Sobrepeso       | -0,020          | 0,020           | Sim         |

Como no caso da MRM, na Mastectomia Simples se detectam diferenças significativas entre as médias das intervenções de pacientes obesos e eutróficos, sendo as cirurgias realizadas a pacientes obesas, em média, 13 minutos mais longas.

Na análise de regressão se observará se é um fator importante de previsão de novas observações.

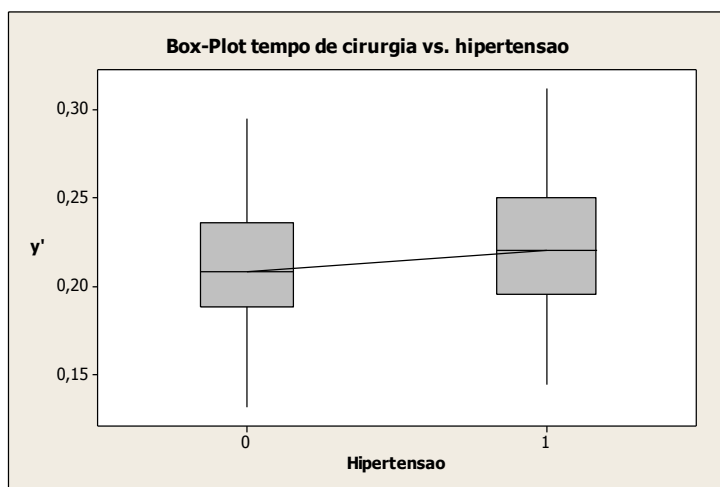
- Hipertensão

A tabela a seguir mostra o número de observações que pertencem a cada um dos dois grupos.

**Tabela 37. Pacientes hipertensas e não hipertensas (MS)**

| <b>Grupo de pacientes</b> | <b>Nº de observaciones</b> |
|---------------------------|----------------------------|
| No hipertensas [0]        | 121                        |
| Hipertensas [1]           | 84                         |

E a seguir encontra-se o Box-Plot (variável transformada), que representa os casos em que a paciente é ou não é hipertensa [1,0].



**Figura 35. Box-Plot de duração das cirurgias vs. hipertensão (MS)**

Existe uma diferença de modo que, se o paciente é hipertenso, a cirurgia leva mais tempo. Porém, a pendente não parece ser excessivamente pronunciada.

A seguir se mostram as médias nos dois casos.

**Tabela 38. Média das durações de cirurgia de pacientes hipertensas e não hipertensas (MS)**

| Hipertensao | Média (variável transformada) | Média (variável original) |
|-------------|-------------------------------|---------------------------|
|             |                               |                           |



|                             |       |        |
|-----------------------------|-------|--------|
| Paciente nao hipertensa [0] | 0,211 | 1:06 h |
| Paciente hipertensa [1]     | 0,220 | 1:11 h |

E a tabela a seguir mostra os limites do intervalo de 95% de confiança da diferença de médias para o teste  $t$ .

**Tabela 39. Resultado do intervalo de 95% de confiança da diferença de médias de duração de cirurgias de pacientes hipertensas e não hipertensas (MS).**

| diferença                         | Límite inferior | Límite superior | Contiene el 0? |
|-----------------------------------|-----------------|-----------------|----------------|
| No hipertensa [0] –Hipertensa [1] | -0,022          | 0,004           | Sim            |

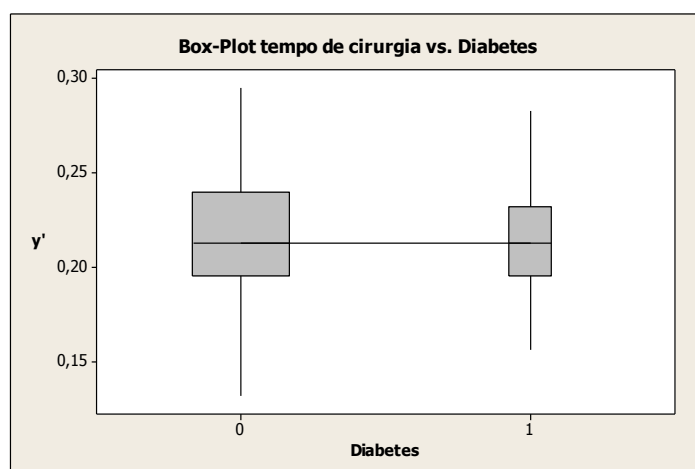
Embora exista uma ligeira diferença apreciada no Box-Plot no caso em que a paciente é hipertensa, não parece que seja um fator determinante de influência na duração da cirurgia.

- Diabetes.

A seguir se apresentam os tamanhos amostrais e o gráfico Box-Plot, que mostra a comparação entre o grupo de pacientes com diabetes [1] e o grupo de pacientes não-diabéticos [0].

**Tabela 40. Pacientes diabéticas e não diabéticas (MS)**

| Grupo de pacientes | N   |
|--------------------|-----|
| Nao Diabéticas [0] | 109 |
| Diabéticas [1]     | 21  |



**Figura 36. Box-Plot pacientes diabéticas e não diabéticas (MS).**

Não se percebe diferença entre os dois grupos de pacientes.

Na tabela abaixo podemos ver as médias de cada grupo:

**Tabela 41. Médias de duração das pacientes diabéticas e não diabéticas (MS)**

| Diabetes                   | Média (variável transformada) | Média (variável original) |
|----------------------------|-------------------------------|---------------------------|
| Paciente nao diabética [0] | 0,214                         | 1:08 h                    |
| Paciente diabética [1]     | 0,221                         | 1:12 h                    |

E a seguir se mostram os limites do intervalo do teste  $t$  de diferença de médias.

**Tabela 42. Resultado do intervalo de 95% de confiança de diferença das médias de duração de cirurgias de pacientes diabéticas e não diabéticas (MS)**

| Diferença                           | Límite inferior | Límite superior | Contiene el 0? |
|-------------------------------------|-----------------|-----------------|----------------|
| Nao diabéticos [0] – Diabéticos [1] | -0,024          | 0,010           | Sim            |

O fato de a paciente ser ou não diabética não parece, portanto, um fator que seja significativo, embora a variável seja levada em conta na análise de regressão posterior.

### **Fatores associados ao cirurgião**

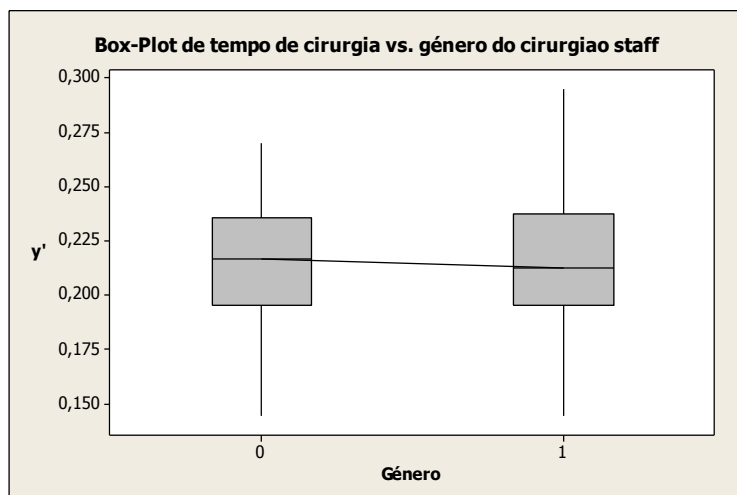
- Género do cirurgião.

A seguir se mostra a tabela com os tamanhos amostrais do grupo de cirurgias praticadas por homens e por mulheres.

**Tabela 43. Cirurgias realizadas por homens e por mulheres (MS)**

| Género do cirurgião | N  |
|---------------------|----|
| Homens [0]          | 64 |
| Mulheres [1]        | 66 |

No Box-Plot a seguir pode-se observar a tendência na duração das cirurgias nos dois grupos.



**Figura 37. Box-Plot cirurgias realizadas por homens e por mulheres (MS)**

Percebe-se uma leve diferença, o que indica maior velocidade das mulheres (1) na realização das cirurgias do que os homens (0).

A tabela a seguir mostra as médias.

**Tabela 44. Média das durações de cirurgias realizadas por homens e mulheres (MS)**

| Género do cirurgiao responsável | Média (variável transformada) | Média (variável original) |
|---------------------------------|-------------------------------|---------------------------|
| Homem [0]                       | 0,215                         | 1:08 h                    |
| Mulher [1]                      | 0,216                         | 1:09 h                    |

Do que resulta que, como existe uma diferença mínima entre os dois grupos, não é necessário realizar o teste *t* para a diferença das durações porque as médias são equivalentes quando a intervenção cirúrgica é efetuada por um homem ou uma mulher.

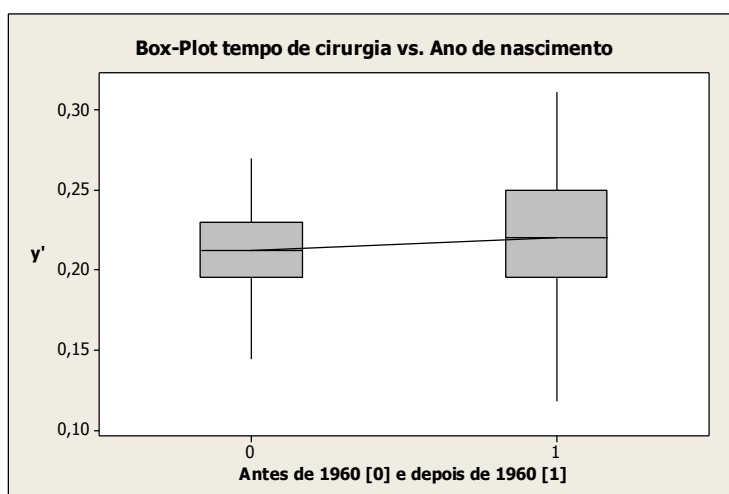
- Ano de nascimento do cirurgião

A tabela a seguir mostra os tamanhos amostrais do grupo de cirurgiões que nasceu depois e antes de 1960.

**Tabela 45. Cirurgias praticadas por cirurgiões que nasceram antes e depois de 1960 (MS)**

| Grupo de cirurgiões         | N  |
|-----------------------------|----|
| Nasceram antes de 1960 [0]  | 66 |
| Nasceram depois de 1960 [1] | 64 |

O Box-Plot a seguir mostra a tendência dos dois grupos.



**Figura 38. Cirurgias realizadas por cirurgiões que nasceram antes e depois de 1960 (MS)**

Percebe-se não só uma leve diferença na duração das cirurgias, mas também maior variabilidade nas intervenções dos cirurgiões menos experientes.

A seguir se mostram as médias e os limites do intervalo de 95% de confiança da diferença.

**Tabela 46. Médias das cirurgias realizadas por cirurgiões que nasceram antes e depois de 1960 (MS)**

| Grupo de cirurgiões     | Média variável transformada | Média variável original |
|-------------------------|-----------------------------|-------------------------|
| Nasceram antes de 1960  | 0,210                       | 1:05 h                  |
| Nasceram depois de 1960 | 0,220                       | 1:12 h                  |

**Tabela 47. Resultado do intervalo de 95% de confiança da diferença de médias entre as cirurgias realizadas por cirurgiões que nasceram antes e depois de 1960 (MS)**

| Diferença de médias              | Limite Inferior | Limite superior | Contem o o? |
|----------------------------------|-----------------|-----------------|-------------|
| Antes 1960 [0] – depois 1960 [1] | -0,023          | 0,003           | Sim         |

Ao contrario da MRM, na Mastectomia simples não se detectam diferenças significativas entre as cirurgias que foram realizadas por cirurgiões que nasceram antes e depois de 1960.

Porém, se observará na análise de regressão se o ano de nascimento tratada como uma variável quantitativa é importante para explicar a variabilidade da resposta.

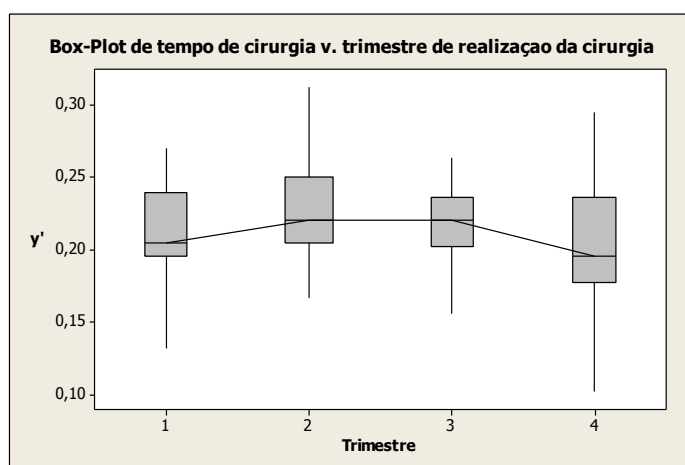
- Trimestre da cirurgia

A seguir estão listados os tamanhos amostrais de cada grupo de cirurgias, de acordo com o trimestre no qual a intervenção foi realizada.

**Tabela 48. Cirurgias realizadas nos diferentes trimestres do ano (MS)**

| Trimestre de realização da cirurgia | N  |
|-------------------------------------|----|
| Trimestre 1 [1]                     | 29 |
| Trimestre 2 [2]                     | 29 |
| Trimestre 3 [3]                     | 30 |
| Trimestre 4 [4]                     | 32 |

No Box-Plot a seguir se pode detectar a tendência das observações dependendo do seu trimestre de realização.



**Figura 39. Box-Plot de cirurgias realizadas nos 4 trimestres do ano (MS)**

Tendência de aumento é detectada entre o primeiro e o segundo trimestre, e uma tendência decrescente entre o terceiro e quarto trimestres.

**Tabela 49. Média de duração das cirurgias realizadas nos quatro trimestres do ano (MS)**

| Trimestre | Média (variável transformada) | Média (variável original) |
|-----------|-------------------------------|---------------------------|
| 1         | 0,213                         | 1:07 h                    |
| 2         | 0,224                         | 1:14 h                    |
| 3         | 0,215                         | 1:08 h                    |
| 4         | 0,207                         | 1:05 h                    |

E a tabela a seguir mostra o resultado dos intervalos de confiança simultâneos de Tukey.

**Tabela 50. Resultado dos intervalos de confiança simultâneos de Tukey (95% de confiança para a família) de diferenças das médias das durações de cirurgias realizadas em cada trimestre do ano (MS)**

| Diferença de médias       | Limite Inferior | Limite Superior | Contem o 0? |
|---------------------------|-----------------|-----------------|-------------|
| Trimestre 4 – trimestre 1 | -0,030          | 0,019           | Sim         |
| Trimestre 3 – trimestre 1 | -0,023          | 0,027           | Sim         |
| Trimestre 2 – trimestre 1 | -0,012          | 0,035           | Sim         |
| Trimestre 4 – trimestre 2 | -0,040          | 0,006           | Sim         |
| Trimestre 3 – trimestre 2 | -0,033          | 0,014           | Sim         |
| Trimestre 4 – trimestre 3 | -0,032          | 0,017           | Sim         |

Não se detecta uma diferença significativa em nenhum deles. Porém, a análise de regressão determinará se é um fator que se deve levar em conta.

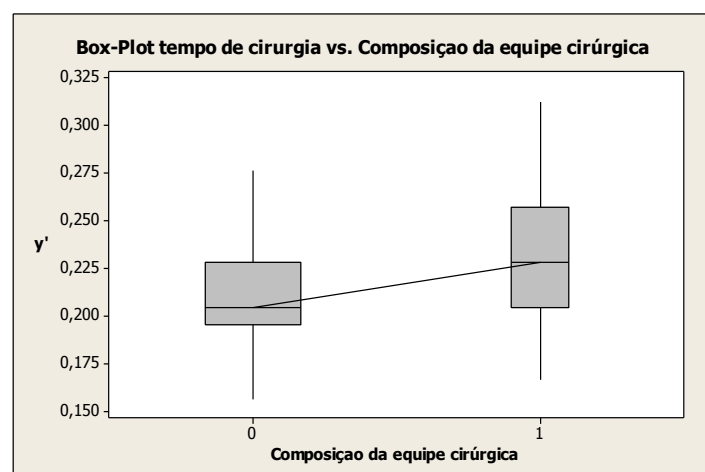
- Composição da equipe cirúrgica.

A continuação se mostram os tamanhos amostrais.

**Tabela 51. Cirurgias em que o cirurgião *staff* participa ou não na equipe cirúrgica (MS).**

| <b>Equipe cirúrgica</b>                            | <b>N</b> |
|--|----------|
| <i>Staff</i> não participa na equipe cirúrgica (0) | 95       |
| <i>Staff</i> participa na equipe cirúrgica (1)     | 35       |

E a seguir se mostra o Box-Plot.



**Figura 40. Box-Plot das cirurgias em que o cirurgião *staff* participa ou não na equipe cirúrgica (MS)**

Detecta-se uma diferença na duração das cirurgias quando o cirurgião *staff* participa da equipe cirúrgica.

A seguir se mostram as médias dos dois grupos e os limites do intervalo de 95% de confiança da diferença de médias.

**Tabela 52. Médias das durações das cirurgias em que o cirurgião staff participou ou não na equipe cirúrgica (MS)**

| Composição da equipe cirúrgica    | Média (variável transf.) | Média (variável original) |
|-----------------------------------|--------------------------|---------------------------|
| Staff participa não participa [0] | 0,209                    | 1:05 h                    |
| Staff participa [1]               | 0,232                    | 1:19 h                    |

E a seguir se mostra o resultado do intervalo de confiança de diferença de médias.

**Tabela 53. Resultado do intervalo de 95% de confiança da diferença das médias de cirurgias em que o cirurgião staff participou ou não na equipe cirúrgica (MS)**

| Diferença composição    | Limite inferior | Limite superior | Contem el 0? |
|-------------------------|-----------------|-----------------|--------------|
| equipe [0] – equipe [1] | -0,037          | -0,010          | Nao          |

As diferenças resultam significativas, resultando em um aumento de 14 minutos no caso em que o cirurgião *staff* participa na equipe cirúrgica.

