

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE  
CENTRO DE CIÊNCIAS EXATAS DA TERRA  
DEPARTAMENTO DE ESTATÍSTICA  
CURSO DE ESTATÍSTICA**

**DIOGO HENRIQUE LIMA BRASIL**

**ASSOCIAÇÃO ENTRE VARIÁVEIS EM TABELAS DE  
CONTINGÊNCIA USANDO DADOS DE ACIDENTES DO  
TRABALHO**

**NATAL - RN  
NOVEMBRO 2017**

DIOGO HENRIQUE LIMA BRASIL

ASSOCIAÇÃO ENTRE VARIÁVEIS EM TABELAS DE  
CONTINGÊNCIA USANDO DADOS DE ACIDENTES DO  
TRABALHO

Monografia de Graduação apresentada  
ao Departamento de Estatística da  
Universidade Federal do Rio Grande  
do Norte, como requisito para à  
obtenção do título de bacharel em  
Estatística.

Orientador: Prof. Dr. André Luís  
Santos de Pinho

NATAL - RN  
NOVEMBRO 2017

Universidade Federal do Rio Grande do Norte - UFRN  
Sistema de Bibliotecas - SISBI  
Catalogação de Publicação na Fonte. UFRN - Biblioteca Setorial Prof. Ronaldo Xavier de Arruda - CCET

Brasil, Diogo Henrique Lima.

Associação entre variáveis em tabelas de contingência usando dados de acidentes do trabalho / Diogo Henrique Lima Brasil. - 2017.

45 f.: il.

Monografia (graduação) - Universidade Federal do Rio Grande do Norte. Centro de Ciências Exatas e da Terra. Departamento de Estatística. Natal, RN, 2017.

Orientador: André Luís Santos de Pinho.

1. Tabela de contingência - Monografia. 2. Estatística descritiva - Monografia. 3. Teste Qui-quadrado - Monografia. 4. Teste  $G^2$  - Monografia. 5. Análise de correspondência - Monografia. I. Pinho, André Luís Santos de. II. Título.

RN/UF/CCET

CDU 519.235

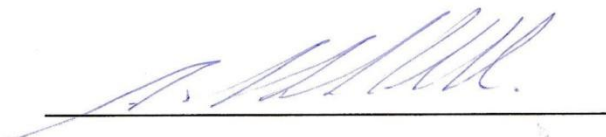
DIOGO HENRIQUE LIMA BRASIL

ASSOCIAÇÃO ENTRE VARIÁVEIS EM TABELAS DE  
CONTINGÊNCIA USANDO DADOS DE ACIDENTES DO  
TRABALHO

Monografia de Graduação apresentada ao Departamento de Estatística da Universidade Federal do Rio Grande do Norte, como requisito para a obtenção do título de bacharel em Estatística.

Aprovada em: 24/11/17

**BANCA EXAMINADORA**



Prof. Dr. André Luís Santos de Pinho

Universidade Federal do Rio Grande do Norte



Prof.ª Dr.ª Carla Almeida Vivacqua

Universidade Federal do Rio Grande do Norte



Prof. Francisco de Assis Medeiros da Silva

Universidade Federal do Rio Grande do Norte

Dedico especialmente à minha avó Pedrina, que infelizmente não está mais conosco, mas que sempre me deu todo suporte que precisei.

## AGRADECIMENTOS

Agradeço, primeiramente a Deus, pois sem Ele nada seria possível.

Aos meus pais Hilton e Sonia pelo suporte de sempre, e ao meu irmão Thiago por estar sempre presente.

Aos meus amigos de curso, que sempre me apoiaram e não me deixaram desistir ao longo desse ciclo.

A todos meus professores, especialmente o Prof. Medeiros e o Prof. André, que sempre foram compreensivos comigo e me deram o suporte necessário para que eu chegasse até aqui.

A Junior, mais conhecido como amigão, por sempre me manter alimentado com seus salgados.

A Ythalo, que por intermédio dele consegui acesso ao banco de dados utilizado nesse estudo.

À minha avó Pedrina, que sem ela, talvez eu nem haveria ingressado na UFRN.

"Não é o que eu sou por dentro,  
mas o que eu faço que me  
define."

- Batman Begins

## RESUMO

O presente estudo tem como objetivo verificar a existência de associação entre o tipo de acidente de trabalho e as unidades de um hospital do Rio Grande do Norte no ano de 2016. Os dados de acidentes de trabalho foram agrupados em uma tabela de contingência, e foram feitas análises descritivas e análises visuais através de gráficos. Para verificar a existência de associação entre as variáveis foram utilizados os testes de independência Qui-quadrado e o teste  $G^2$ , assim como, também foi feita uma análise de correspondência, sendo assim, possível detectar como se dá a associação entre as unidades desse hospital e o tipo de acidente de trabalho.

**Palavras-Chaves:** Tabela de Contingência, Estatística Descritiva, Teste Qui-quadrado, Teste  $G^2$ , Análise de Correspondência.



## **ABSTRACT**

The present study aims to verify the existence of an association between the type of work accident and the units of a hospital in Rio Grande do Norte in the year 2016. The data of work accidents were grouped in a contingency table, and were made descriptive analyzes and visual analysis through graphs. In order to verify the existence of an association between the variables, the Chi-square independence tests and the  $G^2$  test were used, as well as a correspondence analysis. Thus, it is possible to detect the association between the units of this hospital and the type of work accident.

**Keys Words:** Contingency Table, Descriptive Analyzes, Chi-square,  $G^2$  Test, Correspondence Analysis.

## LISTA DE FIGURAS

<i>Figura 1: Gráfico da quantidade de acidentes ocorridos nas unidades do Hospital por tipo, em 2016</i>	<i>21</i>
<i>Figura 2: Gráfico do total de acidentes ocorridos nas unidades do Hospital, em 2016 (%)</i>	<i>22</i>
<i>Figura 3: Gráfico da quantidade de acidentes do tipo Típico ocorridos nas unidades do Hospital, em 2016 (%)</i>	<i>22</i>
<i>Figura 4: Gráfico da quantidade de acidentes do tipo Trajeto ocorridos nas unidades do Hospital, em 2016 (%)</i>	<i>23</i>
<i>Figura 5: Gráfico da quantidade de acidentes do tipo Biológico ocorridos nas unidades do Hospital, em 2016 (%)</i>	<i>23</i>
<i>Figura 6: Mosaico da tabela de contingência (Tabela 2)</i>	<i>26</i>
<i>Figura 7: Gráfico da análise de correspondência</i>	<i>29</i>

## LISTA DE TABELAS

<i>Tabela 1: Registro dos 10 primeiros casos de acidente do trabalho no Hospital, em 2016</i>	<i>18</i>
<i>Tabela 2: Tabela de continência dos tipos de acidentes por unidades do Hospital ocorridos, em 2016</i>	<i>21</i>
<i>Tabela 3: Número médio de acidentes no Hospital por tipo de acidente, em 2016</i>	<i>24</i>
<i>Tabela 4: Número médio de acidentes no Hospital por unidade, em 2016</i>	<i>24</i>
<i>Tabela 5: Valores esperados na tabela de contingência sobre a hipótese de independência</i>	<i>24</i>
<i>Tabela 6: Resíduos padronizados do teste <math>X^2</math></i>	<i>25</i>
<i>Tabela 7: Sub-tabela 1</i>	<i>27</i>
<i>Tabela 8: Sub-tabela 2</i>	<i>27</i>
<i>Tabela 9: Sub-tabela 3</i>	<i>27</i>
<i>Tabela 10: Sub-tabela 4</i>	<i>27</i>
<i>Tabela 11: Resultado do teste <math>G^2</math> e valor-p associado em cada sub-tabela</i>	<i>27</i>
<i>Tabela 12: Escores associados as variáveis Unidade do Hospital e Tipo de Acidente</i>	<i>28</i>

## SUMÁRIO

<b>CAPÍTULO 1 - INTRODUÇÃO</b>	<b>12</b>
1.1 - Segurança do Trabalho	12
1.2 - Acidente de Trabalho	13
<b>CAPÍTULO 2 - OBJETIVOS</b>	<b>15</b>
2.1 - Objetivos Gerais	15
2.2 - Objetivos Específicos	15
<b>CAPÍTULO 3 - DESENVOLVIMENTO</b>	<b>16</b>
3.1 - Estatística e Segurança do Trabalho em âmbito geral	16
3.2 - Conjunto de dados	16
3.3 - Associação entre variáveis em tabelas de contingência	19
<b>CAPÍTULO 4 - ANÁLISE DOS DADOS</b>	<b>21</b>
4.1 - Análises Descritivas dos dados	21
4.2 - Testes de independência	24
<b>CAPÍTULO 5 - CONSIDERAÇÕES FINAIS</b>	<b>30</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>31</b>
<b>ANEXOS</b>	<b>32</b>
ANEXO 1 - Conceitos básicos de Independência em tabelas	32
ANEXO 2 - Banco de Dados	35
ANEXO 3 - Lista de Comandos no R	40
ANEXO 4 - Conceito básico de análise de correspondência	43

## CAPÍTULO 1 - INTRODUÇÃO

---

### 1.1 - Segurança do Trabalho

O entendimento do que é segurança do trabalho pode variar de acordo com a área de trabalho em questão, mas há um consenso entre todos de que se trata de um "conjunto de medidas que são adotadas visando minimizar os acidentes de trabalho, doenças ocupacionais, bem como proteger a integridade e a capacidade de trabalho do trabalhador." (ALBUQUERQUE, 2012)

A segurança do trabalho deve existir em qualquer ambiente que possua algum risco, seja ele físico ou material.