- Momento do dia em que começa a cirurgia

A tabela a seguir resume os tamanhos amostrais do conjunto de observações em que a cirurgia começou com anterioridade às 12:00 h [0] e o grupo no qual a intervenção começou depois das 12:00 h [1].

**Tabela 54. Cirurgias que começaram antes e depois das 12:00 h (MS)**

| Começo da cirurgia                      | N  |
|---|----|
| Cirurgia começou antes das 12:00 h [0]  | 84 |
| Cirurgia começou depois das 12:00 h [1] | 46 |

O Box-Plot a seguir ilustra a possível diferença em relação ao fato.

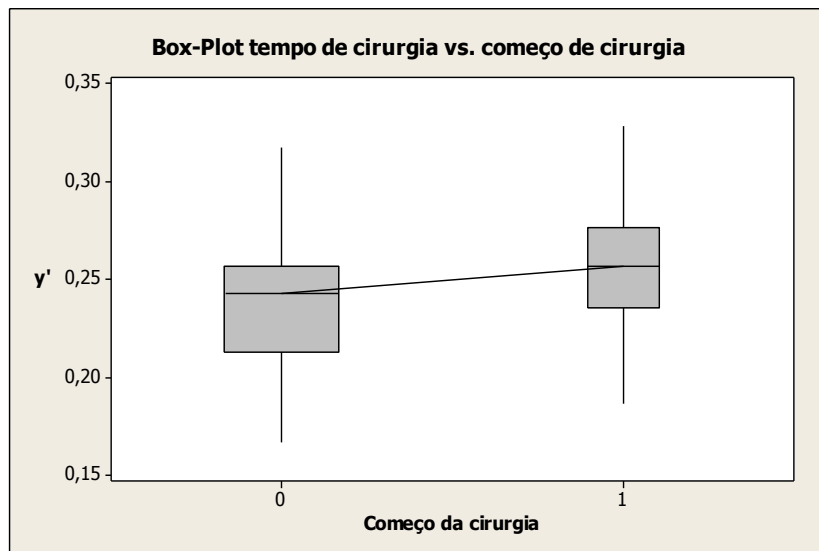


Figura 41. Box-Plot de cirurgias que começaram antes e depois das 12:00 h (MS)

Pode-se notar um aumento na duração da intervenção em aquelas cirurgias que começaram depois de meio-dia.

As médias são mostradas na tabela a seguir.

Tabela 55. Médias de duração das cirurgias que começaram antes e depois das 12:00 h (MS)

| Começo da cirurgia    | Média (variável transf.) | Média (variável original) |
|-----------------------|--------------------------|---------------------------|
| Antes das 12:00 h [0] | 0,209                    | 1:05 h                    |
| Após às 12:00 h [1]   | 0,227                    | 1:15 h                    |

E a tabela a seguir mostra os limites do intervalo de confiança de 95%.

Tabela 56. Resultado do intervalo de 95% de confiança da diferença de médias das durações de cirurgias que começaram antes e depois das 12:00 h (MS)

| Diferença começo cirurgia     | Limite inferior | Limite superior | Contem o 0? |
|-------------------------------|-----------------|-----------------|-------------|
| Começo AM [0] – Começo PM [1] | - 0,030         | - 0,005         | Não         |

De modo que pode ser um fator importante a levar em conta na equação de regressão final. Se percebe que as cirurgias que começaram após às 12:00 h demoram, em média, 10 minutos mais.



## Fatores associados com o procedimento

### - Biópsia de Linfonodo Sentinela (BLS)

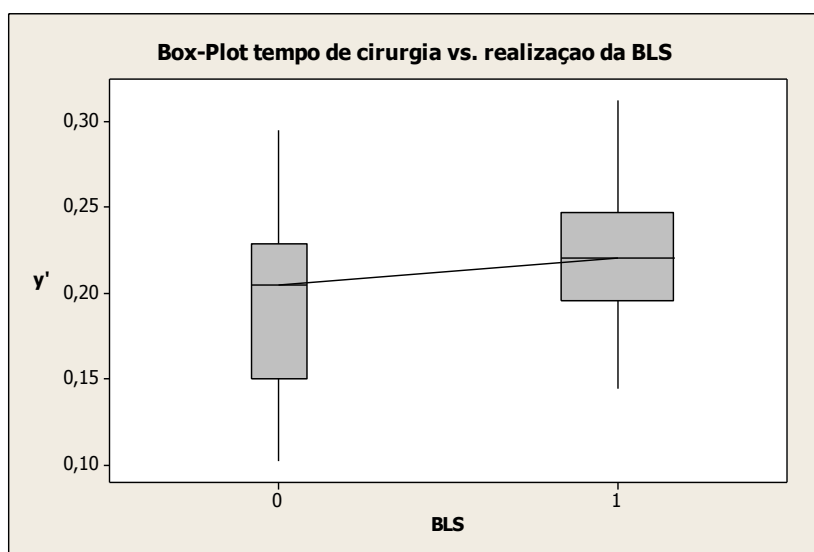
Como foi explicado na descrição dos fatores, a realização ou não da BLS pode ser um elemento que agrega tempo de duração à cirurgia..

Os tamanhos amostrais são mostrados a seguir.

**Tabela 57. Cirurgias em que se realizou ou não a Biopsia Linfonodo Sentinela (MS)**

| <b>Realização da BLS</b>  | <b>N</b> |
|---------------------------|----------|
| Não se realizou a BLS [0] | 25       |
| Realizou-se a BLS [1]     | 125      |

O seguinte Box-Plot indica se foi realizada tal atividade [1], ou pelo contrario só foi removida a mama sem biópsia posterior [0].



**Figura 42. Box-Plot com as cirurgias em que se realizou ou não a Biopsia (MS)**

Pode-se notar um aumento na duração da cirurgia no caso em que se realizou a biópsia.

As médias e o resultado do intervalo de 95% de confiança de diferença são mostrados a seguir.

**Tabela 58. Média da duração das cirurgias em que se realizou ou não a Biopsia Linfonodo Sentinela (MS)**

| Realização da BLS      | Média (variável transf.) | Média (variável original) |
|------------------------|--------------------------|---------------------------|
| No se realizou BLS [0] | 0,195                    | 0:58 h                    |
| Se realizou BLS [1]    | 0,220                    | 1:11 h                    |

**Tabela 59. Resultado do intervalo de 95% de confiança da diferença de médias da duração das cirurgias em que se realizou ou não a Biopsia Linfonodo Sentinela (MS)**

| Diferença                 | Limite inferior | Limite superior | Contem o 0? |
|---------------------------|-----------------|-----------------|-------------|
| Nao BLS [0] – Sim BLS [1] | -0,045          | -0,005          | Não         |

Conclui-se que o fato de realizar a biópsia é um fator que, em princípio, pode afetar a duração da cirurgia (a BLS adiciona 13 minutos para o tempo de operação), pois são detectadas diferenças entre as médias dos dois grupos.

### 3.3.4.2 Interações consideradas no modelo de regressão linear.

O gráfico a seguir mostra as variáveis de interação que foram criadas para a *Stepwise Regression* posterior.

**Tabela 60. Interações consideradas no modelo de Regressao Linear (MS)**

|                  | IM C | Hip.     | Diab.    | Gén.     | Ano nasc. | TRI1      | TRI2      | TRI3      | TRI4      | Começo cirurgia | Comp. equipe cirúrgica | BLS       |
|------------------|------|----------|----------|----------|-----------|-----------|-----------|-----------|-----------|-----------------|------------------------|-----------|
| <b>IMC</b>       |      | IMC xHIP | IMCx DIA | IMCxG EN | IMCxA NO  | IMCxT RI1 | IMCxT RI2 | IMCxT RI3 | IMCxT RI4 | IMCxCO OMI      | IMCxCO MP              | IMCxBL S  |
| <b>Hipert .</b>  |      |          | HIPx DIA | HIPxG EN | HIPxA NO  | HIPxT RI1 | HIPxT RI2 | HIPxT RI3 | HIPxT RI4 | HIPxCO MI       | HIPxCO MP              | HIPxBL S  |
| <b>Diab-</b>     |      |          |          | DIAxG EN | DIAxA NO  | DIAxT RI1 | DIAxT RI2 | DIAxT RI3 | DIAxT RI4 | DIAxCO MI       | DIAxCO MP              | DIAxBL S  |
| <b>Gén.</b>      |      |          |          |          | GENxA NO  | GENxT RI1 | GENxT RI2 | GENxT RI3 | GENxT RI4 | GENxCO OMI      | GENxCO MP              | GENxB LS  |
| <b>Ano nasc.</b> |      |          |          |          |           | TRI1xA NO | TRI2xA NO | TRI3xA NO | TRI4xA NO | ANOxC OMI       | ANOxC OMI              | ANOxB LS  |
| <b>TRI1</b>      |      |          |          |          |           | -         | -         | -         | -         | TRI1xC OMI      | TRI1xC OMI             | TRI1xB LS |
| <b>TRI2</b>      |      |          |          |          |           |           | -         | -         | -         | TRI2xC OMI      | TRI2xC OMI             | TRI2xB LS |
| <b>TRI3</b>      |      |          |          |          |           |           |           | -         | -         | TRI3xC OMI      | TRI3xC OMI             | TRI3xB LS |
| <b>TRI4</b>      |      |          |          |          |           |           |           |           | -         | TRI4xC OMI      | TRI4xC OMI             | TRI4xB LS |
| <b>Come</b>      |      |          |          |          |           |           |           |           |           | -               | COMixC                 | COMixB    |

|                         |  |  |  |  |  |  |  |  |  |  |     |              |
|-------------------------|--|--|--|--|--|--|--|--|--|--|-----|--------------|
| ço<br>cirur.            |  |  |  |  |  |  |  |  |  |  | OMP | LS           |
| Comp<br>Equip<br>cirúr. |  |  |  |  |  |  |  |  |  |  | -   | COMPx<br>BLS |

De modo se vai levar em conta mais 12 variáveis do que na MRM (uma variável Dummy no BLS, duas interação quantitativa da BLS com as variáveis IMC e Ano de nascimento e 9 variáveis Dummy de interação da BLS os com outros fatores de binário).

No total são 72 variáveis explicativas potenciais que entrarão no processo de seleção de variáveis. A tabela a seguir divide as variáveis explicativas de acordo com a sua natureza quantitativa, Dummy, original ou de interação.

**Tabela 61. Natureza das variáveis originais e de interação consideradas no modelo de Regressão Linear (MS)**

| Natureza da variável | Variáveis originais | Variáveis de interação |
|----------------------|---------------------|------------------------|
| Dummy                | 10                  | 39                     |
| Quantitativa         | 2                   | 21                     |

### 3.3.4.3 Análise de Regressão e interpretação dos resultados

#### Stepwise Regression

A seguir se mostra a saída de *Minitab* para a Stepwise Regression ( $\alpha_{entrada} = 0,05$ ;  $\alpha_{saida} = 0,05$ ) levando em conta as variáveis nomeadas acima.

72 variáveis, com N = 130

| Paso               | 1       | 2       | 3       | 4       | 5       |
|--------------------|---------|---------|---------|---------|---------|
| Constante          | 0,1929  | 0,1885  | 0,1860  | 0,1831  | -1,0389 |
| IMCxBLS            | 0,00098 | 0,00102 | 0,00107 | 0,00100 | 0,00097 |
| Valor T            | 4,01    | 4,33    | 4,62    | 4,34    | 4,24    |
| Valor P            | 0,000   | 0,000   | 0,000   | 0,000   | 0,000   |
| HIPxCOMP           |         | 0,0303  | 0,0291  | 0,0272  | 0,0257  |
| Valor T            |         | 3,26    | 3,20    | 3,03    | 2,88    |
| Valor P            |         | 0,001   | 0,002   | 0,003   | 0,005   |
| DIAxTRI2           |         |         | 0,041   | 0,043   | 0,040   |
| Valor T            |         |         | 2,74    | 2,90    | 2,69    |
| Valor P            |         |         | 0,007   | 0,004   | 0,008   |
| ANOxCOMI           |         |         |         | 0,00001 | 0,00001 |
| Valor T            |         |         |         | 2,16    | 2,20    |
| Valor P            |         |         |         | 0,033   | 0,030   |
| Ano Nascimento     |         |         |         |         | 0,00062 |
| Valor T            |         |         |         |         | 2,02    |
| Valor P            |         |         |         |         | 0,046   |
| S                  | 0,0351  | 0,0338  | 0,0330  | 0,0325  | 0,0321  |
| R-cuad.            | 11,17   | 18,03   | 22,64   | 25,43   | 27,80   |
| R-cuad. (ajustado) | 10,47   | 16,74   | 20,80   | 23,04   | 24,89   |

O resumo dos diferentes passos é o seguinte:

- **Passo 1:** A variável *IMCxBLS* é adicionada
- **Passo 2:** A variável *HIPxCOMP* é adicionada
- **Passo 3:** A variável *DIAxTRI2* é adicionada
- **Passo 4:** A variável *ANOxCOMI* é adicionada
- **Passo 5:** A variável *Ano Nascimento* é adicionada

E a figura a seguir mostra a análise de variância para a regressão.

Análise da variância

| Fuente         | GL  | SC       | MC       | F    | P     |
|----------------|-----|----------|----------|------|-------|
| Regresión      | 5   | 0,049268 | 0,009854 | 9,55 | 0,000 |
| Error residual | 124 | 0,127956 | 0,001032 |      |       |
| Total          | 129 | 0,177225 |          |      |       |

A qual resulta significativa.

A figura a seguir mostra os diferentes coeficientes de regressão individuais para cada fator, a estatística  $t$ , o p-valor e os fatores de inflação da variância (VIF). Além disso, se fornecem o coeficiente de determinação da regressão normal e ajustado (explicação total da variabilidade) e o coeficiente de determinação da previsão ajustado.

| Variável       | Coef       | SE Coef    | T     | P     | VIF   |
|----------------|------------|------------|-------|-------|-------|
| Constante      | -1,0389    | 0,6051     | -1,72 | 0,089 |       |
| IMCxBLS        | 0,0009660  | 0,0002280  | 4,24  | 0+    | 1,034 |
| HIPxCOMP       | 0,025664   | 0,008917   | 2,88  | 0,005 | 1,022 |
| DIAXTRI2       | 0,03991    | 0,01483    | 2,69  | 0,008 | 1,025 |
| ANOxCOMI       | 0,00000671 | 0,00000305 | 2,20  | 0,030 | 1,033 |
| Ano Nascimento | 0,0006233  | 0,0003086  | 2,02  | 0,046 | 1,025 |

S = 0,0321233    R-cuad. = 27,8%    R-cuad. (ajustado) = 24,9%

PRESS = 0,140945    R-cuad. (pred) = 20,47%

Detecta-se na Mastectomia Simples que o fato de que o paciente tenha um IMC elevado é um fator influente ao realizar a BLS (variável ativa: *IMCxBLS*).

Do mesmo modo que na MRM, a variável *HIPxCOMP* indica que quando a equipe cirúrgica é composta por cirurgiões *staff* como auxiliares, o fato da paciente ser hipertensa é um fator que agrega tempo de duração da cirurgia.

Em relação à variável *DIAXTRI2* é uma interação que pode refletir que quando se operou no segundo trimestre, as paciente diabéticas foram submetidas a cirurgias que demoraram mais.

A interação *ANOxCOMI* indica que quando as cirurgias começam depois das 12:00 h tem maior demora nos casos em que os cirurgiões são menos expertos.

E, finalmente, de um modo geral, o a variável *Ano de Nascimento* com coeficiente de regressão positiva indica que quando os cirurgiões *staff* possuem menor idade, as cirurgias tendem a demorar mais.

Pode-se verificar que todos os Fatores Inflação da Variância são próximos a um, de modo que não há variáveis colineares.

Finalmente, é preciso dizer que se pode explicar 24,9% da variabilidade da resposta e produz um coeficiente de determinação da predição de 20,47%.

Por tanto, a equação de regressão final resulta,

$$Y' = - 1,04 + 0,000966 * \text{IMCxBLS} + 0,0257 * \text{HIPxCOMP} + 0,0399 * \text{DIAxTRI2} + 0,000007 * \text{ANOxCOMI} + 0,000623 * \text{Ano Nascimento}$$

E, por último, a seguir se apresentam as gráficas de resíduos.

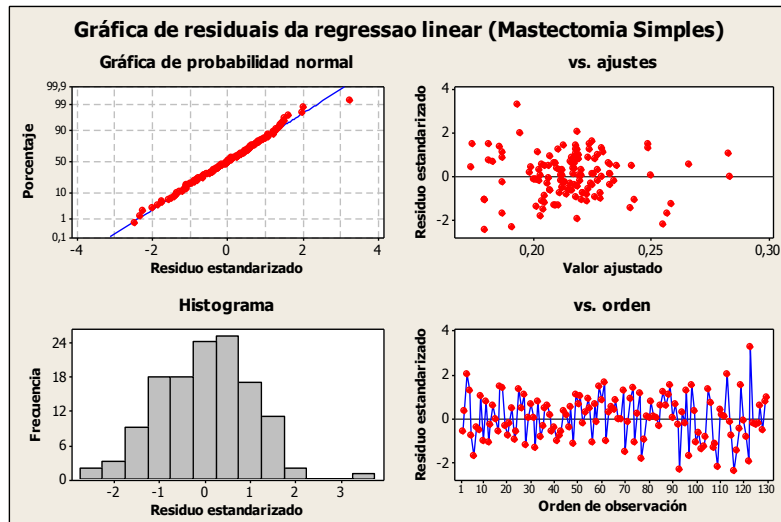


Figura 43. Gráficos de resíduos da Regressão Linear (MS)

Onde não se percebem desvios sensíveis de normalidade nem tendências identificáveis na gráfica de resíduos vs. valores ajustados, pelo que o modelo será considerado válido.

### 3.3.5 Segmentectomia (SEG)

A seguir se mostra um gráfico de probabilidade normal e um histograma dos dados de duração de cirurgia das diferentes observações de Segmentectomia.

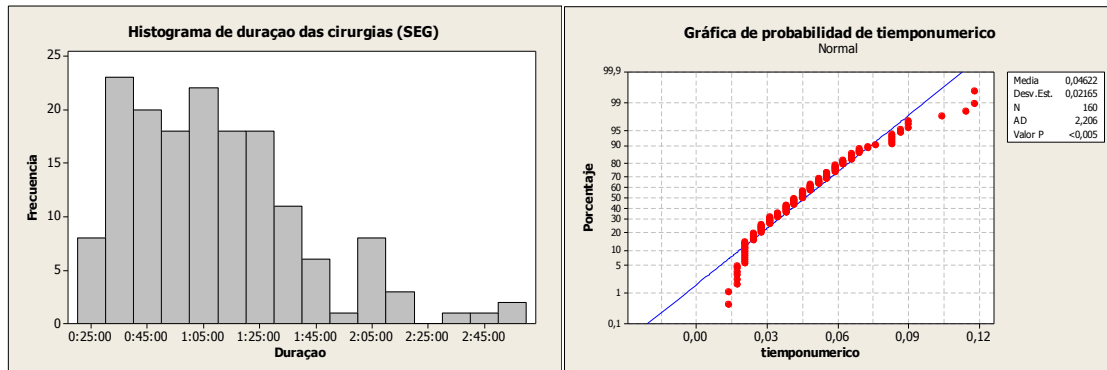


Figura 44. Histograma e gráfica de probabilidade normal dos dados de duração das cirurgias (SEG)

Pode ser visto que a assimetria à direita não permite tratar os dados originais como se fossem normais. Observamos, também, que o valor-p inferior a 0,05 do teste de AD de normalidade rejeita a hipótese nula de que os dados provêm de uma distribuição Normal.

O *Minitab* neste ponto detecta como transformação ótima a transformação logarítmica ( $\lambda = 0$ ). Porém, com o objetivo de se podam realizar comparações entre os três procedimentos (e, portanto, a magnitude da variável transformada deveria ser a mesma) se adotará a transformação de raiz quadrada. Se na análise final dos resíduos os pressupostos da regressão linear são verificados, se terão homogeneizado as transformações das variáveis originais o que facilitara as comparações.

A continuação se pode observar um gráfico de probabilidade normal acumulada da variável transformada de acordo a raiz quadrada.

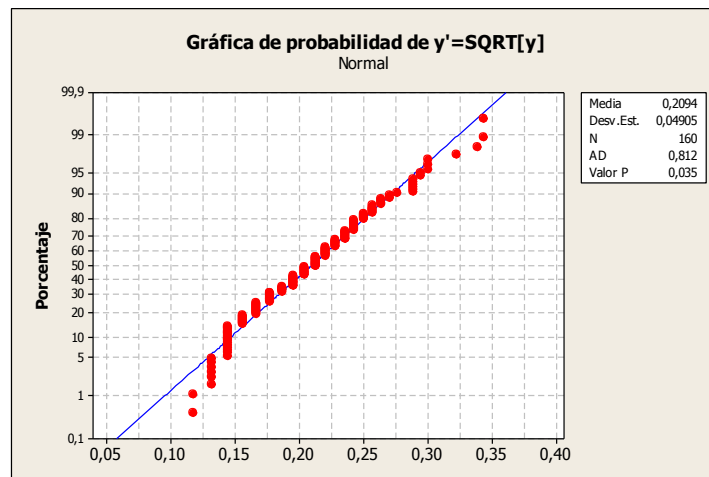


Figura 45. . Gráfica de Probabilidade normal da variável transformada do tempo de cirurgia (SEG)

Pode-se observar que, mesmo que a suposição de normalidade não se verificaria se fixássemos o p-valor mínimo em 0,05 para rejeitar a hipótese nula, o valor real do p-valor se situa relativamente próximo ao nível desejável.

### 3.3.5.1 Análise exploratória dos dados

#### Fatores associados ao paciente

- IMC

A tabela a seguir lista as observações que pertencem a cada um dos intervalos do IMC.

Tabela 62. Quantidade de cirurgias por intervalo de IMC (SEG)

| Intervalo IMC | N  |
|---------------|----|
| Baixo Peso    | 3  |
| Eutrofia      | 26 |
| Sobrepeso     | 69 |
| Obesidade     | 62 |

A seguir apresenta-se um Box-Plot, já com os dados transformados, em que se podem ver diferenças nas durações de acordo com se o doente pertence a um intervalo nutricional ou outro.



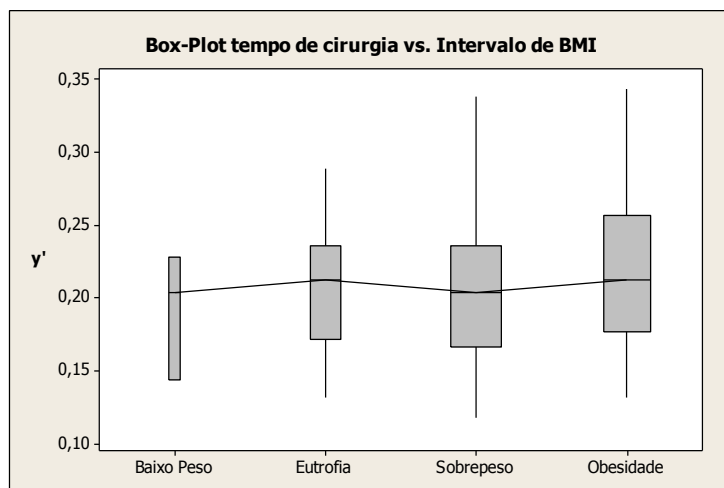


Figura 46. Box-Plot de duração da cirurgia vs. Intervalo de IMC (SEG)

Não se percebem diferenças extremamente significativas nem uma tendência clara.

As tabelas a seguir mostram a média (da variável transformada e a variável original) da duração de cada cirurgia dependendo do intervalo do IMC, e os resultados do teste de Tukey para os intervalos de confiança simultâneos (95% para a família)

Tabela 63. Média de duração de cirurgia para cada intervalo do IMC (SEG)

| Intervalo   | Média (variable normalizada) | Média (variable original) |
|-------------|------------------------------|---------------------------|
| Baixo Peso* | 0,218                        | 0:55 h                    |
| Eutrofia    | 0,205                        | 1:03 h                    |
| Sobrepeso   | 0,204                        | 1:03 h                    |
| Obesidade   | 0,192                        | 1:13 h                    |

Tabela 64. Intervalos de confiança simultâneos de Tukey (95% para a família) de diferença de médias dos intervalos de IMC (SEG)

| Diferença de médias    | Limite Inferior | Limite Superior | Contem o 0? |
|------------------------|-----------------|-----------------|-------------|
| Obesidade – Baixo Peso | -0,065          | 0,101           | Sim         |
| Sobrepeso – Baixo Peso | -0,063          | 0,086           | Sim         |
| Eutrofia – Baixo Peso  | -0,065          | 0,090           | Sim         |
| Obesidade – Eutrofia   | -0,016          | 0,043           | Sim         |
| Sobrepeso – Eutrofia   | -0,030          | 0,028           | Sim         |
| Obesidade – Sobrepeso  | -0,007          | 0,037           | Sim         |

Não se detecta, ao contrario dos procedimentos anteriores, nenhuma diferença significativa.

Será observado no apartado de Regressão Linear se o *IMC*, na sua versão original ou em interação com outra variável, é uma variável explicativa importante para a duração das cirurgias.

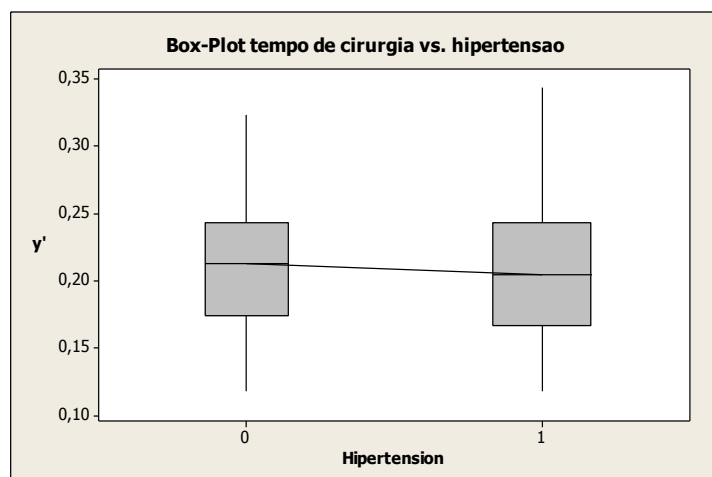
### - Hipertensão

A tabela a seguir mostra o número de observações que pertencem a cada um dos dois grupos.

**Tabela 65. Pacientes hipertensas e não hipertensas (SEG)**

| Grupo de pacientes | Nº de observaciones |
|--------------------|---------------------|
| No hipertensas [0] | 66                  |
| Hipertensas [1]    | 94                  |

E a seguir encontra-se o Box-Plot (da variável transformada), que representa os casos em que a paciente é ou não é hipertensa [1,0].



**Figura 47. Box-Plot de duração das cirurgias vs. hipertensão (SEG)**

Observa-se uma leve diferença entre os dois grupos. Porém, não parece ser significativa.

A seguir se mostram as médias nos dois casos.

**Tabela 66. Média das durações de cirurgia de pacientes hipertensas e não hipertensas (SEG)**

| Hipertensao                 | Média (variável transformada) | Média (variável original) |
|-----------------------------|-------------------------------|---------------------------|
| Paciente nao hipertensa [0] | 0,211                         | 1:07 h                    |
| Paciente hipertensa [1]     | 0,208                         | 1:06 h                    |

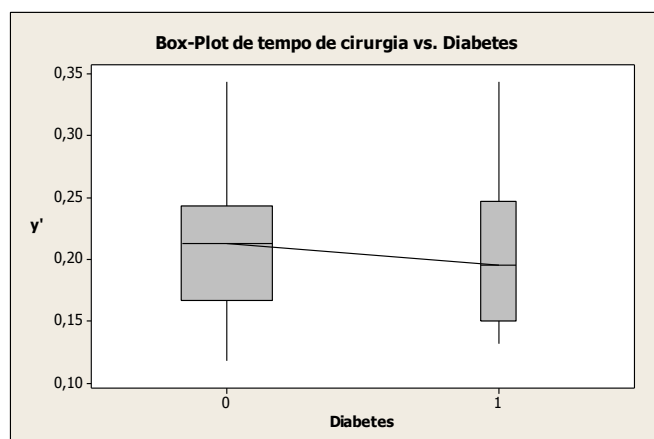
Não será preciso indicar o resultado do teste *t* de diferença de médias, pois já se observa que a diferença não vai ser estatisticamente significativa.

- Diabetes.

A seguir se apresentam os tamanhos amostrais e o gráfico Box-Plot, que mostra a comparação entre o grupo de pacientes com diabetes [1] e o grupo de pacientes não-diabéticos [0].

**Tabela 67. Pacientes diabéticas e não diabéticas (SEG)**

| Grupo de pacientes | N   |
|--------------------|-----|
| Nao Diabéticas [0] | 139 |
| Diabéticas [1]     | 21  |



**Figura 48. Box-Plot pacientes diabéticas e não diabéticas (SEG).**

Pode-se observar como existe uma diferença na linha de conexão da mediana, mas pode ser principalmente devido à assimetria da distribuição dos dois grupos.

Na tabela abaixo podemos ver as médias de cada grupo:

**Tabela 68. Médias de duração das pacientes diabéticas e não diabéticas (SEG)**

| Diabetes                   | Média (variável transformada) | Média (variável original) |
|----------------------------|-------------------------------|---------------------------|
| Paciente nao diabética [0] | 0,209                         | 1:06 h                    |
| Paciente diabética [1]     | 0,208                         | 1:07 h                    |

O fato de a paciente ser ou não diabética não parece, portanto, um fator que seja significativo, uma vez que as médias dos dois grupos são iguais. Não se precisa incluir o resultado do teste  $t$  de diferença, embora a variável seja levada em conta na análise de regressão posterior.

### **Fatores associados ao cirurgião**

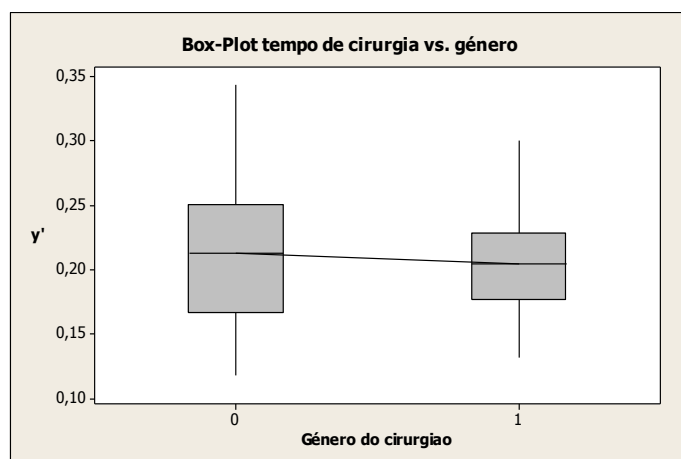
- **Gênero do cirurgião.**

A seguir se mostra a tabela com os tamanhos amostrais do grupo de cirurgias praticadas por homens e por mulheres.

**Tabela 69. Cirurgias realizadas por homens e por mulheres (SEG)**

| <b>Gênero do cirurgião</b> | <b>N</b> |
|----------------------------|----------|
| Homens [0]                 | 110      |
| Mulheres [1]               | 50       |

No Box-Plot a seguir pode-se observar a diferença na duração das cirurgias dos dois grupos.



**Figura 49. Box-Plot cirurgias realizadas por homens e por mulheres (SEG)**

Observa-se diferença entre os dois grupos (as mulheres operam mais rápido do que os homens) e uma menor variância no caso em que as cirurgias foram feitas por mulheres.

A tabela a seguir mostra as médias.

**Tabela 70. Média das durações de cirurgias realizadas por homens e mulheres (SEG)**

| Gênero do cirurgião responsável | Média (variável transformada) | Média (variável original) |
|---------------------------------|-------------------------------|---------------------------|
| Homem [0]                       | 0,212                         | 1:08 h                    |
| Mulher [1]                      | 0,204                         | 1:01 h                    |

E a seguir se mostram os limites do intervalo do 95% de confiança da diferença das médias dos dois grupos.

**Tabela 71. Resultado do intervalo de 95% de confiança da diferença das médias de duração de cirurgias realizadas por cirurgiões mulheres e homens (SEG)**

| Diferença gênero          | Limite inferior | Limite superior | Contem o 0? |
|---------------------------|-----------------|-----------------|-------------|
| Homens [0] – Mulheres [1] | -0,006          | 0,023           | Sim         |

Nesse caso não se detectam diferenças significativas, entre as diferenças de gênero na duração das cirurgias.

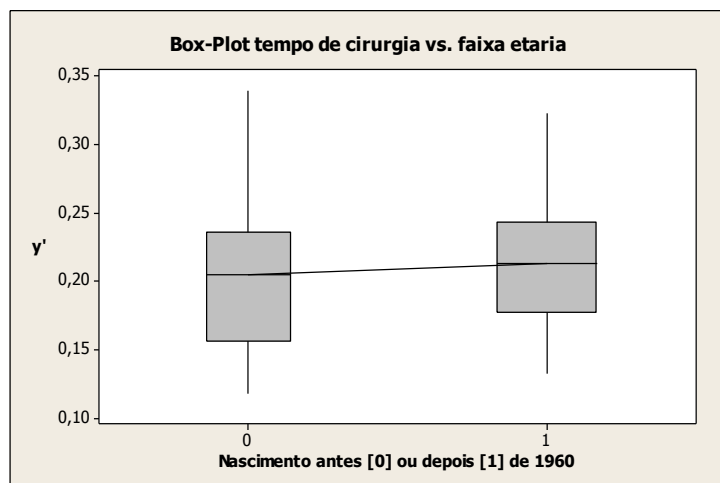
#### - Ano de nascimento do cirurgião

A tabela a seguir mostra os tamanhos amostrais do grupo de cirurgiões que nasceram antes e depois de 1960.

**Tabela 72. Cirurgias praticadas por cirurgiões que nasceram antes e depois de 1960 (SEG)**

| <b>Grupo de cirurgiões</b>  | <b>N</b> |
|-----------------------------|----------|
| Nasceram antes de 1960 [0]  | 67       |
| Nasceram depois de 1960 [1] | 93       |

O Box-Plot a seguir mostra a tendência dos dois grupos.



**Figura 50. Cirurgias realizadas por cirurgiões que nasceram antes e depois de 1960 (SEG)**

Parece existir uma duração menor da cirurgia quando é praticada por cirurgiões com mais experiência (e que nasceram antes de 1960) do que aqueles com menor idade (e que nasceram depois de 1960).

A seguir se mostram as médias e os limites do intervalo de 95% de confiança da diferença.

**Tabela 73. Médias das cirurgias realizadas por cirurgiões que nasceram antes e depois de 1960 (SEG)**

| Faixa etária cirurgião  | Média variável transformada | Média variável original |
|-------------------------|-----------------------------|-------------------------|
| Nasceram antes do 1960  | 0,200                       | 1:01 h                  |
| Nasceram depois do 1960 | 0,216                       | 1:10 h                  |

**Tabela 74. Resultado do intervalo de 95% de confiança da diferença de médias entre as cirurgias realizadas por cirurgiões que nasceram antes e depois de 1960 (SEG)**

| Diferença de médias              | Limite Inferior | Limite superior | Contem o o? |
|----------------------------------|-----------------|-----------------|-------------|
| Antes 1960 [0] – depois 1960 [1] | -0,031          | 0 <sup>+</sup>  | Sim         |

Sendo rigorosos, não se detectam diferenças significativas entre as duas faixas etárias (definidas através da data de nascimento). Porém se observa que o intervalo se encontra quase totalmente no lado esquerdo, o que indica que tem indícios que indicam certa significância que pode ser detectada na análise de regressão.

**- Trimestre da cirurgia**

A seguir estão listados os tamanhos amostrais de cada grupo de cirurgias, de acordo com o trimestre no qual a intervenção foi realizada.