"O quadro de Segurança do Trabalho de uma empresa compõe-se de uma equipe multidisciplinar composta por Técnico de Segurança do Trabalho, Engenheiro de Segurança do Trabalho, Médico do Trabalho e Enfermeiro do Trabalho. Estes profissionais formam o que chamamos de SESMT - Serviço Especializado em Engenharia de Segurança e Medicina do Trabalho." (ALBUQUERQUE, 2012)

No SESMT, "cada profissional desempenha uma função específica dentro da empresa. Os médicos, por exemplo, são responsáveis pelas questões relacionadas à saúde, como prescrição de medicamentos, primeiros socorros, tratamentos, vacinações e diagnósticos e também por tratar dos casos de acidentes que estão sujeitos a acontecer em qualquer empresa. Os engenheiros e técnicos fazem parte da seção operacional, que é responsável por prevenir que aconteçam acidentes de trabalho, mantendo os equipamentos em bom funcionamento." (GRUPO SAÚDE E VIDA, sem data)

Além da SESMT, quando uma empresa possui um quadro de funcionários com mais de 20 trabalhadores, ela deve por lei contar com membros da CIPA - Comissão Interna de Prevenções de Acidentes, a norma que regulamenta essa necessidade é a Norma Regulamentadora 5 - NR5.

Os membros da CIPA têm como missão conscientizar e informar os trabalhadores sobre tudo que englobe segurança do trabalho, normas, uso adequado de equipamentos de proteção, sobre a semana interna de prevenções de acidentes e outros

temas que abordam segurança de modo permanentemente visando sempre a preservação da vida e a saúde.

"Devem constituir CIPA, por estabelecimento, e mantê-la em regular funcionamento as empresas privadas, públicas, sociedades de economia mista, órgãos da administração direta e indireta, instituições beneficentes, associações recreativas, cooperativas, bem como outras instituições que admitam trabalhadores como empregados." (PORTO DE ITAJAÍ, sem data)

## **1.2 - Acidente de Trabalho**

O artigo 19 da Lei nº8.213/91 determina que um acidente de trabalho é aquele que ocorre durante a execução de um trabalho a serviço de uma empresa, provocando lesão corporal ou perturbação funcional que cause a morte, perda ou redução da capacidade do trabalhador.

"Acidente do trabalho ocorre no exercício do trabalho a serviço da empresa, com o segurado empregado, trabalhador avulso, médico residente, bem como com o segurado especial, no exercício de suas atividades." (EQUIPE GUIA TRABALHISTA, sem data)

Além de lesão físicas, existem acidentes que envolvem apenas danos materiais ao empregador, como por exemplo, deixar cair uma prateleira de equipamentos em um estoque, apesar de não existir uma lesão ao trabalhador também é caracterizado um acidente de trabalho.

Os acidentes de trabalho são divididos em três categorias: Típico, Trajeto, Doença de Trabalho(Atípico).

Acidentes do tipo Típico: "Este é o tipo de acidente mais comum, e acontece dentro da empresa durante o horário de expediente. É o caso, por exemplo, de quando o trabalhador cai de uma escada ou se machuca ao manusear um equipamento pesado." (TUIUTI, 2016)

Acidentes do tipo Trajeto: "Acontece durante o percurso do trabalhador de sua casa até o local de trabalho, tanto no início e final do expediente quando no horário de almoço." (TUIUTI, 2016)

Acidentes do tipo Doença de Trabalho(Atípico): "São os acidentes que acontecem dentro ou fora da empresa, devido ao exercício do trabalho, que a lei assemelha aos acidentes de trabalho típico." (TUIUTI, 2016)

O acidente do trabalho será caracterizado tecnicamente pela perícia médica do INSS - Instituto Nacional do Seguro Social, mediante a identificação do nexo entre o trabalho e o agravo.

"A empresa é responsável pela adoção e uso das medidas coletivas e individuais de proteção e segurança da saúde do trabalhador, sendo também seu dever prestar informações pormenorizadas sobre os riscos da operação a executar e do produto a manipular." (EQUIPE GUIA TRABALHISTA, sem data)

Para informar que houve um acidente de trabalho, a empresa deve emitir uma CAT - Comunicado de Acidente de Trabalho.

"A empresa deverá comunicar o acidente do trabalho à Previdência Social até o primeiro dia útil seguinte ao da ocorrência e, em caso de morte, de imediato, à autoridade competente, sob pena de multa variável entre o limite mínimo e o limite máximo do salário-de-contribuição, sucessivamente aumentada nas reincidências, aplicada e cobrada pela Previdência Social. Desta comunicação receberão cópia fiel o acidentado ou seus dependentes, bem como o sindicato a que corresponda a sua categoria. Deverá ser comunicado os acidentes ocorridos com o segurado empregado (exceto o doméstico), o trabalhador avulso, o segurado especial e o médico-residente." (EQUIPE GUIA TRABALHISTA, sem data)

## **CAPÍTULO 2 - OBJETIVOS**

---

### **2.1 - Objetivos Gerais**

Verificar se existe associação entre a variável Tipo de Acidente de Trabalho e a variável Unidade do Hospital, em um Hospital do Rio Grande do Norte.

### **2.2 - Objetivos Específicos**

Conhecer como se dá a natureza da associação entre o Tipo de Acidente de Trabalho e a Unidade do Hospital, em um Hospital do Rio Grande do Norte.



## **CAPÍTULO 3 - DESENVOLVIMENTO**

---

### **3.1 - Estatística e Segurança do Trabalho em âmbito geral**

"As Estatísticas de Saúde e Segurança no Trabalho são utilizadas para acompanhar e monitorar os registros de acidentes do trabalho e as doenças ocupacionais em todo país." (PREVIDÊNCIA SOCIAL, 2014)

"A NR 4 determina que a empresa deve preencher anualmente as estatísticas de acidentes de trabalho e deixar arquivado na empresa, isso se encontra na NR 4, item 4.12 letra i." (NETO, sem data)

Anualmente é apresentado a previdência social o Anuário Estatístico de Acidentes do Trabalho - AEAT, "neste anuário são apresentados dados sobre acidentes do trabalho, suas principais consequências, os setores de atividades econômicas e a localização geográfica de ocorrência dos eventos." (EMPRESA DE TECNOLOGIA E INFORMAÇÕES DA PREVIDÊNCIA SOCIAL, 2015)

O AEAT trabalha com estatísticas descritivas dos acidentes em cada município do país, ou seja, através dele, por exemplo, é possível checar quais tipos de acidentes são mais recorrentes, sendo assim, dando idéia aos responsáveis qual tipo de providência deverá ser tomada para resolver esse problema.

### **3.2 - Conjunto de dados**

Toda empresa é obrigada por lei a possuir um registro de acidentes de trabalho, nesse registro existem dados referente a data do acidente, o tipo de acidente, o lugar do acidente, entre outras variáveis que dependem do ramo da empresa.

Para nosso estudo, tivemos acesso a um banco de dados de um Hospital do município de Natal - Rio Grande do Norte do ano de 2016, mas por fins sigilosos, o nome do hospital e o nome de seus funcionários não serão divulgados neste trabalho.

Esta pesquisa tem como objetivo ir além da estatística descritiva dos dados, e verificar se existe uma diferença estatisticamente significativa no número de acidentes de trabalho por tipo de acidente entre diferentes unidades desse hospital.