Tabela 75. Cirurgias realizadas nos diferentes trimestres do ano (SEG)

| Trimestre de realização da cirurgia | N  |
|-------------------------------------|----|
| Trimestre 1 [1]                     | 47 |
| Trimestre 2 [2]                     | 47 |
| Trimestre 3 [3]                     | 36 |
| Trimestre 4 [4]                     | 30 |

No Box-Plot a seguir se pode observar a tendência das observações dependendo do seu trimestre de realização.

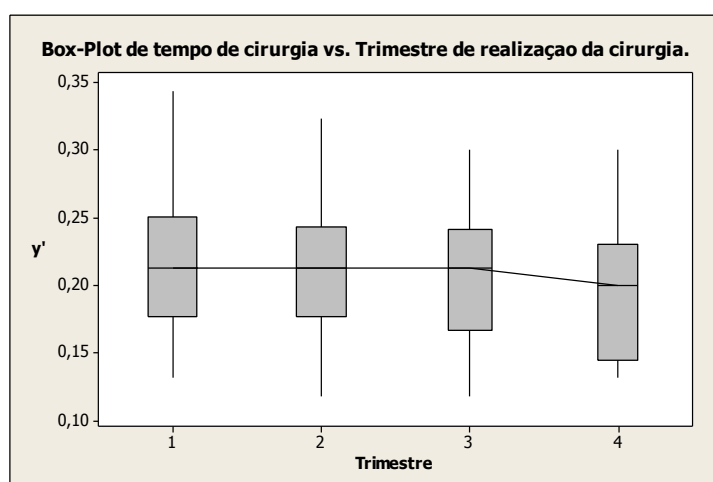


Figura 51. Box-Plot de cirurgias realizadas nos 4 trimestres do ano (SEG)

Percebe-se um descenso na duração das cirurgias no trimestre 4.

E as tabelas a seguir mostram, respectivamente, as médias de cada trimestre e o resultado dos intervalos de confiança simultâneos de Tukey.

Tabela 76. Média de duração das cirurgias realizadas nos quatro trimestres do ano (SEG)

| Trimestre | Média (variável transformada) | Média (variável original) |
|-----------|-------------------------------|---------------------------|
| 1         | 0,214                         | 1:09 h                    |
| 2         | 0,215                         | 1:10 h                    |
| 3         | 0,105                         | 1:03 h                    |
| 4         | 0,199                         | 1:00 h                    |

**Tabela 77. Resultado dos intervalos de confiança simultâneos de Tukey (95% de confiança para a família) de diferenças das médias das durações de cirurgias realizadas em cada trimestre do ano (SEG)**

| Diferença de médias       | Limite Inferior | Limite Superior | Contem o 0? |
|---------------------------|-----------------|-----------------|-------------|
| Trimestre 4 – trimestre 1 | -0,045          | 0,015           | Sim         |
| Trimestre 3 – trimestre 1 | -0,037          | 0,020           | Sim         |
| Trimestre 2 – trimestre 1 | -0,026          | 0,027           | Sim         |
| Trimestre 4 – trimestre 2 | -0,045          | 0,014           | Sim         |
| Trimestre 3 – trimestre 2 | -0,037          | 0,019           | Sim         |
| Trimestre 4 – trimestre 3 | -0,038          | 0,025           | Sim         |

Não se detecta uma diferença significativa entre as médias de duração das intervenções realizadas nos diferentes trimestres do ano.

A análise de regressão determinará se é um fator que se deve levar em conta.

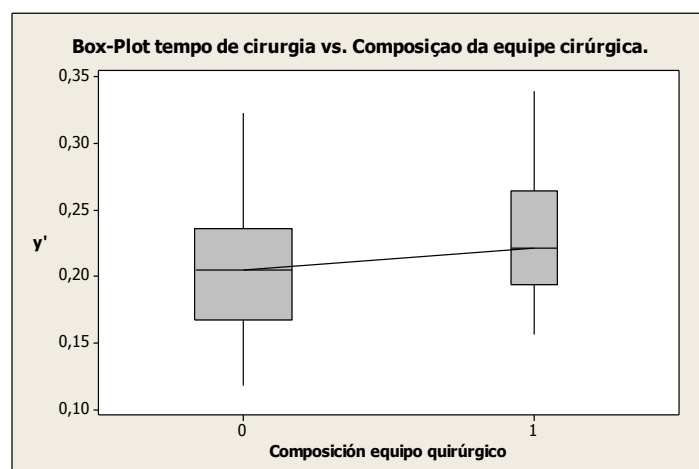
- Composição da equipe cirúrgica.

A continuação se mostram os tamanhos amostrais.

**Tabela 78. Cirurgias em que o cirurgião *staff* participa ou não na equipe cirúrgica (SEG)**

| <b>Equipe cirúrgica</b>                            | <b>N</b> |
|--|----------|
| <i>Staff</i> não participa na equipe cirúrgica (0) | 130      |
| <i>Staff</i> participa na equipe cirúrgica (1)     | 30       |

A seguir se mostra o Box-Plot



**Figura 52. Box-Plot das cirurgias em que o cirurgião *staff* participa ou não na equipe cirúrgica (SEG)**



Detecta-se uma diferença na duração das cirurgias quando o cirurgião *staff* participa da equipe cirúrgica (demoram mais).

A seguir se mostram as médias dos dois grupos e os limites do intervalo de 95% de confiança da diferença de médias.

**Tabela 79. Médias das durações das cirurgias em que o cirurgião *staff* participou ou não na equipe cirúrgica (SEG)**

| Composição da equipe cirúrgica           | Média (variável transf.) | Média (variável original) |
|--|--------------------------|---------------------------|
| <i>Staff</i> participa não participa [0] | 0,205                    | 1:04 h                    |
| <i>Staff</i> participa [1]               | 0,227                    | 1:18 h                    |

E a seguir se mostra o resultado do intervalo de confiança de diferença de médias.

**Tabela 80. Resultado do intervalo de 95% de confiança da diferença das médias de cirurgias em que o cirurgião *staff* participou ou não na equipe cirúrgica (SEG)**

| Diferença composição    | Límite inferior | Límite superior | Contiene el 0? |
|-------------------------|-----------------|-----------------|----------------|
| equipe [0] – equipe [1] | -0,022          | -0,003          | Nao            |

De modo que o fator pode ser de interesse para explicar as demoras das cirurgias.

- Momento do dia em que começa a cirurgia

A tabela a seguir resume os tamanhos amostrais do conjunto de observações em que a cirurgia começou com anterioridade às 12:00 h [0] e o grupo no qual a intervenção começou depois das 12:00 h [1].

**Tabela 81. Cirurgias que começaram antes e depois das 12:00 h (SEG)**

| Começo da cirurgia                      | N  |
|---|----|
| Cirurgia começou antes das 12:00 h [0]  | 64 |
| Cirurgia começou depois das 12:00 h [1] | 96 |

O Box-Plot a seguir ilustra a possível diferença em relação ao fato.

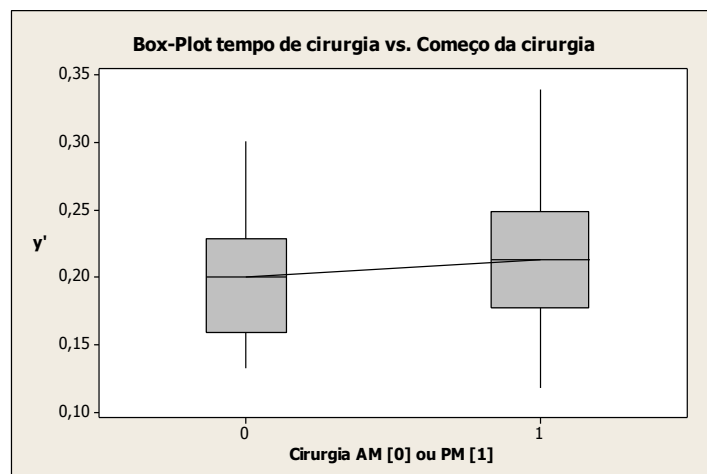


Figura 53. Box-Plot de cirurgias que começaram antes e depois das 12:00 h (SEG)

Detecta-se um aumento na duração das cirurgias no caso em que começaram depois das 12:00h

As médias são mostradas na tabela a seguir.

Tabela 82. Médias de duração das cirurgias que começaram antes e depois das 12:00 h (SEG)

| Começo da cirurgia         | Média (variável transf.) | Média (variável original) |
|----------------------------|--------------------------|---------------------------|
| Antes de las 12:00 h [0]   | 0,202                    | 1:02 h                    |
| Después de las 12:00 h [1] | 0,214                    | 1:09 h                    |

E a tabela a seguir mostra os limites do intervalo de confiança de 95%.

Tabela 83. Resultado do intervalo de 95% de confiança da diferença de médias das durações de cirurgias que começaram antes e depois das 12:00 h (SEG)

| Diferença começo cirurgia      | Limite inferior | Limite superior | Contem o 0? |
|--------------------------------|-----------------|-----------------|-------------|
| Começoo AM [0] – Começo PM [1] | -0,028          | 0,003           | Sim         |

Ao contrario da MS, na Segmentectomia não se percebem diferenças estatisticamente significativas entre os casos das cirurgias que começaram antes ou depois de meio-dia, mesmo se a tendência do Box-Plot e o cálculo de médias revelam menor duração em aquelas que começaram antes.

### **- Fatores associados ao procedimento**

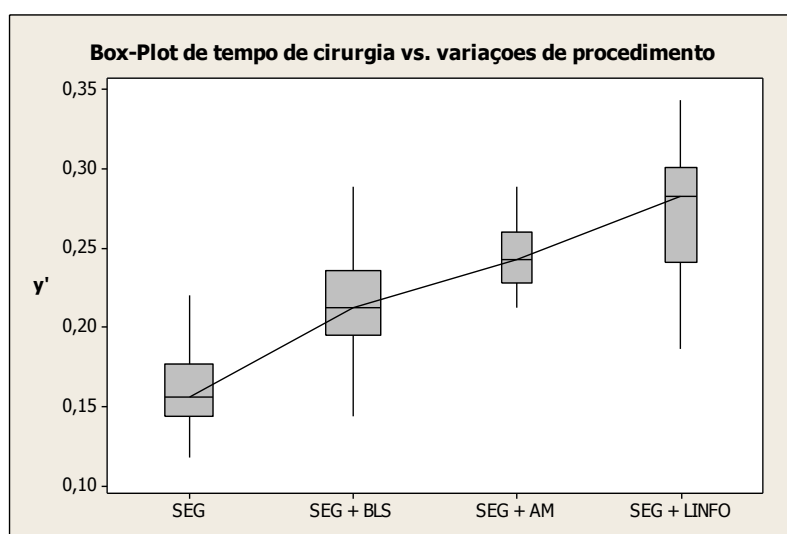
A seguir se analisam as diferentes variações do procedimento associado com a Segmentectomia.

Na tabela a seguir se apresentam os tamanhos amostrais de cada variação.

**Tabela 84. Quantidade de cirurgias realizadas de acordo com a variação do procedimento (SEG)**

| Variações do procedimento | N  |
|---------------------------|----|
| SEG                       | 52 |
| SEG+BLS                   | 65 |
| SEG+AM                    | 21 |
| SEG+LINFO                 | 22 |

E o seguinte Box-Plot mostra as diferenças de demora de cirurgia.



**Figura 54. Box-Plot de cirurgias realizadas de acordo à variação do procedimento (SEG)**

Percebe-se claramente que é o fator mais importante. Também se detecta maior variabilidade nos casos em que se realizou a Linfodectomia Axilar do que nos outros 3.

A continuação se apresentam as médias e o resultado dos intervalos de confiança simultâneos de Tukey.

**Tabela 85. Médias de duração das cirurgias de acordo a cada variação do procedimento (SEG)**

| Variações no procedimento | Média (variável transf.) | Média (variável original) |
|---------------------------|--------------------------|---------------------------|
| SEG                       | 0,164                    | 0:40 h                    |
| SEG+BLS                   | 0,213                    | 1:07 h                    |
| SEG+AM                    | 0,245                    | 1:27 h                    |
| SEG+LINFO                 | 0,273                    | 1:50 h                    |

**Tabela 86. Resultado dos intervalos de confiança simultâneos de Tukey (95% para a família) das diferenças de médias de duração da cirurgia de acordo as variações do procedimento (SEG)**

| Diferença de médias     | Limite Inferior | Limite Superior | Contem o 0? |
|-------------------------|-----------------|-----------------|-------------|
| [SEG+LINFO] – SEG       | 0,088           | 0,130           | Não         |
| [SEG+AM] – SEG          | 0,060           | 0,102           | Não         |
| [SEG + BLS] – SEG       | 0,034           | 0,064           | Não         |
| [SEG+LINFO] – [SEG+BLS] | 0,040           | 0,081           | Não         |
| [SEG+AM] – [SEG+BLS]    | 0,011           | 0,053           | Não         |
| [SEG+LINFO] – [SEG+AM]  | 0,003           | 0,028           | Não         |

É claro, portanto, que o fato de realizar a Biopsia, uma ampliação de Margem ou a Linfodectomia axilar são fatores muito a levar em conta na análise de regressão.

### 3.3.5.2 Interações consideradas no modelo de regressão linear.

O gráfico a seguir mostra as variáveis de interação que foram criadas para a *Stepwise Regression* posterior.

**Tabela 87. Interações consideradas no modelo de Regressão Linear (SEG)**

|         | Variáveis originais                                  | SEG      | SEG+BLS  | SEG+AM  | SEG+LINFO  |
|---------|--|----------|----------|---------|------------|
| IMC     | Interações já mostradas nos procedimentos anteriores | SEGxIMC  | BLSxIMC  | AMxIMC  | LINFOxIMC  |
| HIP.    |  | SEGxHIP  | BLSxHIP  | AMxHIP  | LINFOxHIP  |
| DIA     |  | SEGxDIA  | BLSxDIA  | AMxDIA  | LINFOxDIA  |
| GEN     |  | SEGxGEN  | BLSxGEN  | AMxGEN  | LINFOxGEN  |
| ANONASC |  | ANOxSEG  | ANOxBLS  | ANOxAM  | ANOxLINFO  |
| TRI1    |  | SEGxTRI1 | BLSxTRI1 | AMxTRI1 | LINFOxTRI1 |
| TRI2    |  | SEGxTRI2 | BLSxTRI2 | AMxTRI2 | LINFOxTRI2 |
| TRI3    |  | SEGxTRI3 | BLSxTRI3 | AMxTRI3 | LINFOxTRI3 |
| TRI4    |  | SEGxTRI4 | BLSxTRI4 | AMxTRI4 | LINFOxTRI4 |
| COMI    |  | SEGxCOMI | BLSxCOMI | AMxCOMI | LINFOxCOMI |
| COMP    |  | SEGxCOMP | BLSxCOMP | AMxCOMP | LINFOxCOMP |

Em total resultam 105 variáveis explicativas, potenciais que entrarão no processo de seleção de variáveis. A tabela a seguir divide as variáveis explicativas de acordo com a sua natureza quantitativa, Dummy, original ou de interação.

**Tabela 88. Natureza das variáveis originais e de interação consideradas no modelo de Regressão Linear (SEG)**

| Natureza da variável | Variáveis originais | Variáveis de interação |
|----------------------|---------------------|------------------------|
| Dummy                | 13                  | 65                     |
| Cuantitativa         | 2                   | 25                     |

### 3.3.5.3 Análise de Regressão e interpretação dos resultados

#### Stepwise Regression

A seguir se mostra a saída de *Minitab* para a Stepwise Regression ( $\alpha_{entrada} = 0,05$ ;  $\alpha_{saida} = 0,05$ ) levando em conta as variáveis nomeadas acima.

105 variáveis, com N = 160

| Paso               | 1        | 2        | 3        | 4        | 5        | 6        |
|--------------------|----------|----------|----------|----------|----------|----------|
| Constante          | 0,2302   | 0,2193   | 0,2113   | 0,2117   | 0,2060   | -0,8749  |
| IMCxSEG            | -0,00220 | -0,00184 | -0,00157 | -0,00158 | -0,00139 | -0,00139 |
| Valor T            | -10,07   | -9,21    | -7,85    | -8,33    | -7,18    | -7,25    |
| Valor P            | 0,000    | 0,000    | 0,000    | 0,000    | 0,000    | 0,000    |
| IMCxLINFO          |          | 0,00170  | 0,00194  | 0,00252  | 0,00269  | 0,00266  |
| Valor T            |          | 6,88     | 8,04     | 9,41     | 10,14    | 10,15    |
| Valor P            |          | 0,000    | 0,000    | 0,000    | 0,000    | 0,000    |
| IMCxAM             |          |          | 0,00119  | 0,00118  | 0,00137  | 0,00141  |
| Valor T            |          |          | 4,26     | 4,43     | 5,18     | 5,39     |
| Valor P            |          |          | 0,000    | 0,000    | 0,000    | 0,000    |
| GENxLINFO          |          |          |          | -0,055   | -0,055   | -0,056   |
| Valor T            |          |          |          | -4,20    | -4,28    | -4,39    |
| Valor P            |          |          |          | 0,000    | 0,000    | 0,000    |
| COMPxBLS           |          |          |          |          | 0,0295   | 0,0298   |
| Valor T            |          |          |          |          | 3,23     | 3,30     |
| Valor P            |          |          |          |          | 0,001    | 0,001    |
| Ano Nascimento     |          |          |          |          |          | 0,00055  |
| Valor T            |          |          |          |          |          | 2,25     |
| Valor P            |          |          |          |          |          | 0,026    |
| S                  | 0,0384   | 0,0338   | 0,0321   | 0,0305   | 0,0296   | 0,0292   |
| R-cuad.            | 39,11    | 53,21    | 58,08    | 62,36    | 64,75    | 65,88    |
| R-cuad. (ajustado) | 38,73    | 52,61    | 57,28    | 61,39    | 63,61    | 64,54    |
| Paso               | 7        |          |          |          |          |          |
| Constante          | -1,039   |          |          |          |          |          |
| IMCxSEG            | -0,00149 |          |          |          |          |          |
| Valor T            | -7,70    |          |          |          |          |          |
| Valor P            | 0,000    |          |          |          |          |          |
| IMCxLINFO          | 0,00259  |          |          |          |          |          |
| Valor T            | 9,95     |          |          |          |          |          |
| Valor P            | 0,000    |          |          |          |          |          |
| IMCxAM             | 0,00137  |          |          |          |          |          |
| Valor T            | 5,30     |          |          |          |          |          |
| Valor P            | 0,000    |          |          |          |          |          |
| GENxLINFO          | -0,059   |          |          |          |          |          |
| Valor T            | -4,72    |          |          |          |          |          |

|                    |         |
|--------------------|---------|
| Valor P            | 0,000   |
| COMPxBLS           | 0,0281  |
| Valor T            | 3,15    |
| Valor P            | 0,002   |
| Ano Nascimento     | 0,00063 |
| Valor T            | 2,60    |
| Valor P            | 0,010   |
| IMCxTRI1           | 0,00043 |
| Valor T            | 2,37    |
| Valor P            | 0,019   |
| S                  | 0,0288  |
| R-cuad.            | 67,10   |
| R-cuad. (ajustado) | 65,58   |

O resumo dos diferentes passos é o seguinte:

- **Passo 1:** A variável *IMCxSEG* é adicionada
- **Passo 2:** A variável *IMCxLINFO* é adicionada
- **Passo 3:** A variável *IMCxAM* é adicionada
- **Passo 4:** A variável *GENxLINFO* é adicionada
- **Passo 5:** A variável *COMPxBLS* é adicionada
- **Passo 6:** A variável *Ano Nascimento* é adicionada
- **Passo 7:** A variável *IMCxTRI1* é adicionada

E a figura a seguir mostra a análise de variância para a regressão.