Esta pesquisa será uma análise do tipo quantitativa, ou seja, as informações serão traduzidas em números, com esses resultados será possível dizer estatisticamente se existe ou não diferença entre unidades desse hospital.

Serão feitas análises estatísticas usando as frequências dos eventos de interesse, análises descritivas dos dados, testes de aderência, tabelas e gráficos que facilitarão o entendimento dos resultados.

Os dados usados neste estudo são de origem populacional, pois se trata de um registro que abrange todos os funcionários desse hospital, ou seja, essa pesquisa trabalhará resultados censitários.

Como se trata de uma empresa do ramo hospitalar, os tipos de acidentes registrados são: Típico, Trajeto, Biológico e Doença Ocupacional. No ano de 2016 não houve casos de acidentes do tipo "Doença Ocupacional" no hospital, por este motivo, não há dados que possam ser usados para fins comparativos entre as unidades.

O banco de dados foi reorganizado para facilitar a sua visualização, dessa forma ele possui 169 observações e 7 variáveis, e está tabelado da seguinte forma:

*Tabela 1: Registro dos 10 primeiros casos de acidente do trabalho no Hospital, em 2016*

UNIDADE	H. TRAB	DATA DO ACIDENTE	HORA DO ACIDENTE	TÍPICO	TRAJETO	BIOLÓGICO
UNIDADE C	04:30	08/jan	23:30	QUÍMICO		
UNIDADE A	11:30	11/jan	19:40		QUEDA	
UNIDADE B	01:00	11/jan	14:00	QUEDA		
UNIDADE B	01:30	12/jan	14:00	QUEIMADURA		
UNIDADE C	04:09	19/jan	11:09	APRISIONAMENTO		
UNIDADE A	03:10	19/jan	17:10			SANGUE
UNIDADE C	00:30	21/jan	12:30	CORTE		
UNIDADE A	02:40	21/jan	15:40			PERFUROCORTANTE
UNIDADE C	03:30	21/jan	16:30	QUEDA		

*Fonte: Registro do Hospital*

Para facilitar a leitura dos parâmetros vamos usar as seguintes notações:

$F$  = Frequência de acidentes

$$j = \text{Tipo do acidente} \begin{cases} j = 1, & \text{TipoTípico} \\ j = 2, & \text{TipoTrajeto} \\ j = 3, & \text{TipoBiológico} \end{cases}$$

$$i = \text{Unidade do Hospital} \begin{cases} i = 1, & \text{UnidadeA} \\ i = 2, & \text{UnidadeB} \\ i = 3, & \text{UnidadeC} \end{cases}$$

Para o cálculo da média ( $\mu_j$ ) de qualquer um dos tipos ( $j$ ) de acidentes será usado a seguinte fórmula:

$$\mu_j = \sum_{i=1}^3 \frac{F_i}{3}, \quad i = 1, 2, 3$$

em que  $F_i$  é o número de acidentes do tipo em questão da unidade  $i$  do hospital. E a constante 3 corresponde ao número de unidades do hospital.

Para o cálculo da variância dos acidentes( $\sigma_j^2$ ) do tipo( $j$ ) em questão, será usado a seguinte fórmula:

$$\sigma_j^2 = \frac{1}{3} \sum_{i=1}^3 (Fi - \mu_j)^2, i = 1,2,3,$$

Desta forma, o desvio padrão( $\sigma_j$ ) dos acidentes do tipo( $j$ ) em questão se dará da seguinte forma:

$$\sigma_j = \sqrt{\sigma_j^2}$$

### 3.3 - Associação entre variáveis em tabelas de contingência

Para verificar se existe associação entre as variáveis da nossa tabela de contingência, usaremos os testes estatísticos  $X^2$  e  $G^2$ , que têm como principal função averiguar a existência de associação.

Em ambos os testes, se for constatada uma associação entre as variáveis, não será possível saber como se dá a natureza dessa associação, por isso, se faz necessário um passo a mais, que nesse estudo, tratasse da análise dos resíduos padronizados.

Um outro passo além das estatísticas testes que também será abordado nesse estudo é a decomposição do teste  $G^2$ , que será possível subdividindo a tabela de contingência original em sub-tabelas, deste modo, será mais fácil de perceber quais variáveis estão associadas com outras. Para usar deste método se faz necessário seguir algumas regras:

- 1) A soma dos graus de liberdade das sub-tabelas deve ser igual aos graus de liberdade da tabela original.
- 2) A frequência de cada célula na tabela original deve ser a frequência em uma e somente uma sub-tabela.
- 3) Em cada total marginal da tabela original deve ser um total marginal para uma e somente uma sub-tabela.

Para mais detalhes técnicos sobre as estatísticas  $X^2$  e  $G^2$  verificar o anexo (ANEXO 1) deste trabalho.

Também faremos uma de análise de correspondência, que é uma técnica exploratória, e tem como objetivo analisar tabelas de contingência de dupla ou múltipla entrada, usando medidas de correspondência entre linhas e colunas.

No nosso trabalho, faremos uma análise de correspondência simples, pois nossa tabela de contingência é de dupla entrada, neste método as frequências da tabela serão padronizadas, de modo que a soma relativa das células totalizem 1. O objetivo desse método é representar os valores da tabela em relação as distancias da Unidade do Hospital e do Tipo de Acidente no espaço com poucas dimensões.

Para mais detalhes sobre análise de correspondência verificar o anexo (ANEXO 4) deste trabalho.

## CAPÍTULO 4 - ANÁLISE DOS DADOS

### 4.1 - Análises Descritivas dos dados

Nosso interesse com este trabalho é verificar se existe associação entre as variáveis tipos de acidentes de trabalho e as unidades desse Hospital, mas antes de partir para os testes estatísticos, vamos conhecer melhor os dados fazendo uma análise descritiva.

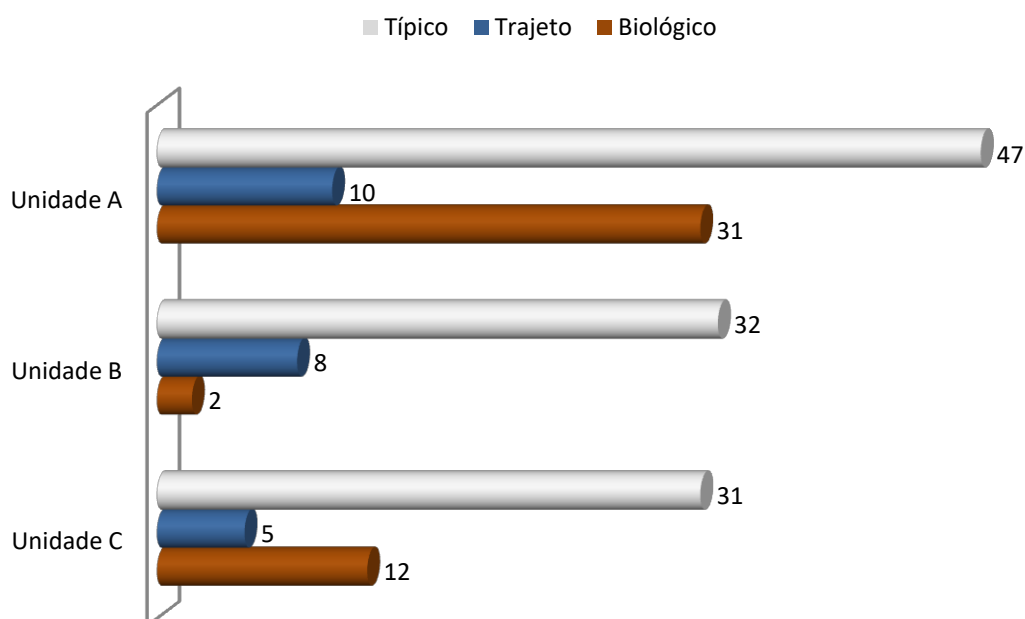
Tabela 2: Tabela de continência dos tipos de acidentes por unidades do Hospital ocorridos, em 2016

	Típico	Trajeto	Biológico
<b>Unidade A</b>	47	10	31
<b>Unidade B</b>	32	8	2
<b>Unidade C</b>	31	5	12

Fonte: Registro do Hospital

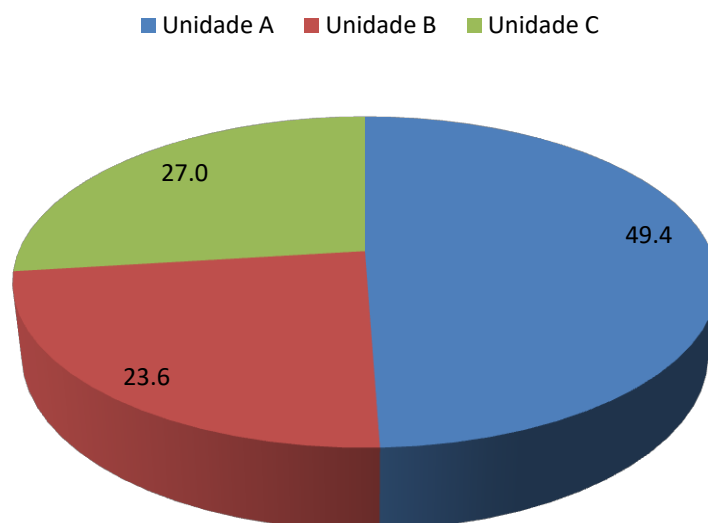
Com base nos dados da Tabela 2 vamos criar uma série gráficos de modo que facilite o entendimento de como se comportam os nossos dados.

Figura 1: Gráfico da quantidade de acidentes ocorridos nas unidades do Hospital por tipo, em 2016



Fonte: Registro do Hospital

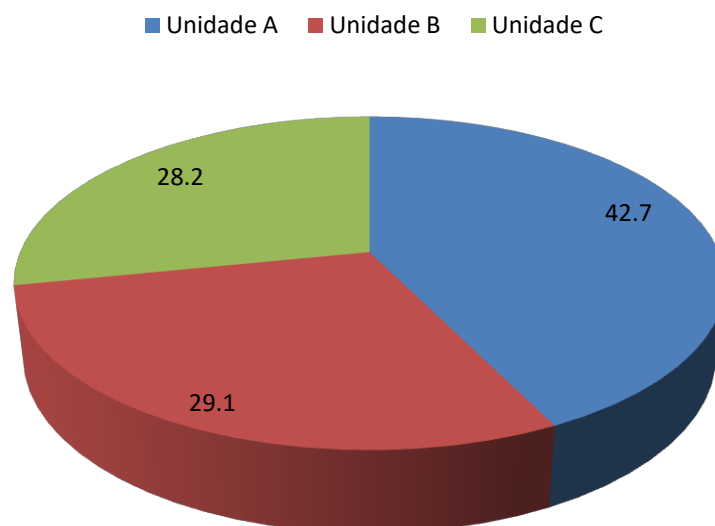
Figura 2: Gráfico do total de acidentes ocorridos nas unidades do Hospital, em 2016 (%)



Fonte: Registro do Hospital

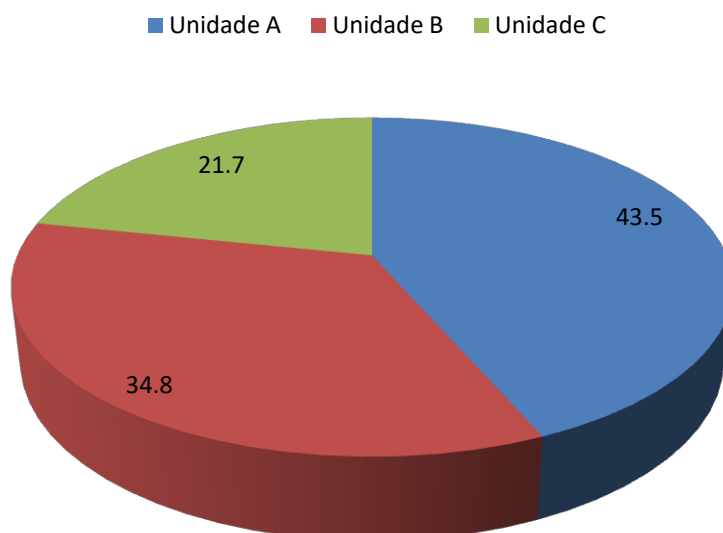
Notamos na Figura 2 que a "Unidade A" detém quase 50% dos casos de acidente de trabalho do Hospital.

Figura 3: Gráfico da quantidade de acidentes do tipo Típico ocorridos nas unidades do Hospital, em 2016 (%)



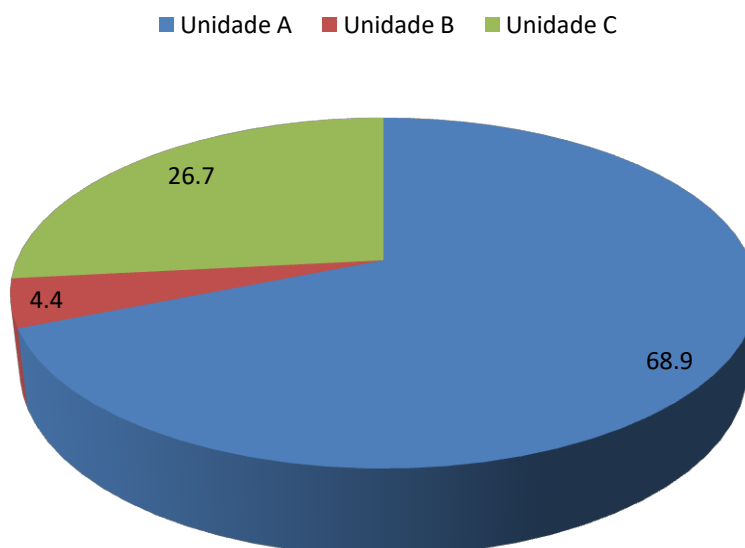
Fonte: Registro do Hospital

Figura 4: Gráfico da quantidade de acidentes do tipo Trajeto ocorridos nas unidades do Hospital, em 2016 (%)



Fonte: Registro do Hospital

Figura 5: Gráfico da quantidade de acidentes do tipo Biológico ocorridos nas unidades do Hospital, em 2016 (%)



Fonte: Registro do Hospital



Tabela 3: Número médio de acidentes no Hospital por tipo de acidente, em 2016

	Típico	Trajeto	Biológico
<b>Número Médio de Acidentes</b>	36.7	7.7	15.0
<b>Desvio Padrão</b>	7.3	2.1	12.0

Fonte: Registro do Hospital

Tabela 4: Número médio de acidentes no Hospital por unidade, em 2016

	Unidade A	Unidade B	Unidade C
<b>Número Médio de Acidentes</b>	29.3	14.0	16.0
<b>Desvio Padrão</b>	15.2	13.0	11.0

Fonte: Registro do Hospital

Baseando-se nas Figuras 3, 4 e 5, como também na Tabela 4 notamos que a "Unidade A" sempre registra um número maior de acidentes, uma das possíveis explicações para isso é se a "Unidade A" possuir um número maior de funcionário que as demais unidades, com isso, aumentando o número de possíveis ocorrências de acidentes em relação às outras unidades desse Hospital.

#### 4.2 - Testes de independência

Para verificarmos se a variável tipo de acidente tem alguma associação com as unidades do Hospital, iremos utilizar o teste de independência  $X^2$  e  $G^2$ , nesses testes utilizaremos um nível de significância( $\alpha$ ) de 0.05.

Para a realização do teste  $X^2$  se faz necessária a construção de uma tabela de valores esperados.

Tabela 5: Valores esperados na tabela de contingência sobre a hipótese de independência

	Típico	Trajeto	Biológico
<b>Unidade A</b>	52.6	12.0	23.4
<b>Unidade B</b>	19.7	4.5	8.8
<b>Unidade C</b>	28.7	6.5	12.8

Fonte: Registro do Hospital

Notamos que uma das células dos valores esperados na Tabela 5 é menor que 5, porém não corresponde a mais de 20% das frequências esperadas, sendo essa uma das condições para aplicação o teste  $X^2$ .

Utilizando o (R Core Team, 2017) na sua versão 1.0.153 para realização dos testes, primeiro faremos o teste  $X^2$ .

Depois de feito o teste, temos um  $X^2 = 12.487$ , com 4 graus de liberdade e com um valor-p de  $0.0143 < 0.05 = \alpha$ , dessa forma rejeitamos a hipótese nula de independência entre as variáveis, sendo assim, existe uma associação entre a unidade do hospital e o tipo de acidente.

Notado há existência de associação, agora estaremos interessados em saber quais categorias nos levaram a rejeitar a hipótese de independência, primeiro vamos analisar como se comportam os resíduos padronizados no teste  $X^2$ .