Análise da variância

| Fuente         | GL  | SC       | MC       | F     | P     |
|----------------|-----|----------|----------|-------|-------|
| Regresión      | 7   | 0,256663 | 0,036666 | 44,28 | 0,000 |
| Error residual | 152 | 0,125855 | 0,000828 |       |       |
| Total          | 159 | 0,382518 |          |       |       |

A qual resulta significativa.

A figura a seguir mostra os diferentes coeficientes de regressão individuais para cada fator, a estatística *t*, o p-valor e os fatores de inflação da variância (VIF). Além disso, se fornecem o coeficiente de determinação da regressão normal e ajustado (explicação total da variabilidade) e o coeficiente de determinação da previsão ajustado.

| Variável       | Coef       | SE Coef   | T     | P     | VIF   |
|----------------|------------|-----------|-------|-------|-------|
| Constante      | -1,0391    | 0,4784    | -2,17 | 0,031 |       |
| IMCxSEG        | -0,0014864 | 0,0001929 | -7,70 | 0+    | 1,388 |
| IMCxLINFO      | 0,0025909  | 0,0002603 | 9,95  | 0+    | 1,644 |
| IMCxAM         | 0,0013708  | 0,0002585 | 5,30  | 0+    | 1,211 |
| GENxLINFO      | -0,05937   | 0,01257   | -4,72 | 0+    | 1,451 |
| COMPxBLS       | 0,028068   | 0,008913  | 3,15  | 0,002 | 1,146 |
| Ano Nascimento | 0,0006330  | 0,0002435 | 2,60  | 0,010 | 1,035 |
| IMCxTRI1       | 0,0004251  | 0,0001793 | 2,37  | 0,019 | 1,105 |

S = 0,0287749 R-cuad. = 67,1% R-cuad. (ajustado) = 65,6%

PRESS = 0,141523 R-cuad. (pred) = 63,00%

Detecta-se na Segmentectomia que todos os fatores associados às variações do procedimento foram selecionados, e em interação com o IMC. Portanto, o fato de fazer uma linfodectomia axilar, uma Biopsia Linfonodo Sentinela ou uma ampliação de margens, implica que quando o paciente possui um IMC mais elevado, a cirurgia demora mais.

Novamente, neste caso o género do cirurgião é um fator ativo. Na Segmentectomia em concreto é uma variável significativa quando interaciona com a linfodectomia axilar.

A variável *COMPxBLS* indica que quando o cirurgião *staff* participa na equipe cirúrgica e se realiza a Biopsia Linfonodo Sentinela, a cirurgia possui maior duração.

O ano de nascimento é também um fator a levar em conta. Quando o cirurgião tem mais idade, a cirurgia demora menos.

A significância da variável *IMCxTRI1* indica que o IMC do paciente influencia a duração da cirurgia no primeiro trimestre do ano (demorando mais a intervenção). Isto pode indicar que, o fato de que os cirurgiões residentes começam a operar a inícios de ano, a sua falta de experiência na cirurgia leva a uma maior necessidade de tempo para realizar a Segmentectomia quando o paciente possui um IMC elevado.

A equação de regressão final resulta,

$$Y' = - 1,04 - 0,00149 * \text{IMCxSEG} + 0,00259 * \text{IMCxLINFO} + 0,00137 * \text{IMCxAM} - 0,0594 * \text{GENxLINFO} + 0,0281 * \text{COMPxBLS} + 0,000633 * \text{Ano Nascimento} + 0,000425 * \text{IMCxTRI1}$$

A seguir se mostram as gráficas de resíduos nas quais se podem verificar os pressupostos de normalidade e homocedasticidade.

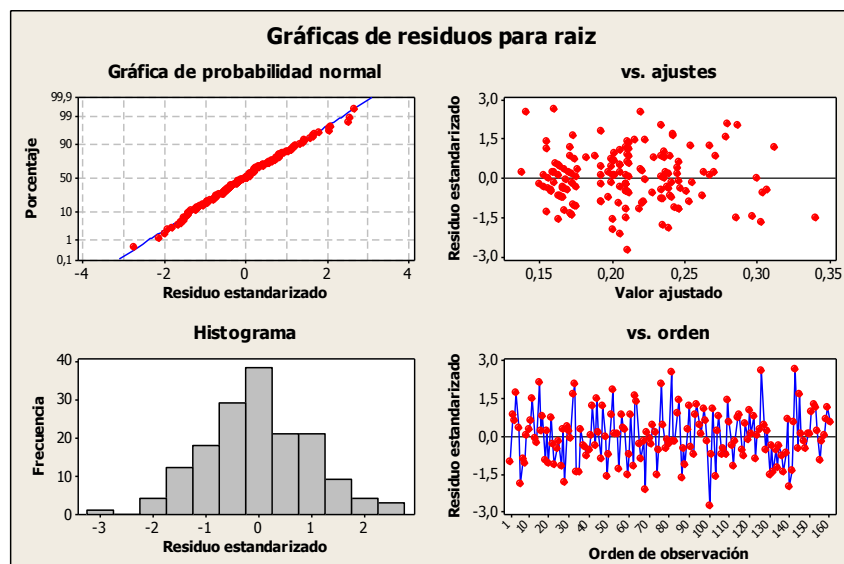


Figura 55. Gráficos de resíduos da Regressão Linear (SEG)

### 3.4 Verificação do modelo ajustado

Para avaliar os resultados da regressão foram analisadas as intervenções que se realizaram durante a semana de 3 de junho a 7 de junho para poder fazer uma comparação entre a duração prevista para a cirurgia via regressão linear, e a duração real das intervenções.

A seguir as equações de regressão estimadas para cada procedimento e os coeficientes de determinação ajustado de cada uma delas são apresentados.

Tabela 89. Lembrete das equações de regressão estimadas para cada procedimento

| Proced. | Equação de regressão  | R <sup>2</sup> -ajustado |
|---------|---|--------------------------|
| MRM     | $Y' = - 0,909 + 0,000001 * IMC \times ANO + 0,000008 * COMI \times ANO - 0,0184 * Trimestre\ 4 - 0,000520 * IMC \times GEN + 0,000566 * ANONASC + 0,0173 * HIP \times COMP$ | 21,6%                    |
| MS      | $Y' = - 1,04 + 0,000966 * IMC \times BLS + 0,0257 * HIP \times COMP + 0,0399 * DIA \times TRI2 + 0,000007 * ANO \times COMI + 0,000623 * ANONASC$                           | 24,9%                    |
| SEG     | $Y' = - 1,04 - 0,00149 * IMC \times SEG + 0,00259 * IMC \times LINFO +$   | 65,6%                    |



|  |  |  |
|--|--|--|
|  | $+ 0,00137 * IMC \times AM - 0,0594 * GEN \times LINFO +$<br>$+ 0,0281 * COMP \times BLS + 0,000633 * ANONASC +$<br>$+ 0,000425 IMC \times TRI1$ |  |
|--|--|--|

Pode-se ver que há uma disparidade significativa em termos da quantidade de variabilidade explicada entre cada equação de regressão (Segmentectomia com respeito aos outros dois procedimentos), o que será discutido mais adiante na seção de conclusões.

### Intervenções realizadas

O gráfico a seguir mostra os grupos de intervenções durante a semana de referência, de acordo foram cirurgias de tipo mastológico, plástico ou de ambas.

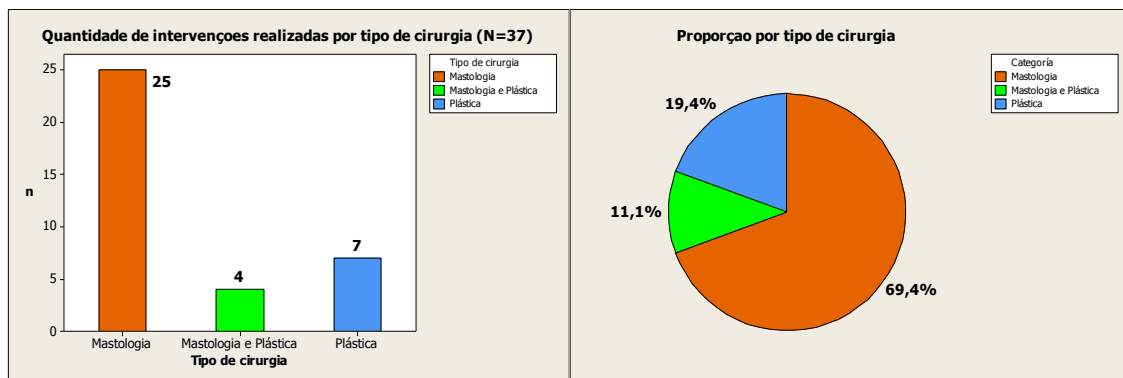


Figura 56. Cirurgias realizadas por área cirúrgica do hospital

Pode ser visto como o maior número de cirurgias pertence à área de mastologia, sendo a proporção ainda maior se considerar também as intervenções que incluem uma equipa de cirurgia mastológica e outra equipa de cirurgia plástica (operações que incluem alguma forma de reconstrução imediata após a remoção do tumor).

O gráfico a seguir mostra o número de cirurgias realizadas na área de mastologia durante o período estudado.

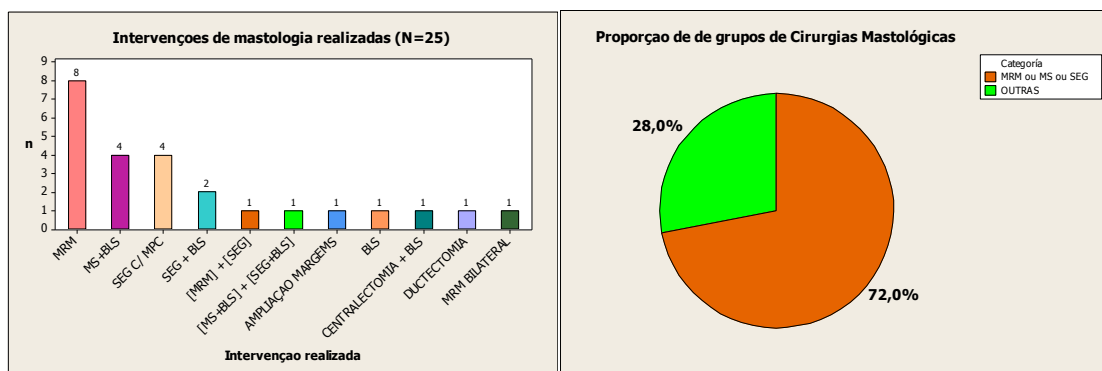


Figura 57. Cirurgias realizadas na área de mastologia

Note-se que os três procedimentos estudados neste projeto (MRM, MS e SEG) representaram uma alta porcentagem do total de procedimentos realizados durante a semana de junho 03-07, da área de mastologia.

### Previsão de observações através da equação de regressão linear

As três tabelas a seguir (em ordem, MRM, MS e SEG) contêm o resultado da aplicação da equação de regressão e a duração real da cirurgia.

Tabela 90. Resultado de aplicar a equação de regressão linear às intervenções MRM

| Cirurgia | Estimado<br>$[\hat{Y}']$ | Estimado em tempo<br>$[\hat{Y} = (\hat{Y}')^2]$ | Duração real | Erro (e)<br>$[Real - Estimado]$ |
|----------|--------------------------|---|--------------|---------------------------------|
| 1        | 0,259                    | 1:37 h  | 1:20 h       | - 0:17 h                        |
| 2        | 0,252                    | 1:32 h  | 1:50 h       | - 0:18 h                        |
| 3        | 0,237                    | 1:21 h  | 1:20 h       | - 0:01 h                        |
| 4        | 0,261                    | 1:38 h  | 1:15 h       | - 0:23 h                        |
| 5        | 0,236                    | 1:20 h  | 1:15 h       | - 0:05 h                        |
| 6        | 0,246                    | 1:27 h  | 1:45 h       | 0:18 h                          |
| 7        | 0,265                    | 1:41 h  | 1:50 h       | 0:09 h                          |
| 8        | 0,246                    | 1:27 h  | 1:45 h       | 0:18 h                          |

Tabela 91. Resultado de aplicar a equação de regressão linear às intervenções MS

| Cirurgia | Estimado<br>[ $\hat{Y}'$ ] | Estimado em tempo<br>[ $\hat{Y} = (\hat{Y}')^2$ ] | Duração real | Erro (e)<br>[Real – Estimado] |
|----------|----------------------------|---|--------------|-------------------------------|
| 1        | 0,281                      | 1:53 h  | 1:00         | - 0:53 h                      |
| 2        | 0,234                      | 1:19 h  | 1:00         | - 0:18 h                      |
| 3        | 0,228                      | 1:15 h  | 1:25         | - 0:10 h                      |
| 4        | 0,254                      | 1:33 h  | 1:10         | - 0:23 h                      |

Tabela 92. Resultado de aplicar a equação de regressão linear às intervenções SEG

| Cirurgia | Estimado<br>[ $\hat{Y}'$ ] | Estimado em tempo<br>[ $\hat{Y} = (\hat{Y}')^2$ ] | Duração real | Erro (e)<br>[Real – Estimado] |
|----------|----------------------------|---|--------------|-------------------------------|
| 1        | 0,201                      | 0:58 h  | 0:25 h       | - 0:33 h                      |
| 2        | 0,172                      | 0:42 h  | 0:30 h       | - 0:12 h                      |
| 3        | 0,191                      | 0:53 h  | 1:05 h       | 0:12 h                        |
| 4        | 0,161                      | 0:37 h  | 0:30 h       | - 0:07 h                      |
| 5        | 0,156                      | 0:35 h  | 0:40 h       | 0:05 h                        |
| 6        | 0,211                      | 1:04 h  | 1:30 h       | 0:26 h                        |

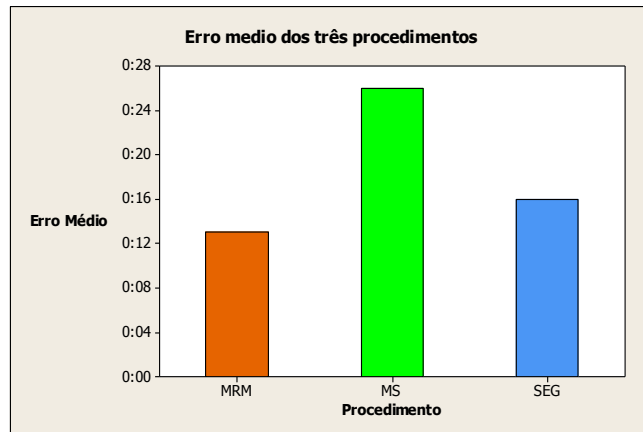
Pode-se definir o erro médio da predição como,

$$\text{erro medio de predição} = \frac{1}{n} \sum_{i=1}^{n \text{ cirurgias}} |e_i|$$

E a tabela a seguir resume os erros médios em cada um dos três procedimentos.

Tabela 93. Erro médio de estimação da previsão para os três procedimentos.

| Procedimento | Erro Médio |
|--------------|------------|
| MRM          | 0:13 h     |
| MS           | 0:26 h     |
| SEG          | 0:16 h     |



**Figura 58. Comparação visual dos erros de estimação**

Observa-se que o procedimento do que foi obtido um nível de previsão pior foi a Mastectomia Simples sendo ainda duas vezes maior que no caso de a Mastectomia Radical Modificada.

Porém, as poucas observações obtidas (no caso da Mastectomia Simples foram 4) não permitem o teste ser conclusivo.

## PARTE 4: CONCLUSÕES E DISCUSSÃO

---

Da análise exploratória de dados, os testes de hipóteses e as variáveis selecionadas na análise de regressão foram detectados fatores que afetam a duração da cirurgia.

Dentro do grupo de fatores associados com o procedimento, o IMC surgiu como o fator mais influente na demora das intervenções. Mesmo na sua forma original como em interação com outras variáveis, percebe-se que nos três procedimentos se produz um aumento na duração da cirurgia, quando o paciente tem um grau mais elevado de gordura corporal. Em segundo lugar, a hipertensão aparece como uma variável a ser considerada quando a composição da equipe cirúrgica inclui o cirurgião *staff* na MRM e na MS. Ela sugere que, quando o cirurgião responsável pela operação é um cirurgião menos experiente, a cirurgia transcorre de forma diferente (e demora mais) do que quando o cirurgião responsável pela intervenção é um médico com um maior número de cirurgias anteriores realizadas na sua vida laboral. No que diz respeito à diabetes, em nenhum dos três procedimentos foi detectado como um fator significativo a levar em conta na análise de demora das intervenções.