*Tabela 6: Resíduos padronizados do teste  $X^2$*

	<b>Típico</b>	<b>Trajeto</b>	<b>Biológico</b>
<b>Unidade A</b>	-1.8	-0.9	2.6
<b>Unidade B</b>	1.3	2.0	-3.0
<b>Unidade C</b>	0.8	-0.8	-0.3

*Fonte: Registro do Hospital*

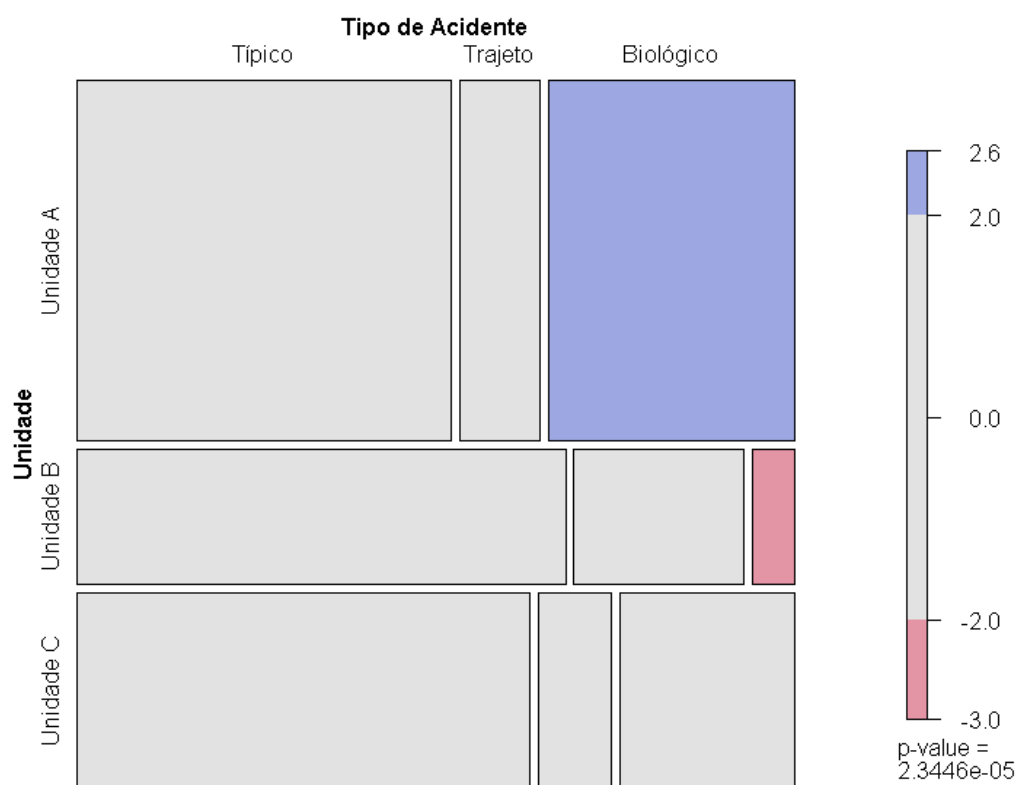
Notamos que dois resultados se destacam a mais que os outros nessa análise de resíduos:

a) O número de acidentes do tipo Biológico na unidade B é menor do que seria se as variáveis fossem independentes.

b) O número de acidentes do tipo Biológico na Unidade A é maior do que seria se as variáveis fossem independentes.

Para uma melhor visualização dessas informações vamos utilizar um gráfico de mosaico.

Figura 6: Mosaico da tabela de contingência (Tabela 2)



Fonte: Registro do Hospital

Na Figura 6, o comprimento dos retângulos representa o número de pessoas acidentadas por categoria, a largura dos retângulos representa a probabilidade de um indivíduo ser daquela Unidade. Além disso, a cor está associada à análise dos resíduos padronizados, em que azul significa que se esperava um número menor de indivíduos naquela categoria, considerando a hipótese de independência entre as variáveis, e a cor vermelha significa que se esperava mais indivíduos naquela categoria, também sobre a hipótese de independência.

Agora usaremos também o teste  $G^2$  para verificar se existe ou não associação entre as unidades do hospital e o tipo de acidente.

Feito o teste, temos um  $G^2 = 14.1915$ , com 4 graus de liberdade e com um valor-p de  $0.0067 < 0.05 = \alpha$ , desta forma, rejeitamos a hipótese de independência entre as variáveis unidade do hospital e tipo de acidente, de modo que fica constatada a existência de associação entre elas.

Agora iremos dividir a nossa Tabela 2 em 4 sub-tabelas, dessa forma será possível verificar como é a natureza dessa associação, utilizaremos uma propriedade do teste  $G^2$ , em que a soma dos testes  $G^2$  aplicado em cada sub-tabela resultará no teste  $G^2$  aplicado na tabela original.

*Tabela 7: Sub-tabela 1*

	<b>Típico</b>	<b>Trajeto</b>
<b>Unidade A</b>	47	10
<b>Unidade B</b>	8	23

*Fonte: Registro do Hospital*

*Tabela 8: Sub-tabela 2*

	<b>Típico + Trajeto</b>	<b>Biológico</b>
<b>Unidade A</b>	57	31
<b>Unidade B</b>	31	2

*Fonte: Registro do Hospital*

*Tabela 9: Sub-tabela 3*

	<b>Típico</b>	<b>Trajeto</b>
<b>Unidade A + Unidade B</b>	70	18
<b>Unidade C</b>	31	5

*Fonte: Registro do Hospital*

*Tabela 10: Sub-tabela 4*

	<b>Típico + Trajeto</b>	<b>Biológico</b>
<b>Unidade A + Unidade B</b>	88	33
<b>Unidade C</b>	36	12

*Fonte: Registro do Hospital*

Aplicando o teste  $G^2$  em cada sub-tabela, teremos:

*Tabela 11: Resultado do teste  $G^2$  e valor-p associado em cada sub-tabela*

<b>Sub-Tabela</b>	<b>Graus de Liberdade</b>	<b><math>G^2</math></b>	<b>Valor-p</b>
<b>Sub-tabela 1</b>	1	0.82288599	0.3643377204
<b>Sub-tabela 2</b>	1	12.51460184	0.0004037838
<b>Sub-tabela 3</b>	1	0.76240999	0.3825753726
<b>Sub-tabela 4</b>	1	0.09160578	0.7621455769

*Fonte: Registro do Hospital*

Note que rejeitamos a hipótese de independência apenas na Sub-tabela 2, na qual o teste  $G^2$  corresponde a 12.51460184 com 1 grau de liberdade e com um valor-p associado de  $0.0004037838 < 0.05 = \alpha$ , ou seja, temos fortes indícios que o acidente do tipo Biológico tem associação com a Unidade A e Unidade B desse Hospital.

Faremos agora uma análise de correspondência para verificar a estrutura de relações entre as variáveis Unidade do Hospital e Tipo de Acidente, primeiro encontraremos os "escores" associados.

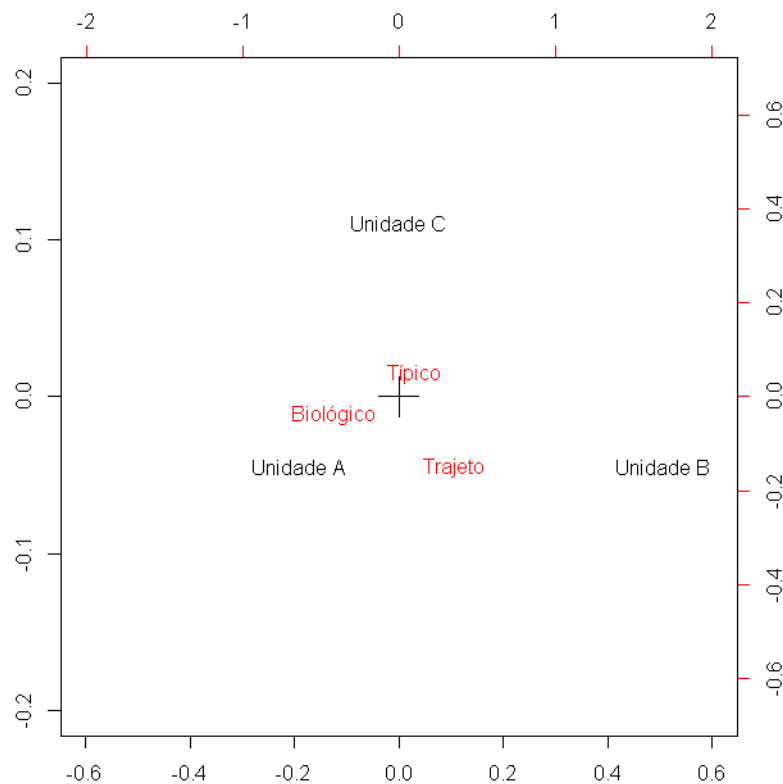
*Tabela 12: Escores associados as variáveis Unidade do Hospital e Tipo de Acidente*

<b>Unidade</b>	<b>Escores</b>
<b>Unidade A</b>	-0.725452279
<b>Unidade B</b>	1.928151457
<b>Unidade C</b>	0.004391719
<b>Tipo de Acidente</b>	<b>Escores</b>
<b>Típico</b>	0.3915752
<b>Trajeto</b>	1.3562114
<b>Biológico</b>	-1.5720435

*Fonte: Registro do Hospital*

Agora faremos o gráfico da análise de correspondência para visualizar a existência de associação entre as variáveis. Para mais informações sobre os métodos para análise de correspondência verificar o anexo (ANEXO 4) deste trabalho.