Com respeito aos fatores associados ao cirurgião, a sua idade é, entre as variáveis estudadas, aquela com maior peso na duração da cirurgia (nos três procedimentos foi detectada como uma variável significativa). Em todos os procedimentos cirúrgicos se repetiu o mesmo fenômeno: os cirurgiões *staff* com maior idade, em geral, realizam cirurgias que demoram menos. A composição da equipe cirúrgica tem sido também um fator que afeta à duração da intervenção. Nos casos da MS e a MRM se mostrou significativa na regressão linear quando interagindo com a hipertensão, o que acaba de ser discutido. Em terceiro lugar, quando as cirurgias começam depois das 12:00 h também é percebido um aumento da duração da intervenção nos casos em que o cirurgião *staff* é mais jovem. Uma interpretação possível é que a carga mental que envolve a realização da cirurgia afeta mais os cirurgiões que são menos experientes (em comparação com aqueles com muitos anos de experiência) na hora de realizar várias intervenções consecutivas em uma jornada cirúrgica. Em quarto lugar, o gênero do cirurgião responsável foi outro fator detectado como influente. As intervenções em que o cirurgião é uma mulher, em média, demoraram menos. Finalmente, o trimestre de realização da cirurgia parece ser o fator menos importante. Unicamente na MRM foi detectado que quando a cirurgia é realizada no último trimestre tem a duração mais

curta. Como observado na seção explicando os fatores, pode ser devido a que os cirurgiões residentes tornaram-se mais experientes depois de fazer certo número de cirurgias ao longo do ano.

Com relação aos fatores associados ao procedimento, todos os fatores estudados foram significativos tanto na análise exploratória dos dados, como nos testes estatísticos e a análise de regressão. Quando se realiza uma Biópsia Linfonodo Sentinela em Mastectomia Simples e Segmentectomia, a duração da cirurgia aumenta. Em segundo lugar, ao realizar uma expansão da margem cirúrgica também se detecta um acréscimo significativo no tempo cirúrgico. Por último, nos casos em que a BLS foi positiva e procedeu-se à linfadenectomia axilar, o tempo de duração da intervenção foi consideravelmente maior.

Com relação aos fatores estudados, as variações no procedimento produzem maiores desvios em relação à demora das intervenções e, em segundo lugar, deve enfatizar-se a influência do IMC e da idade do cirurgião.

Um dos objetivos iniciais básicos do projeto era estabelecer uma equação de regressão linear que, através da combinação dos fatores discutidos acima, permitisse a previsão de novas observações com maior precisão.

No entanto, apesar de ter identificado vários fatores que influenciam a variabilidade na duração das cirurgias, uma das conclusões, em termos de variabilidade explicada e previsão, é que a precisão da estimativa de novas durações é relativamente baixa. Na MRM e na MS o coeficiente de determinação da regressão não excede a 25%. Na Segmentectomia, onde o  $R^2$  ajustado é de 64%, mas as variáveis que permitiram uma explicação da variabilidade tão elevada da variável resposta são não programáveis. Ou seja, se pode explicar a variabilidade, mas não se pode prever com alta precisão. Especificamente, o objetivo de uma Biópsia Linfonodo Sentinela é determinar se é necessário proceder ao esvaziamento axilar, de modo que um eventual programador não pode saber de antemão o resultado da biópsia e, portanto, não pode estabelecer uma duração tentativa do procedimento considerando a realização ou não da linfadenectomia (o que agrega cerca de 25 minutos à intervenção).

Em relação ao resultado do teste de avaliação das cirurgias também resulta insatisfatório no que diz respeito à previsão de cirurgias analisadas de 3 a 7 de Junho. Em particular, a duração real das cirurgias que consistiram em uma Mastectomia Simples foram as que estiveram mais afastadas da estimativa da equação de regressão, sendo o erro de estimação até maior do que a variabilidade da cirurgia.

Quanto à Segmentectomia (procedimento dos três que foram analisados com maior variabilidade associada), a precisão das estimativas é menor do que a variabilidade associada com a cirurgia (o que significa que se possui maior precisão na hora de estimar os tempos cirúrgicos). Porém, para avaliar a equação de regressão foi utilizado o resultado das biopsias *a posteriori*, com o objetivo de analisar a precisão da equação de regressão em função do tempo cirúrgico. No entanto, como discutido acima, um programador não pode conhecer o resultado da BLS, ou se na Segmentectomia se produzirá uma ampliação das margens. Em relação à Mastectomia Radical Modificada, a variabilidade associada com a previsão das observações é a menor de todos eles (o erro de predição é o menor). No entanto, deve notar-se que é dos três procedimentos o que contém menor de variabilidade (não é realizada BLS toda vez que o procedimento já inclui o esvaziamento axilar, após a remoção do tumor).

Em relação a trabalhos futuros, propostas de fatores a estudar que poderiam melhorar a percentagem de variabilidade explicada são a composição da equipe cirúrgica e o tipo de anestesia realizada. Em relação ao primeiro, no presente projeto tem-se refletido apenas uma variável qualitativa que enumera os casos em que no prontuário se inclui ao cirurgião *staff* na equipe cirúrgica. Porém, deve notar-se que uma cirurgia é um trabalho em equipe, e que o cirurgião *staff* é responsável pelo bom desenvolvimento da intervenção (o que implica óbvias cargas de responsabilidade para a saúde atual e futura de pacientes com câncer) e, ao mesmo momento, professor dos cirurgiões residentes em treinamento envolvidos no procedimento. Portanto, muito provavelmente, existem diferentes métodos de ensino utilizados pelos médicos responsáveis pela cirurgia, bem como diferentes níveis e formas de participação ativa na intervenção por parte dos cirurgiões residentes. Definir e analisar fatores que recolham esta informação pode resultar em um melhor entendimento das causas que afetam a variabilidade da duração das intervenções considerando que a cirurgia é um trabalho manual realizado por várias pessoas ao mesmo tempo e, assim, sujeito à variabilidade das decisões tomadas e executadas por indivíduos diferentes. No que diz respeito ao tipo de anestesia, foi observado que existem muitos casos em que é realizada por indução. Este fator não foi levado em conta, e poderia afetar o tempo cirúrgico.

Outro trabalho futuro que se poderia realizar é analisar a dependência das variáveis explicativas com um Modelo Linear Generalizado. Eles não possuem a necessidade de realizar transformações da variável resposta e se eliminaria a possibilidade de obter resultados incongruentes como intervalos de confiança negativos.

Um fenômeno que tem sido detectado na análise dos registros médicos é que às vezes existem discrepâncias entre os tempos de operação que figuram no Relatório da Cirurgia e a Folha de Enfermagem. Pode ser o caso que os critérios de entrada dos dados nos prontuários sejam diferentes dependendo de quem os insere. Isto afetaria às análises de tempos (qualquer uma). Em consequência, o que se propõe é estudar a validade e os critérios para inserção de informações do pessoal médico e de enfermagem, para uma melhoria na padronização do procedimento de entrada de dados nos relatórios médicos.

### **A Programação de cirurgias e a Variabilidade: Possíveis Indicadores em simulação**

Como concluído em pontos anteriores, os procedimentos cirúrgicos possuem inevitavelmente certo grau de variabilidade associada, de modo que uma melhoria presente da programação das salas deve, na opinião do autor, incorporar às análises alguns indicadores que permitam quantificar as consequências de não prever com precisão o tempo de operação. Particularmente **em um estudo de simulação das salas de cirurgia**, algumas magnitudes identificadas que poderiam ser utilizadas como medidas da qualidade da programação são descritas a seguir:

#### **a) Taxa de utilização da sala de cirurgia**

Tal como descrito no marco teórico, é um indicador da quantidade de tempo útil (tempo de preparação + tempo de cirurgia + setup) entre o tempo de sala disponível. O maior benefício de atingir taxas de utilização elevadas é um aumento do número de pacientes operados em um horizonte dado.

#### **b) Risco de atraso das cirurgias**

Este é um indicador geral da adequação da programação proposta. Uma definição possível é mostrada a seguir,

$$\text{risco de atraso da cirurgia} = \frac{\sum_{i=1}^n t_i}{n}$$



Com  $n$  o número de corridas realizadas para cada combinação de cirurgias, o seu sequenciamento e as durações tentativas; e  $t_i$  é o tempo excedido (em horas, minutos, etc.) da cirurgia acima do planejado em cada corrida tal que,

$$t_i = \begin{cases} t_{simulado} - t_{planejado}, & t_{simulado} > t_{planejado} \\ 0, & t_{simulado} \leq t_{planejado} \end{cases}$$

Na simulação de cenários, um risco elevado de atraso levaria à rejeição da programação simulada por inadequada à duração real das cirurgias.

**c) Risco de cancelamento após demora previa de cirurgia**

Para ilustrar a diferença entre o risco de atraso e o risco de cancelamento, se pode partir das duas figuras a seguir, as quais mostram dois exemplos básicos de simulação (com 3 corridas) de duas cirurgias respeito à mesma programação.

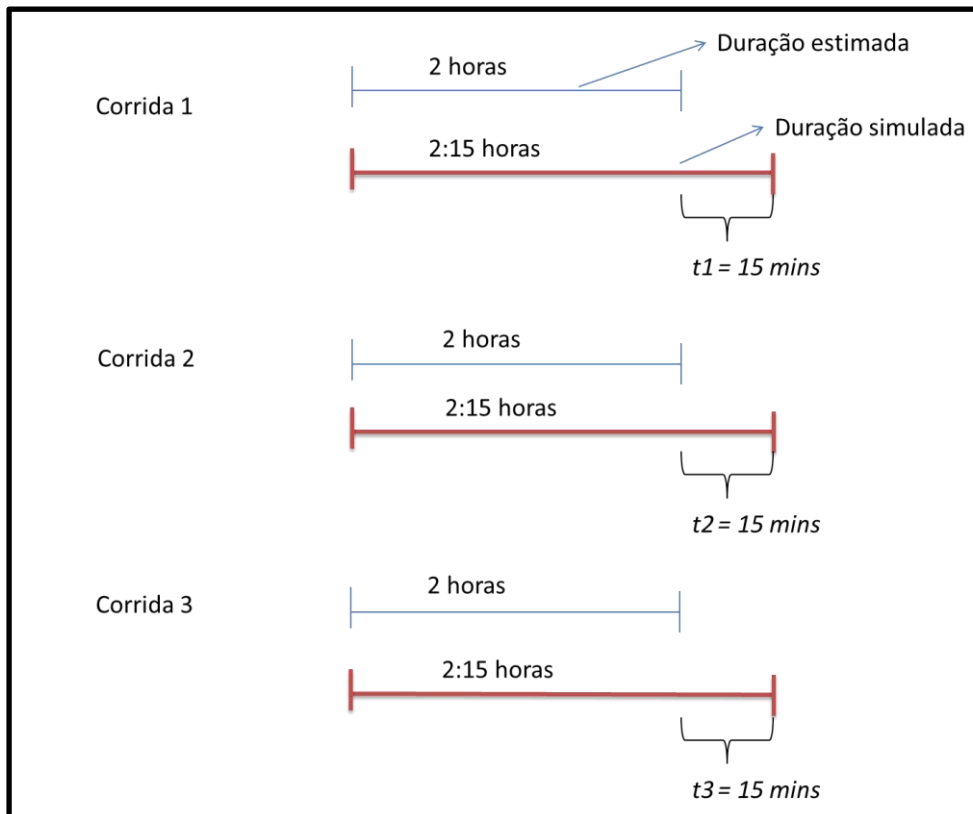


Figura 59. Simulação da duração da cirurgia 1.

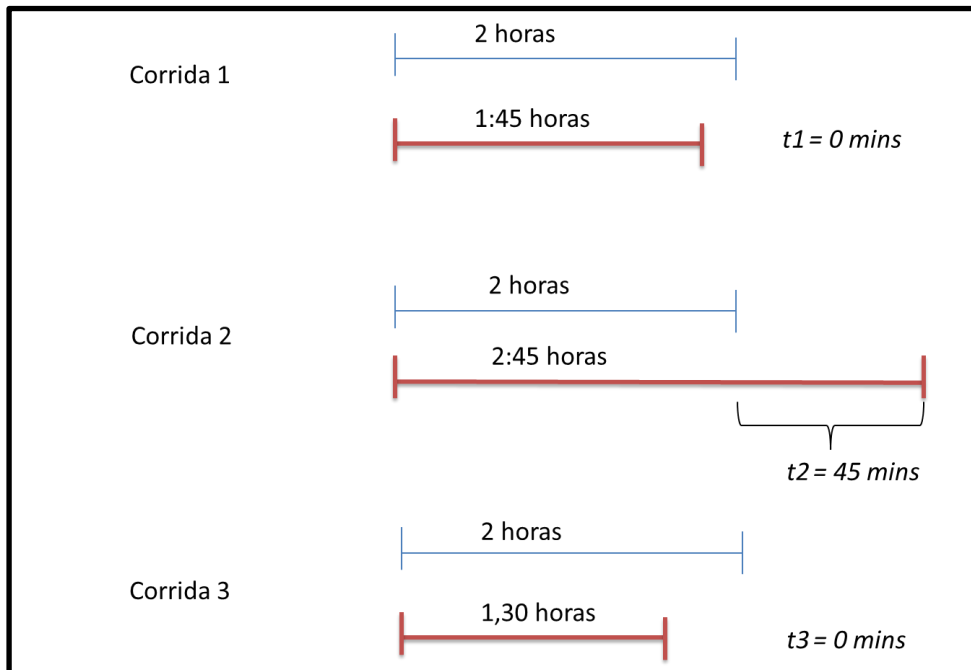


Figura 60. Simulação da duração da cirurgia 2.

O risco de atraso, em ambos os casos, tal como descrito no ponto anterior, seria,

$$\text{risco de atraso da primeira cirurgia} = \frac{15 + 15 + 15}{3} = 15 \text{ minutos}$$

$$\text{risco de atraso da segunda cirurgia} = \frac{0 + 45 + 0}{3} = 15 \text{ minutos}$$

Na ausência de uma definição mais concreta do que seria considerado um atraso aceitável, um risco de 15 minutos parece um valor admissível de risco que se estaria disposto a assumir na programação das operações. No entanto, na primeira das cirurgias simuladas, a distribuição do risco varia consideravelmente em relação à segunda. No primeiro caso, em nenhuma das corridas pode ser assumido que a demora da cirurgia derivasse em um atraso dramático do começo da seguinte intervenção. Em contraste, na segunda das corridas da segunda cirurgia simulada houve um atraso de 45 minutos em comparação com a duração planejada. Este cenário pode ser extremamente prejudicial para as ulteriores cirurgias, mesmo resultando em um possível cancelamento. Considerando o fato de que em cada jornada cirúrgica em uma sala de cirurgia se realizam entre 3 e 5 intervenções, um cancelamento supõe uma diminuição drástica na taxa de utilização da sala cirúrgica. Assim, para avaliar o risco de cancelamento, uma possível solução seria definir, **de acordo com os cirurgiões**, o valor máximo  $t_0$  de atraso a partir do qual toda demora

$t_j > t_o$  pode ser considerada como um caso de possível anulação da intervenção posterior. De modo que o risco de cancelamento se poderia definir como,

$$\text{risco de cancelamento da cirurgia [j + 1]} = \frac{\sum_{j=1}^n t_j}{n^{\circ} \text{ de atrasos}}$$

Com  $t_j$  uma função tal que,

$$t_j = \begin{cases} t_{\text{simulado}} - t_{\text{planif.}}, & t_{\text{simulado}} - t_{\text{planif.}} > t_o \\ 0, & t_{\text{simulado}} - t_{\text{planif.}} \leq t_o \end{cases}$$

# PARTE 5: BIBLIOGRAFIA

---

[1] HEALTH CARE FINANCIAL MANAGEMENT ASSOCIATION. "Achieving Operating Room Efficiency Through Process Integration". Technical Report. 2005.

[2] MOLINA J.M., FRAMINAN J.M., "Testing planning policies for solving the Elective Case Scheduling phase: A real application". *Operational Research Applied to Health Services- ORAHS 2009. International Conference on 2009*.

[3] VAN OOSTRUM, J.M., 2009, *Applying Mathematical models to surgical patient planning*, Ph.D. Dissertation, Erasmus Universiteit, Rotterdam, Netherlands, 2009.

[4] BELIEN, J., ***Exact and Heuristic Methodologies for Scheduling in Hospitals: Problems, Formulations and Algorithms***, Katholieke Universiteit Leuven, Bélgica, 2005.

[5] MARQUES PROENÇA, I. *Planeamento de cirurgias electivas. Abordagens em Programação Inteira*, Ph.D. Dissertação, Universidade de Lisboa, Lisboa, Portugal, 2011.

[6] DENTON, B.T., RAHMAN, A.S., NELSON, H., BAILEY, A. C. "Simulation of a multiple Operating Room Surgical Suite". In: *Proceedings of the 2006 Winter Simulation Conference*. Monterey, California, U.S.A., 3-6 Decembro. 2006.

[7] MILEWSKY, F. *Operating room utilization and perioperative process flow*. Disponível em: [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&ved=0CDMQFjAA&url=https%3A%2F%2Fwww.iienet2.org%2Fuploadedfiles%2FSHS\\_Community%2FPremier%2520Hospital-Operating%2520Room%2520Utilization%2520and%2520Perioperative%2520Process%2520Flow.pdf&ei=wyPOUf3yFOSP0QHms4GYAQ&usq=AFQjCNFIHm4IZ6CTJiJgBDpL26yUYLOKSQ&bvm=bv.48572450,d.dmQ](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&ved=0CDMQFjAA&url=https%3A%2F%2Fwww.iienet2.org%2Fuploadedfiles%2FSHS_Community%2FPremier%2520Hospital-Operating%2520Room%2520Utilization%2520and%2520Perioperative%2520Process%2520Flow.pdf&ei=wyPOUf3yFOSP0QHms4GYAQ&usq=AFQjCNFIHm4IZ6CTJiJgBDpL26yUYLOKSQ&bvm=bv.48572450,d.dmQ) Acceso em: 30 Fev, 2013. 8:30.