Figura 7: Gráfico da análise de correspondência



Fonte: Registro do Hospital

Através da Figura 7 notamos a existência de uma associação entre a Unidade A do Hospital e o Acidente do Tipo Biológico.

## CAPÍTULO 5 - CONSIDERAÇÕES FINAIS

---

Este estudo nos permitiu checar a existência de associação entre as variáveis Unidade e Tipo de Acidente de Trabalho em um Hospital do Rio Grande do Norte no ano de 2016. Além disso, nos permitiu saber quais unidades estão associadas a esses acidentes, requerendo assim, mais conhecimento na área de estatística.

Alguns acidentes de trabalho costumam incapacitar o trabalhador em fazer as suas atividades dentro da empresa, sendo assim, o empregado é obrigado a se ausentar da sua jornada diária de trabalho por dias ou meses, o que para a empresa pode significar diretamente uma perda de desempenho. Saber qual tipo de acidente é mais frequente, nos ajuda a perceber quais atitudes devemos tomar para sanar esse problema.

Ao fazer os testes estatísticos notamos a existência de associação entre essas variáveis, além disso, temos razões para acreditar que essa associação se dá entre o Tipo de Acidente Biológico com as Unidade A e Unidade B deste hospital.

Os acidentes do Tipo Biológico, segundo os registros do Hospital, são aqueles casos em que os trabalhadores foram expostos a sangue, perfuro cortante e secreções. Este tipo de acidente demonstrou ter uma associação prejudicial com a Unidade A, pois se não houvesse associação entre as variáveis, o número de casos de acidentes do Tipo Biológico na Unidade A deveria ser menor do que o constatado.

Uma das possíveis soluções para resolver este problema na Unidade A seria orientar os membros da CIPA do Hospital a fiscalizar com mais rigor as áreas onde as chances desse tipo de acidente sejam maiores, e orientar também aos demais funcionários a prestar um pouco mais de atenção ao mapa de risco do hospital, assim, evitando além de acidentes do Tipo Biológico, também os outros demais Tipo de Acidentes de Trabalho.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALBUQUERQUE, D. (2012). Acesso em 1 de Agosto de 2017, disponível em Templum: <http://certificacaoiso.com.br/e-seguranca-trabalho/>
- COSTA, A. (26 de Agosto de 2016). ANÁLISE DE CORRESPONDÊNCIA SIMPLES COM NOVOS ESCORES E O USO DA ANÁLISE DE CORRESPONDÊNCIA MÚLTIPLA EM DADOS COMPOSICIONAIS DE GRANULOMETRIA DE GRÃOS DE CAFÉ. Lavras, Minas Gerais, Brasil.
- EMPRESA DE TECNOLOGIA E INFORMAÇÕES DA PREVIDÊNCIA SOCIAL. (2015). *Anuário Estatístico de Acidentes do Trabalho*. Acesso em 3 de Agosto de 2017, disponível em previdencia: <http://www.previdencia.gov.br/wp-content/uploads/2017/05/aeat15.pdf>
- EQUIPE GUIA TRABALHISTA. (sem data). Acesso em 1 de Agosto de 2017, disponível em guiatrabalhista: <http://www.guiatrabalhista.com.br/noticias/trabalhista210306.htm>
- FOGO, C. (23 de Setembro de 2017). *DEs - UFSCar*. Acesso em 9 de Novembro de 2017, disponível em ufscar: [http://www.ufscar.br/jcfogo/EMult\\_2/arquivos/Exemplo\\_A\\_Corresp.pdf](http://www.ufscar.br/jcfogo/EMult_2/arquivos/Exemplo_A_Corresp.pdf)
- GRUPO SAÚDE E VIDA. (sem data). Acesso em 15 de Agosto de 2017, disponível em Saudeevida: <http://www.saudeevida.com.br/o-que-e-sesmt/>
- NETO, W. (sem data). *SEGURANÇA DO TRABALHO NWN*. Acesso em 31 de Agosto de 2017, disponível em segurancadotrabalhonwn: <https://segurancadotrabalhonwn.com/estatistica-aplicada-a-seguranca-do-trabalho/>
- PORTO DE ITAJAÍ. (sem data). Acesso em 26 de Agosto de 2017, disponível em portoitajai: <http://www.portoitajai.com.br/cipa/legislacao.html>
- PREVIDÊNCIA SOCIAL. (25 de Agosto de 2014). Acesso em 31 de Agosto de 2017, disponível em previdencia: <http://www.previdencia.gov.br/saude-e-seguranca-do-trabalhador/estatisticas-de-saude-e-seguranca-no-trabalho/>
- SILVA, D. (6 de Abril de 2017). Análise de Dados Categorizados. *Inferência Para Tabelas de Contigência*, pp. 25-39.
- TUIUTI. (4 de Novembro de 2016). Acesso em 4 de Agosto de 2017, disponível em epi-tuiuti: <http://www.epi-tuiuti.com.br/blog/saiba-quais-sao-os-3-tipos-de-acidente-de-trabalho/>
- R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.



## ANEXOS

### ANEXO 1 - Conceitos básicos de Independência em tabelas

Para estes conceitos vamos usar como base o material de Nobrega (2017) Vamos Considerar uma tabela de contingência com  $i$  linhas e  $j$  colunas. Agora vamos supor que as frequências desta tabela seguem o modelo de amostragem multinomial com tamanho amostral  $n$  fixo e probabilidades  $\{\pi_{ij}\}$ . Definindo o teste de hipóteses como:

$$H_0: \pi_{ij} = \pi_{i+}\pi_{+j}; \text{ para todas as células } (i, j); i = 1, \dots, I; j = 1, \dots, J$$

$$H_1: \pi_{ij} \neq \pi_{i+}\pi_{+j}; \text{ para pelo menos uma das células } (i, j)$$

Este será um teste da hipótese de independência ou não associação entre as variáveis X e Y.

O estimador de máxima verossimilhança(EMV) de  $\pi_{ij}$  sob a hipótese nula é  $\hat{\pi}_{ij} = (n_{i+}/n)(n_{+j}/n)$ , em que  $n_{i+} = \sum_j n_{ij}$  e  $n_{+j} = \sum_i n_{ij}$  são o total das marginais da linha  $i$  e coluna  $j$ , respectivamente. Logo, sob a hipótese  $H_0$ , o EMV da frequência esperada da célula  $(i ; j)$  da tabela é dada por:

$$\hat{\mu}_{ij} = n\hat{\pi}_{ij} = n \frac{n_{i+}}{n} \frac{n_{+j}}{n} = \frac{n_{i+}n_{+j}}{n}$$

Duas estatísticas de teste possíveis para estas frequências estimadas são dadas pela estatística de Pearson:

$$X^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(n_{ij} - \hat{\mu}_{ij})^2}{\hat{\mu}_{ij}}$$

e pela estatística que é -2 vezes o logaritmo da razão de verossimilhança generalizada:

$$G^2 = 2 \sum_{i=1}^I \sum_{j=1}^J n_{ij} \ln \left( \frac{n_{ij}}{\hat{\mu}_{ij}} \right)$$

Os graus de liberdade desta estatística são:

$$g.l. = (IJ - 1) - [I - 1 + J - 1] = (I - 1)(J - 1)$$

Em que o termo  $[I - 1 + J - 1]$  se dá pelo fato que, sob  $H_0$ , as probabilidades conjuntas  $\{\pi_{ij}\}$  são representadas apenas pelas suas marginais  $\{\pi_{i+}\}$  e  $\{\pi_{+j}\}$  e a

primeira distribuição marginal tem  $I - 1$  parâmetros livres, enquanto que a segunda distribuição tem  $J - 1$  parâmetros livres.

O teste para grandes amostras de  $H_0$  é feito com base na distribuição  $X^2_{(I-1)(J-1)}$ , de tal forma que se deve rejeitar  $H_0$  ao nível de significância  $\alpha$  se

$$\text{Valor. } p = P(X^2_{(I-1)(J-1)} > \text{valor observada da estatística}) < \alpha$$

A convergência da distribuição amostral de  $X^2$  ou  $G^2$  é estudada fixando-se o número de células e tomando-se  $n \rightarrow \infty$ , de tal forma que cada  $\mu_{ij} = n\pi_{ij} \rightarrow \infty$ . Em síntese, o crescimento dos valores médios das células torna a distribuição conjunta das frequências  $\{n_{ij}\}$  próxima da distribuição normal multivariada e, por consequência, as distribuições amostrais de  $X^2$  e  $G^2$  próximas da distribuição qui-quadrado. Não é possível dizer com certeza o quão grande deve ser o tamanho das frequências para essa convergência, mas o fato de  $\{\mu_{ij} \geq 5\}$  já é considerado o suficiente.

Apesar de  $X^2$  e  $G^2$  serem estatísticas que testam associação, apenas com elas não é possível dizer como se dá a natureza dessa associação, sendo assim se fazendo necessário um análise mais detalhada para suprir essa deficiência dos testes. Um dos procedimentos usados é a análise dos resíduos das células da tabela, a fim de verificar a diferenças entre os valores observados e os esperados pela hipótese de independência.

Um dos procedimentos para a análise posterior aos testes de independência é a análise dos resíduos padronizados, que se dão da seguinte forma:

$$r_{ij} = \frac{n_{ij} - \hat{\mu}_{ij}}{\sqrt{\hat{\mu}_{ij}(1 - p_{i+})(1 - p_{+j})}}$$

Estes resíduos têm distribuição assintoticamente normal padrão, de tal forma que quando o valor absoluto de um deles for 2 ou 3, há indicação de falta de ajuste com a hipótese de independência.

Além da análise de resíduos, uma outra alternativa para checar a natureza da associação dos dados posterior aos testes  $X^2$  e  $G^2$  é a decomposição da estatística  $G^2$  em parcelas independentes. Este procedimento é baseado na propriedade reprodutiva da distribuição qui-quadrado. Esta propriedade diz que se  $G^2$  tem  $v > 1$  graus de

liberdade, então  $G^2$  pode ser decomposta pela soma de componentes independentes  $G^2 = G_1^2 + G_2^2 + G_3^2 + \dots + G_v^2$ , em que cada  $G_i^2 \sim X_1^2$ ,  $i = 1, \dots, v$ . Este tipo de análise é possível pelo método da partição da tabela em sub-tabelas. Algumas regras que devem ser usadas são as seguintes:

1) Os graus de liberdade para as sub-tabelas devem somar os graus de liberdade da tabela completa.

2) A frequência de cada célula na tabela completa deve ser a frequência em uma e somente uma sub-tabela.

3) Cada total marginal da tabela completa deve ser um total marginal para uma e somente uma sub-tabela

## ANEXO 2 - Banco de Dados

UNIDADE	H. TRAB	DATA DO ACIDENTE	HORA DO ACIDENTE	TÍPICO	TRAJETO	BIOLÓGICO
UNIDADE C	04:30	08/jan	23:30	QUÍMICO		
UNIDADE A	11:30	11/jan	19:40		QUEDA	
UNIDADE B	01:00	11/jan	14:00	QUEDA		
UNIDADE B	01:30	12/jan	14:00	QUEIMADURA		
UNIDADE C	04:09	19/jan	11:09	APRISIONAMENTO		
UNIDADE A	03:10	19/jan	17:10			SANGUE
UNIDADE C	00:30	21/jan	12:30	CORTE		
UNIDADE A	02:40	21/jan	15:40			PERFUROCORTANTE
UNIDADE C	03:30	21/jan	16:30	QUEDA		
UNIDADE A	03:25	28/jan	09:25	IMPACTO		
UNIDADE C	01:30	28/jan	14:30	APRISIONAMENTO		
UNIDADE C	09:30	29/jan	04:30	CORTE		
UNIDADE A	03:00	29/jan	10:00	IMPACTO		
UNIDADE C	02:30	29/jan	14:30	IMPACTO		
UNIDADE A	00:00	01/fev	05:45		QUEDA	
UNIDADE C	04:40	01/fev	11:40	CORTE		
UNIDADE A	11:15	04/fev	17:15	QUEIMADURA		
UNIDADE C	01:30	05/fev	08:30	APRISIONAMENTO		
UNIDADE A	01:40	05/fev	14:40			PERFUROCORTANTE
UNIDADE C	01:45	11/fev	14:45			PERFUROCORTANTE
UNIDADE A	05:30	12/fev	13:30	APRISIONAMENTO		
UNIDADE C	02:00	15/fev	21:00	QUEDA		
UNIDADE B	07:20	17/fev	13:50	QUEIMADURA		
UNIDADE B	00:00	25/fev	05:50		QUEDA	
UNIDADE A	02:00	26/fev	15:00			PERFUROCORTANTE
UNIDADE C	00:00	03/mar	06:40		IMPACTO	
UNIDADE A	02:30	15/mar	09:00	QUEDA		
UNIDADE A	06:20	15/mar	09:05	QUEDA		
UNIDADE C	05:10	26/mar	18:10			PERFUROCORTANTE
UNIDADE B	10:00	21/mar	03:00	ATRITO		
UNIDADE C	15:15	21/mar	08:15	QUEDA		
UNIDADE B	17:00	24/mar	09:00		QUEDA	

UNIDADE A	10:00	25/mar	05:00		PERFUROCORTANTE
UNIDADE B	12:20	26/mar	00:20	IMPACTO	
UNIDADE C	04:00	28/mar	11:10		PERFUROCORTANTE
UNIDADE A	05:50	29/mar	18:50		PERFUROCORTANTE
UNIDADE A	05:35	30/mar	12:35	QUEDA	
UNIDADE C	06:15	01/abr	13:15	QUEDA	
UNIDADE A	02:20	01/abr	15:20		PERFUROCORTANTE
UNIDADE B	02:08	04/abr	09:20		PERFUROCORTANTE
UNIDADE A	06:00	06/abr	13:50		QUEDA
UNIDADE A	02:30	07/abr	15:30	IMPACTO	
UNIDADE C	00:00	08/abr	06:25		QUEDA
UNIDADE B	02:15	08/abr	09:15	IMPACTO	
UNIDADE A	08:00	09/abr	15:00	QUEIMADURA	
UNIDADE B	06:00	10/abr	12:00	QUÍMICO	
UNIDADE B	02:08	11/abr	09:20		SEM INFORMAÇÃO
UNIDADE A	04:40	12/abr	17:40		PERFUROCORTANTE
UNIDADE A	06:00	15/abr	13:22		QUEDA
UNIDADE B	00:10	18/abr	09:10	CORTE	
UNIDADE A	05:20	18/abr	18:20		PERFUROCORTANTE
UNIDADE B	06:00	19/abr	13:00	QUEDA	
UNIDADE A	01:30	19/abr	07:30	CORTE	
UNIDADE A	03:00	24/abr	10:00	QUEDA	
UNIDADE A	04:35	26/abr	16:35	QUEIMADURA	
UNIDADE A	10:00	01/mai	04:20	QUEDA	
UNIDADE C	03:30	02/mai	22:30	QUEDA	
UNIDADE A	10:35	03/mai	05:30	IMPACTO	
UNIDADE A	06:00	03/mai	13:20	IMPACTO	
UNIDADE A	03:20	09/mai	10:20	ATRITO	
UNIDADE A	05:50	09/mai	12:50		SANGUE
UNIDADE A	02:30	11/mai	15:30	QUEDA	
UNIDADE B	05:00	12/mai	12:00	IMPACTO	
UNIDADE B	10:00	16/mai	05:00	CORTE	
UNIDADE C	06:15	23/mai	14:30		QUEDA
UNIDADE C	04:00	26/mai	11:00	IMPACTO	
UNIDADE A	12:00	27/mai	06:40		PERFUROCORTANTE