[8] WEINBROUM, A.A., EKSTEIN, P., EZRI, T. "Efficiency of the operating room suite", ***The American Journal of Surgery***, v 185 pp 244-250. 2003.

[9] BEJARANO, M., "Evaluación cuantitativa de la eficiencia en las salas de cirugía". ***Revista Colombiana de Cirugía*** v.26 nº 4, Dec. 2011.

[10] MULA, J., POLER, R., J.P., GARCÍA-SABATER, F.C. "Models for production planning under uncertainty: A review". ***International Journal of Production and Economics***, v.103, pp 271-285. 2006.

- [11] REYES, A. "Estado del arte en planificación con incertidumbre". Reporte Técnico, Doctorado en Ciencias Computacionales-ITESM Cuernavaca-Mexico, Junho, 2003.
- [12] KARGER, D., STEIN, C., WEIN, J. "Scheduling Algorithms" In: J. Atallah, editor, *Handbook on Algorithms and Theory of Computation*, chapter 35. 1999.
- [13] IERAPETRITOU, M., LI, Z. "Process scheduling under uncertainty: Review and challenges". ***Computers and Chemical Engineering***, v. 32, pp. 715-727, 2008
- [14] HANS, E.W. "Robust Surgery loading". ***European Journal Of Operations Research***, v.185, pp. 1038-1050. 2008.
- [15] GUOXUAN, M., BELIEN, J., DEMEULEMEESTER, E. "Solving the strategic case mix problem optimally by using branch-and-price algorithms". In: *International Conference on Operational Research Applied to Health Services (ORAHS)*. Leuven, Belgium. 12-17 July. 2009.
- [16] UNBIASED SYSTEMS ARCHITECTURE, *Block scheduling in the perioperative environment*. Disponível em: <http://www.unibased.com/blockschedulingor.html>. Acesso em: 15 mar. 2013, 15:30.
- [17] HOSSAIN, M. Z. "The use of Box-Cox Transformation Technique in Economic and Statistical Analyses". ***Journal of Emerging Trends in Economics and Management Sciences***, v. 2, nº 1, pp. 32-39. 2011.
- [18] BOX, G.E.P., COX, D.R. "An analysis of transformations". ***Journal of the Royal Statistical Society***, Series B, v.26, nº2, pp. 211-252. 1964.
- [19] WEISBERG, S., ***Applied Linear Regression***. 3 ed. Minnesota, John Wiley & Sons, 2005.
- [20] DOBSON, A.J., ***An Introduction to Generalized Linear Models***. 2 ed. Boca Raton, Florida, Chapman &Hall/CRC, 2002.

[21] SOCIEDADE BRASILEIRA DE MASTOLOGIA. *Estatísticas de câncer de mama*. Disponível em: <http://www.sbmastologia.com.br/cancer-de-mama/rastreamento-diagnostico-cancer-de-mama/estatisticas-sobre-cancer-de-mama-no-brasil-5.htm>.

Acceso em: 20 Feb, 2013, 9:00.

[22] INSTITUTO NACIONAL DE CÂNCER. *Relatório de Gestão 2011*.

Disponível em: [www.inca.gov.br/](http://www.inca.gov.br/)

Acceso em: 24 Feb, 2013, 10:00.

[23] INSTITUTO NACIONAL DE CÂNCER. *Relatório de Gestão 2009*.

Disponível em: [www.inca.gov.br/](http://www.inca.gov.br/)

Acceso em: 24 Feb, 2013, 10:00.

[24] LIMA, C.F.F. *Relatório de desempenho dos processos de assistência do HCIII*.

Disponível em:

[www.inca.gov.br/](http://www.inca.gov.br/)

Acceso em: 24 Feb, 2013, 10:00.

[25] NEUMAYER, L., *Mastectomy*. Disponível em:

[http://www.emedicinehealth.com/mastectomy/article\\_em.htm](http://www.emedicinehealth.com/mastectomy/article_em.htm)

Acceso em: 16 Feb, 2013, 18:30.

[26] ANJOS, L.A. "Índice de massa corporal ( $\text{massa corporal}/\text{estatura}^2$ ) como indicador do estado nutricional de adultos: revisão da literatura". **Revista Saúde Pública**, v. 26, nº 6, São Paulo, 1992.

[27] NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE.

"Hypertension: Clinical management of primary hipertension in adults". Clinical Guideline 34. 2011.

[28] BANCO DE SAÚDE, *Diabetes Mellitus*. Disponível em:

<http://www.bancodesaude.com.br/diabetes/diabetes-mellitus-tipo-1-2-diabetes-gestacional-pre-diabetes>. Acceso em: 20 Mar., 2013. 8:00.

[29] AMERICAN DIABETES ASSOCIATION, *Diabetes Basics*. Disponível em:

<http://www.diabetes.org/diabetes-basics/?loc=GlobalNavDB>. Acceso: 24 Mar., 2013. 9:30.

[30] STRUM, D.P., A.R. SAMPSON, J.H. MAY, AND L.G. VARGAS, "Surgeon and type of anesthesia predict variability in surgical procedure times". *Anesthesiology*, v.92, pp.1454-1456. 2000.

[31] STEPANIAK, P.S., CHRITIAAN, H., DE VRIES, G. "Modeling and prediction of surgical procedure times". *Econometric Intitute Report*. 2009.

# ANEXO I: ESTUDO DE SIMULAÇÃO DAS SALAS DE CIRURGIA

---

## 1. Introdução

Para melhor compreender a vinculação entre a regressão linear e as salas de cirurgia, se realizou uma simulação da sala de operação.

Não era o objetivo principal deste projeto. No entanto, é interessante saber mais sobre a natureza, em termos de duração das cirurgias, dos diferentes procedimentos, e para gerar uma ideia de quais são os limites de utilização razoáveis para trabalhar nas salas de cirurgia.

## 2. Unidade simulada e tempos assumidos

Para simular, se estudou o comportamento de diferentes programações no contexto de **um dia em uma sala de operações**.

Os dados determinísticos (dos quais se poderia estudar a sua verdadeira natureza estocástica) que têm sido utilizados são os seguintes:

- Tempo de preparação (início do dia): 1h
- Tempo de manutenção (*Setup*. Limpeza e preparação entre cirurgias): 00:30 h
- Tempo total disponível em um dia na sala de cirurgia: 12h

O esquema básico de funcionamento de uma sala cirúrgica é o da Figura 2, tal como descrito no marco teórico.



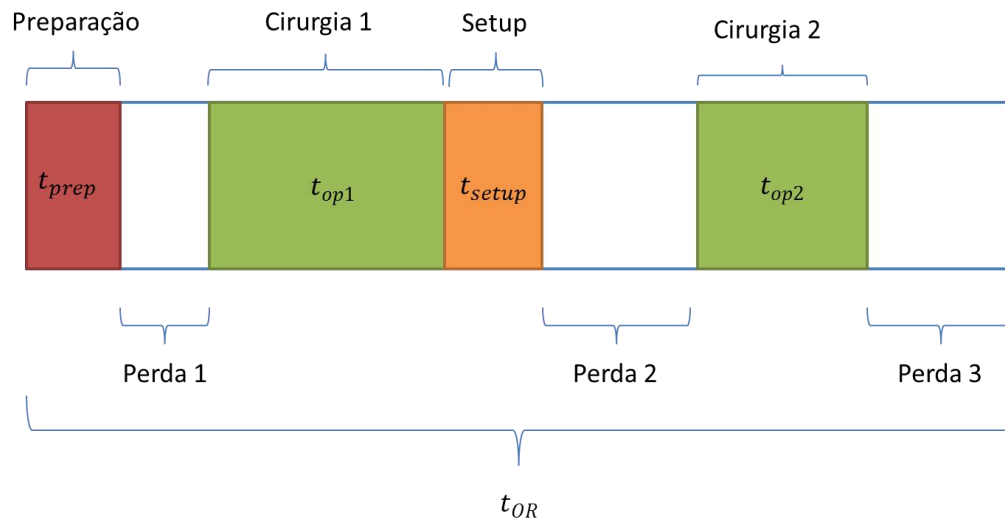


Figura 61. Esquema de tempos de operação, tempo de preparação, setup e perdas em um dia de uma OR. Elaboração própria.

### 3. Casos simulados. Ordem das cirurgias e programações possíveis

Entidades de simulação incluíram os dados que foram analisados para estimar a regressão linear obtidos a partir da amostra aleatória dos últimos três anos, acrescentando que o tempo de permanência anterior (entrada na sala de operações e anestesia) e posterior (fim anestesia e saída da sala de cirurgia).

Para a simulação, se estudaram combinações de **Segmentectomia com BLS (SEG + BLS)**, **Segmentectomia sem BLS (SEGMPC)** **Mastectomia Simples com BLS (MASTEC)** e **Mastectomia Radical Modificada (MRM)**.

O fato de separar a Segmentectomia entre aquelas cirurgias em que a BLS foi realizada ou não responde a que existe um número grande de intervenções em que não se pratica a biópsia (geralmente nos casos em que a Segmentectomia é realizada através da marcação pré-cirúrgica), de modo que o procedimento demora menos (levando em conta o tempo de estadia do paciente na sala de operações é de 1:20 h), nos casos em que não é executada (não existe variabilidade associada com a possibilidade de realização da linfodectomia axilar).

Especificamente se tem simulados três casos diferentes de combinações e sua programação:

- 3 procedimentos em um dia na sala de cirurgia

- 4 procedimentos em um dia da sala de cirurgia

- 5 casos em um dia da sala de cirurgia

A seguinte tabela mostra o de ordens possíveis dos procedimentos (variações com repetição 4 elementos tomados de  $n$  em  $n$ ) e o número de horários (duração de cada procedimento simulado).

| Nº de procedimentos | Ordens possíveis | Programações |
|---------------------|------------------|--------------|
| 3                   | 64               | 78           |
| 4                   | 256              | 120          |
| 5                   | 1024             | 35           |

A unidade de variação mínima entre programação e programação consistiu 0:30h, a que, de fato, é utilizada no hospital.

Como exemplo, na tabela abaixo se mostram os primeiros ordens (em total são total de 64) de procedimentos no caso da simulação de 3 cirurgias.

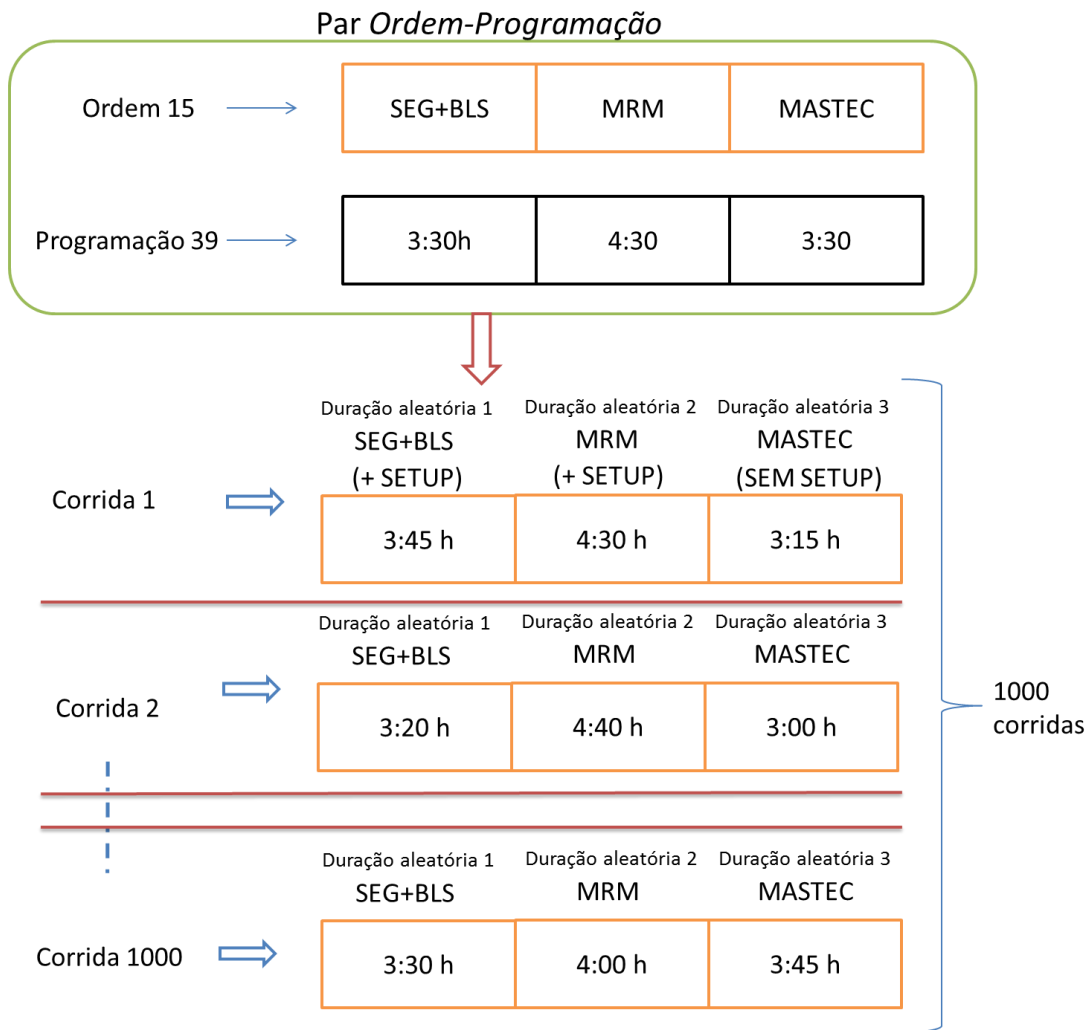
| Ordem de 3 procedimentos | 1       | 2       | 3       |
|--------------------------|---------|---------|---------|
| 1                        | SEG+BLS | SEG+BLS | SEG+BLS |
| 2                        | SEG+BLS | SEG+BLS | SEGMPC  |
| 3                        | SEG+BLS | SEG+BLS | MASTEC  |
| 4                        | SEG+BLS | SEG+BLS | MRM     |
| 5                        | SEG+BLS | SEGMPC  | SEG+BLS |
| 6                        | SEG+BLS | SEGMPC  | SEGMPC  |
| .....                    | .....   | .....   | .....   |
| 64                       | MRM     | MRM     | MRM     |

E, na tabela seguinte se pode ver uma porção do total de durações tentativas que foram avaliadas para cada possível ordem dos três procedimentos.

| Programação [h] | 1       | 2       | 3       |
|-----------------|---------|---------|---------|
| 1               | 2:00:00 | 2:00:00 | 7:00:00 |
| 2               | 2:00:00 | 2:30:00 | 6:30:00 |
| 3               | 2:00:00 | 3:00:00 | 6:00:00 |
| 4               | 2:00:00 | 3:30:00 | 5:30:00 |
| 5               | 2:00:00 | 4:00:00 | 5:00:00 |
| 6               | 2:00:00 | 4:30:00 | 4:30:00 |
| .....           | ...     | ....    | ....    |
| 78              | 7:30:00 | 2:00:00 | 1:30:00 |

Para cada par *ordem de procedimentos - conjunto de durações* foram realizadas 1.000 corridas de simulação selecionando aleatoriamente as durações reais obtidas das amostras dos prontuários médicos (através da geração de números aleatórios no Excel)

A figura abaixo mostra um exemplo de simulação para um dos pares **ordem das cirurgias-programação**.



#### 4. Indicadores da simulação: taxa de utilização e os riscos de atraso

Como foi explicado no marco teórico, há um contraste entre a robustez das programações (entendida como o grau de correspondência relacionada com a sua execução posterior) e a taxa de utilização das salas de operações.

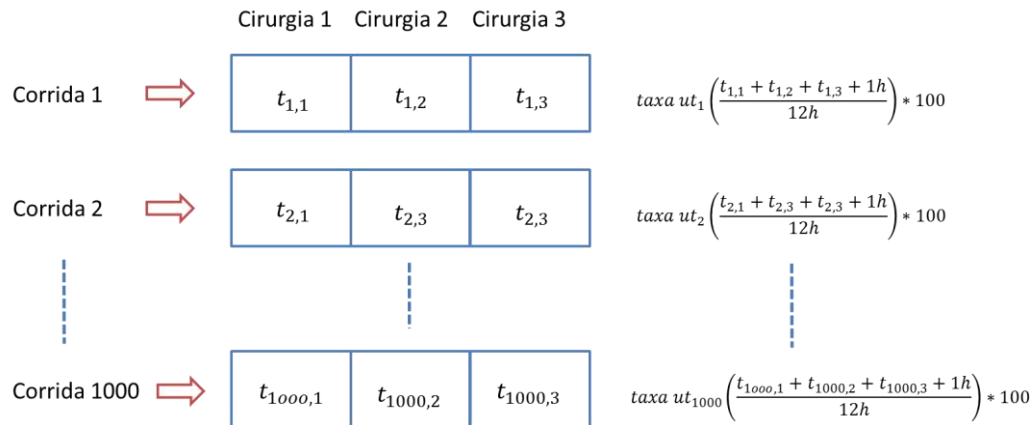
Os indicadores utilizados para avaliar diferentes programações são a **taxa de utilização esperada** e o **risco de demora de uma cirurgia**.

##### 4.1 Simulação da Taxa de utilização esperada

Na seguinte expressão se lembra a taxa de utilização.

$$taxa\ utilização[\%] = \frac{t_{util}}{t_{OR}} = \left( \frac{t_{op} + t_{setup} + t_{prep}}{t_{OR}} \right) * 100$$

Para cada corrida foi calculada a taxa de utilização. A figura a seguir mostra a simulação da taxa de utilização.

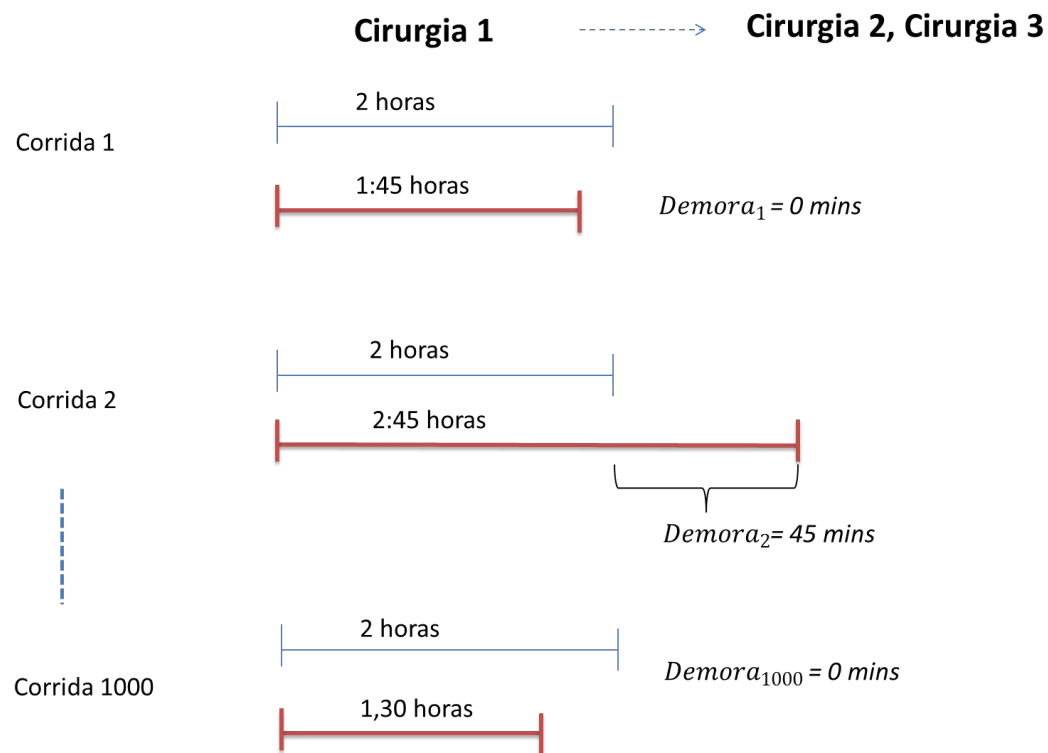


Finalmente, a taxa de utilização esperada para cada ordem de procedimentos é,

$$Taxa\ de\ utilização\ esperada\ [\%] = \frac{\sum_{i=1}^{1000\ corridas} taxa\ ut_i}{1000}$$

#### 4.2 Simulação do risco de demora

Para cada par *ordem-programação* se analisou o atraso de cada cirurgia em uma possível planificação, como mostrado na figura a seguir.



Assim, o risco de atraso de uma cirurgia com respeito ao planejado se define como,

$$Risco\ de\ demora\ [minutos] = \frac{\sum_{i=1}^{1000\ corridas} Demora_i}{1000}$$

## 5. Critérios de seleção da melhor programação

Depois de ter simulado cada possível ordem de procedimentos programação tentativa, foi calculada uma taxa de utilização e o risco de demora associado a cada cirurgia.

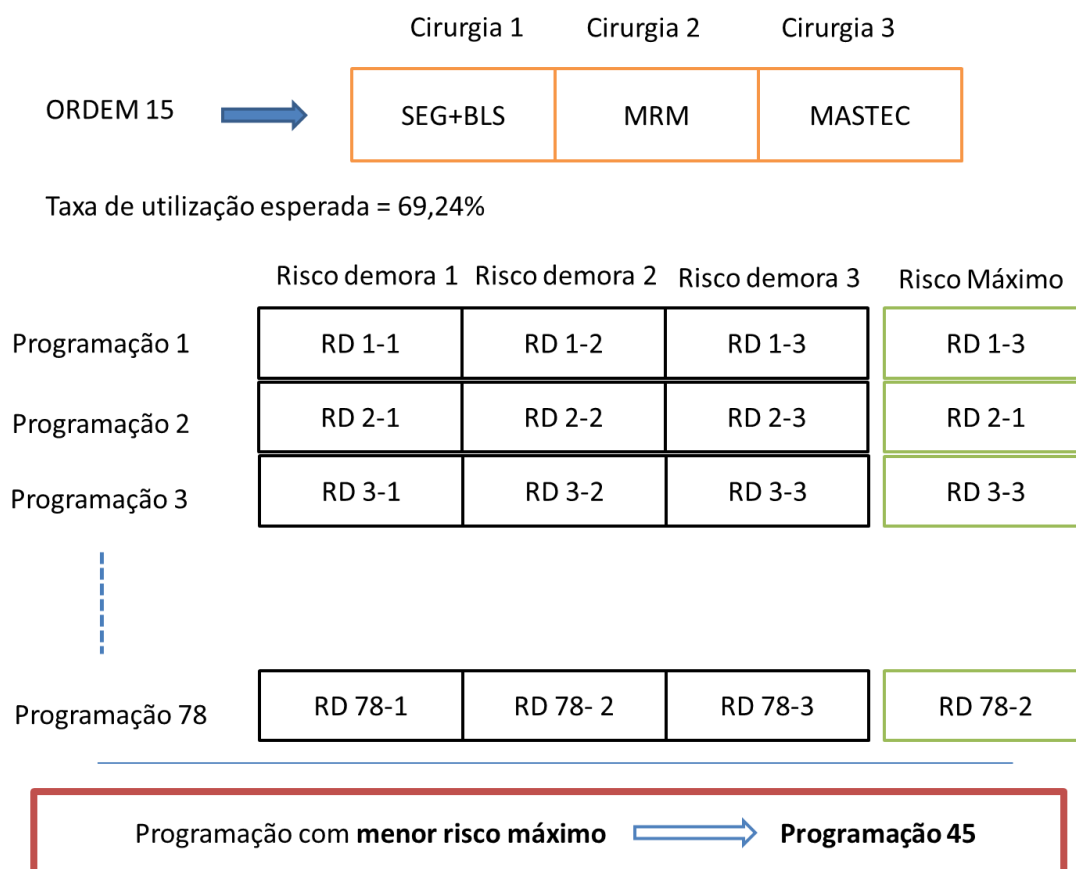
A tabela a seguir mostra um esquema dos resultados obtidos para três procedimentos diferentes simulações nas diferentes simulações.

| Par  | Ordem | Programação | Risco demora |            |            | Taxa de utilização |
|------|-------|-------------|--------------|------------|------------|--------------------|
|      |       |             | Cirurgia 1   | Cirurgia 2 | Cirurgia 3 |                    |
| 1    | 1     | 1           | RD 1         | RD 2       | RD 3       | TU1                |
| 2    | 2     | 1           | RD 1         | RD 2       | RD 3       | TU2                |
| ...  | ...   | ...         | ...          | ...        | ...        | ...                |
| 64   | 64    | 1           | RD 1         | RD 2       | RD 3       | TU64               |
| 65   | 1     | 2           | RD 1         | RD 2       | RD 3       | TU1                |
| 66   | 2     | 2           | RD 1         | RD 2       | RD 3       | TU2                |
| ...  | ...   | ...         | ...          | ...        | ...        | ...                |
| 4992 | 64    | 78          | RD 1         | RD 2       | RD 3       | TU64               |

O resultado é igual para as simulações de 4 e 5 procedimentos.

O critério para a escolha da melhor programação foi selecionar para cada ordem dos procedimentos aquela programação que tem **menor risco de demora máximo por cirurgia**, que seria um critério de otimização.

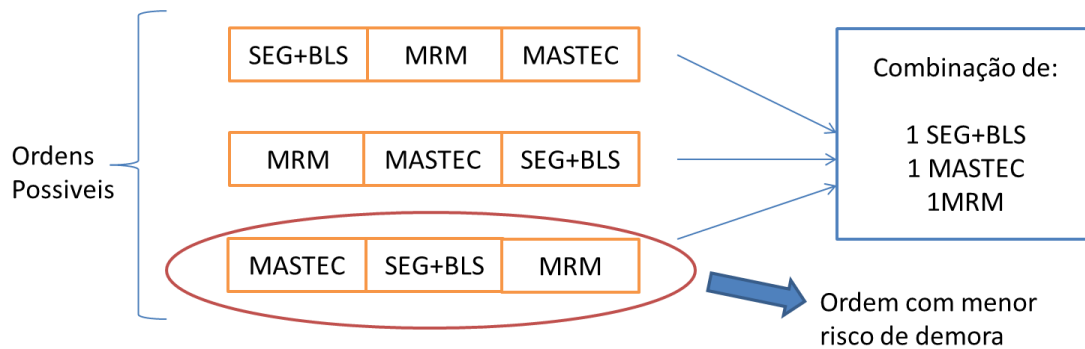
A figura a seguir ilustra o critério de escolha do melhor horário para uma determinada ordem de operações (em três procedimentos).



Foi realizado para todas as ordens possíveis de 3, 4 e 5 cirurgias programadas em uma mesma sala de operações.

## 6. Redução do problema. Seleção de combinações com risco mínimo

Supondo-se que o programador das salas tenha a autoridade para decidir a ordem das cirurgias, se podem selecionar as ordens que minimizem o risco, com os mesmos procedimentos que devem ser realizados no final do dia. A figura a seguir ilustra a seleção da melhor ordem para uma determinada combinação de procedimentos.

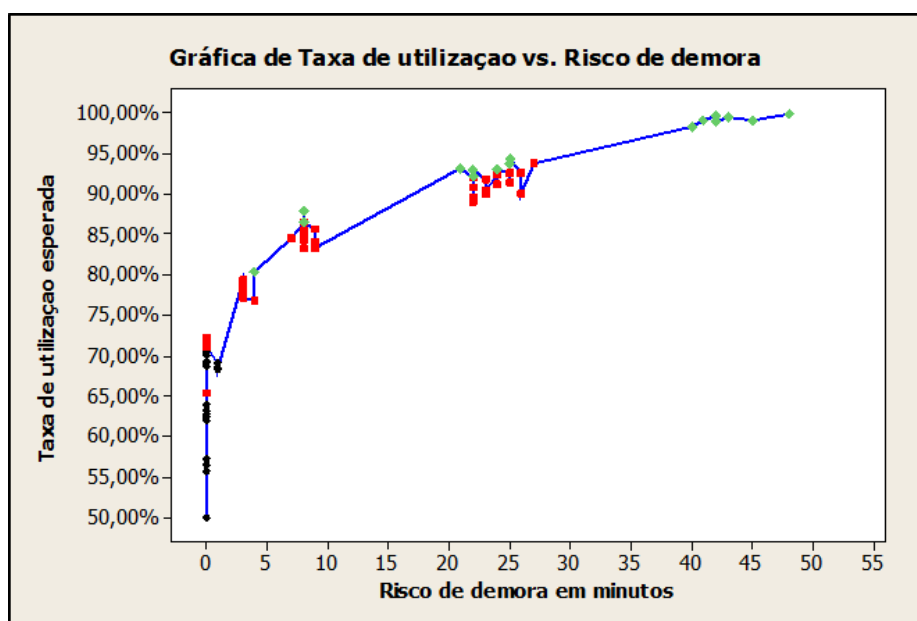


| Nº de procedimentos programados | 3  | 4   | 5   |
|---------------------------------|----|-----|-----|
| Nº de Ordens possíveis          | 64 | 256 | 260 |
| Nº de combinações possíveis     | 20 | 35  | 17* |

Note-se que neste ponto se pode fornecer, para qualquer combinação de procedimentos cirúrgicos a realizar em um dia, a ordem e a duração dos procedimentos que devem ser programados minimizando o risco máximo.

## 7. Taxa de utilização vs. risco máximo de atraso por cirurgia

Uma vez que se tem determinadas as programações ótimas que minimizam o risco de atraso por procedimento, na figura a seguir se tem representadas cada combinação de procedimentos com base na sua taxa de utilização esperada e o risco máximo de atraso.





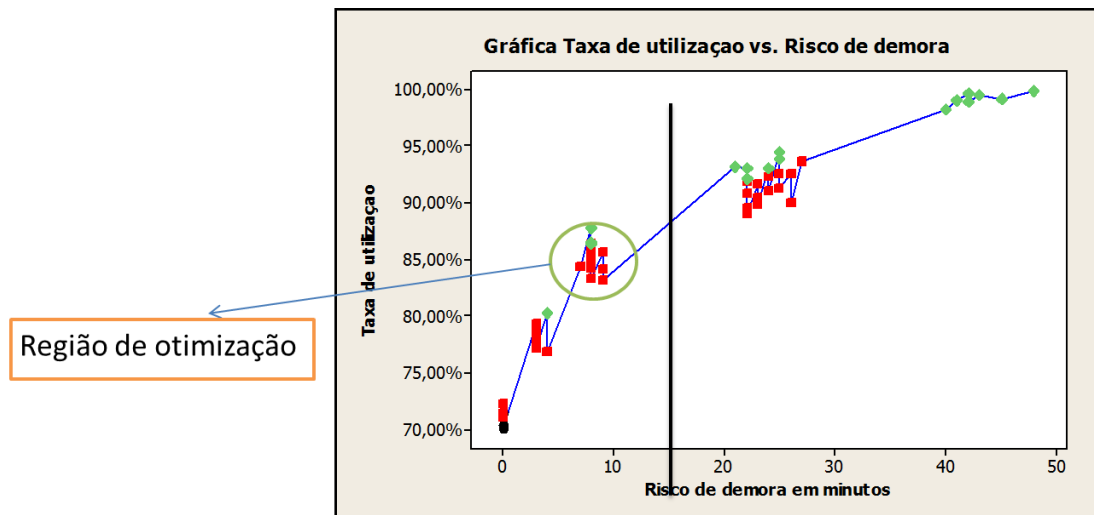
Podem-se realizar as seguintes conclusões:

- a. O aumento em termos de utilização com respeito ao risco não é linear. Há uma fase de aumento elevado da utilização da sala (para 70-75%) em que o aumento do risco máximo de atraso é relativamente baixo (entre 0 e 5 minutos).
- b. Utilizando combinações de programação de procedimentos que excedem o 80% em taxa de utilização supõe um aumento significativo dos riscos de atraso cirurgia.
- c. Existem concentrações elevadas de programações em faixas de risco e taxa de utilização estreitas, o que contribui a determinar regiões de otimização amplias.
- d. Isto também significa que é possível atender grandes flutuações da demanda. O grande número de combinações possíveis que estão em faixas de risco estreitas, permite atender qualquer proporção de procedimentos que se deva realizar.

## 8. Otimização da taxa de utilização das salas de cirurgia

Um critério para determinar quais programações de procedimentos utilizar que maximizem a utilização é, **de acordo com funcionários do hospital**, determinar um risco de atraso máximo que se está disposto a assumir por a cirurgia. Assim, muitos grupos podem ser identificados que maximizariam, com o risco de atraso como restrição, a taxa de utilização da sala de operação.

Imagine-se, como exemplo, que em conversas com os cirurgiões do hospital, determinarem 15 minutos como o risco máximo de atraso por cirurgia que se quer assumir. A figura a seguir ilustra o processo de otimização da sala de cirurgia que poderia ser feito.



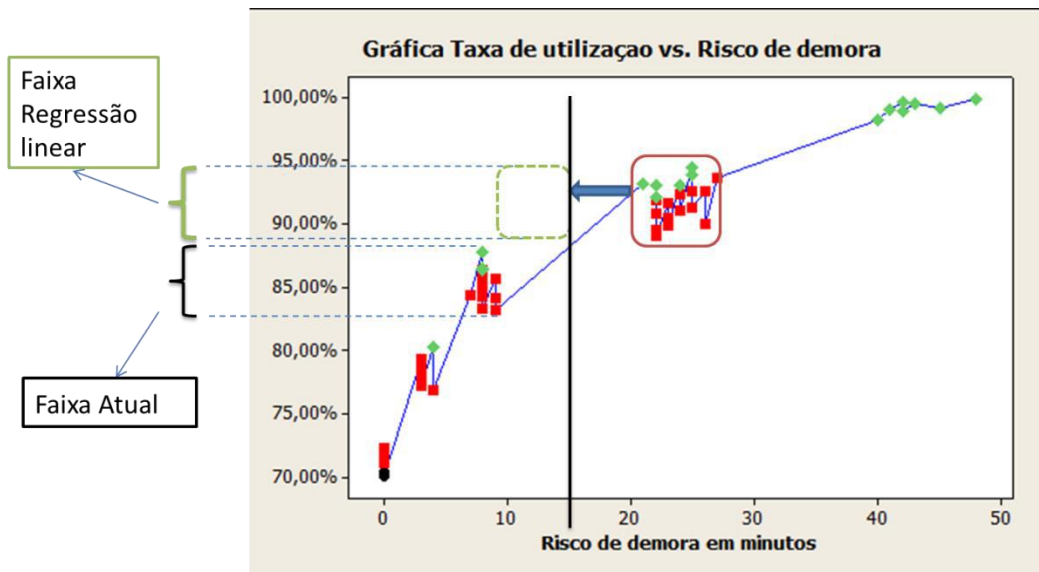
Observa-se que, apenas com um exemplo, um valor de 15 minutos de risco assumido determinaria uma região de otimização consideravelmente rica em combinações de procedimentos, sua ordem e sua programação (em termos de duração tentativa).

9. A regressão linear e o aumento da taxa de utilização da sala de operação através da simulação

No caso em que se pudessem detectar fatores comuns a todas as cirurgias através dos quais obter elevadas percentagens de variabilidade explicada, a redução do risco de atraso de algumas combinações de cirurgias para níveis aceitáveis implicaria a possibilidade de se atingir maiores valores de taxa de utilização da sala de cirurgia.

No caso conseguise reduzir a incerteza associada com a duração dos vários procedimentos, a figura a seguir mostra como você poderia aumentar as taxas de utilização.

A figura a seguir mostra como se poderiam aumentar, neste caso, as taxas de utilização.



Continuando com o exemplo anterior, com um risco reduzido de 10 a 15 minutos, se determinariam novas combinações que aumentariam a cerca de 10% da taxa de utilização de operação, com o consequente aumento do número de pacientes operados no final de ano.