UNIDADE C	03:00	01/jun	10:00	ESCORIAÇÃO	
UNIDADE C	02:00	04/jun	14:00		SANGUE
UNIDADE C	06:15	08/jun	13:30	QUEDA	
UNIDADE C	09:30	08/jun	15:30	CORTE	
UNIDADE B	00:00	11/jun	11:45		IMPACTO
UNIDADE A	00:00	15/jun	06:45		QUEDA
UNIDADE C	02:50	16/jun	12:30	CORTE	
UNIDADE A	02:50	19/jun	15:50		PERFUROCORTANTE
UNIDADE C	03:00	20/jun	10:00	QUEDA	
UNIDADE C	10:00	20/jun	05:00	QUEDA	
UNIDADE B	00:05	22/jun	07:05	ATRITO	
UNIDADE C	06:00	22/jun	13:00	QUEDA	
UNIDADE C	04:20	24/jun	11:20	CORTE	
UNIDADE A	04:00	25/jun	23:00		PERFUROCORTANTE
UNIDADE A	07:30	28/jun	13:30	CORTE	
UNIDADE A	02:10	29/jun	13:30	CORTE	
UNIDADE B	06:15	30/jun	12:30		IMPACTO
UNIDADE B	02:30	04/jul	10:30	IMPACTO	
UNIDADE C	02:30	06/jul	09:30	CORTE	
UNIDADE C	00:15	11/jul	07:15	CORTE	
UNIDADE A	00:00	11/jul	11:40	IMPACTO	
UNIDADE A	00:50	11/jul	13:50		PERFUROCORTANTE
UNIDADE A	05:00	11/jul	18:00	QUEDA	
UNIDADE A	00:00	13/jul	12:00		QUEDA
UNIDADE A	01:30	13/jul	14:30	ATRITO	
UNIDADE A	01:00	14/jul	14:00	IMPACTO	
UNIDADE A	01:30	15/jul	08:30	IMPACTO	
UNIDADE C	01:00	15/jul	09:00	PERFURAÇÃO	
UNIDADE B	01:45	19/jul	15:15	IMPACTO	
UNIDADE A	05:15	20/jul	17:15		PERFUROCORTANTE
UNIDADE C	01:40	26/jul	20:40		SANGUE
UNIDADE B	00:00	27/jul	14:45	QUEDA	
UNIDADE A	00:00	29/jul	06:30		QUEDA
UNIDADE A	04:30	29/jul	11:45	QUEIMADURA	
UNIDADE A	00:20	30/jul	19:05	IMPACTO	

UNIDADE A	04:00	01/ago	17:00	IMPACTO	
UNIDADE A	06:00	01/ago	13:35		APRISIONAMENTO
UNIDADE A	03:30	05/ago	10:30		PERFUROCORTANTE
UNIDADE B	06:15	08/ago	13:16	QUEDA	
UNIDADE C	02:12	10/ago	08:12		PERFUROCORTANTE
UNIDADE A	10:00	12/ago	06:50		SEM INFORMAÇÃO
UNIDADE A	06:30	12/ago	13:30	QUEDA	
UNIDADE A	02:05	12/ago	15:00		SEM INFORMAÇÃO
UNIDADE A	03:10	12/ago	16:10		SEM INFORMAÇÃO
UNIDADE B	03:00	15/ago	15:00	QUEDA	
UNIDADE A	08:10	16/ago	15:10	ATRITO	
UNIDADE C	10:15	17/ago	17:15		PERFUROCORTANTE
UNIDADE A	02:25	19/ago	09:25	ATRITO	
UNIDADE C	04:00	23/ago	22:55		PERFUROCORTANTE
UNIDADE C	02:10	25/ago	09:10		SANGUE
UNIDADE A	02:00	30/ago	08:00	QUEDA	
UNIDADE A	05:00	30/ago	00:00		SECREÇÃO TRAQUEAL SANGUE
UNIDADE C	08:50	31/ago	04:50		
UNIDADE C	08:00	05/set	04:00		PERFUROCORTANTE
UNIDADE B	05:30	05/set	12:30	QUEDA	
UNIDADE C	06:20	05/set	13:20	QUEDA	
UNIDADE C	00:00	05/set	09:40		QUEDA
UNIDADE A	08:00	15/set	15:00	ATRITO	
UNIDADE B	00:00	27/set	06:45		QUEDA
UNIDADE A	05:55	27/set	17:25	QUEDA	
UNIDADE A	01:15	07/out	08:15	CORTE	
UNIDADE B	07:00	11/out	13:00	IMPACTO	
UNIDADE A	04:45	13/out	10:45	IMPACTO	
UNIDADE A	12:00	20/out	18:20		SANGUE
UNIDADE A	02:00	20/out	09:00		SECREÇÃO TRAQUEAL
UNIDADE A	10:00	23/out	05:30		PERFUROCORTANTE
UNIDADE A	03:00	25/out	10:00	CONTUSÃO	
UNIDADE A	01:00	26/out	08:00	CONTUSÃO	
UNIDADE A	02:00	27/out	09:00		SANGUE
UNIDADE A	06:00	31/out	19:15	CONTUSÃO	

UNIDADE A	02:30	31/out	20:30	CORTE	
UNIDADE C	08:00	01/out	15:30	QUEDA	
UNIDADE A	06:00	03/nov	19:00		SECREÇÃO TRAQUEAL
UNIDADE A	00:00	07/nov	08:15		QUEDA
UNIDADE B	12:00	09/nov	07:20		QUEDA
UNIDADE C	08:00	10/nov	15:00	ESCORIAÇÃO	
UNIDADE A	02:15	10/nov	15:15	QUEDA	
UNIDADE B	02:30	11/nov	09:30	ATRITO	
UNIDADE C	01:20	16/nov	08:20	CORTE	
UNIDADE A	03:30	16/nov	10:30		SANGUE
UNIDADE C	06:00	16/nov	12:40		SEM INFORMAÇÃO
UNIDADE C	00:00	16/nov	12:20		IMPACTO
UNIDADE B	06:00	18/nov	13:25		QUEDA
UNIDADE C	00:00	21/nov	06:50	QUEIMADURA	
UNIDADE A	05:30	21/nov	12:30		PERFUROCORTANTE
UNIDADE A	05:00	24/nov	12:00	IMPACTO	
UNIDADE B	08:15	24/nov	15:15	ATRITO	
UNIDADE B	00:55	28/nov	08:55	ATRITO	
UNIDADE B	06:20	29/nov	13:20	ATRITO	
UNIDADE C	06:10	01/dez	12:10	QUEDA	
UNIDADE A	00:00	02/dez	12:45	IMPACTO	
UNIDADE A	04:00	07/dez	17:15	PERFURAÇÃO	
UNIDADE A	04:00	11/dez	10:00	QUEIMADURA	
UNIDADE B	12:00	11/dez	19:15		QUEDA
UNIDADE A	05:30	14/dez	18:30		PERFUROCORTANTE
UNIDADE A	10:00	18/dez	16:00	QUEIMADURA	
UNIDADE A	01:46	26/dez	09:00		SANGUE
UNIDADE A	04:00	27/dez	11:00		SANGUE
UNIDADE A	00:00	28/dez	12:10		IMPACTO
UNIDADE A	07:00	29/dez	14:00	CORTE	
UNIDADE A	03:45	29/dez	16:55		PERFUROCORTANTE
UNIDADE A	02:30	30/dez	09:30	QUEDA	



### ANEXO 3 - Lista de Comandos no R

```

#Entrando com os dados de frequência para a tabela de contingência
d <- c(47, 10, 31,
      23, 8, 2,
      31, 5, 12)

A <- matrix(d, nrow = 3, byrow=TRUE)

#Adicionando a legenda a tabela
dimnames(A) <- list("Unidade" = c("Unidade A", "Unidade B", "Unidade C"),
                  "Tipo de Acidente" = c("Típico", "Trajeto", "Biológico"))

#visualizando a tabela e suas marginais
addmargins(A)

#Fazendo o teste de qui-quadrado com p-valor exato na matriz A,
#pois um dos valores esperados da tabela é menor que 5
tt <- chisq.test(A, simulate.p.value = TRUE, B=10000)

#tabela de valores esperados
round(tt$expected, 1)

#Resultado do teste qui-quadrado
tt

#Teste de Hipótese, onde Ho: Independência e H1: Não há independência
ifelse(tt$p.value<0.05, "Rejeita Ho", "Não Rejeita Ho")

#tabela com os residuos padronizados
round(tt$stdres, 1)

#Criando um grafico de mosaico
library(vcd)
library("vcdExtra")
mosaic(A, shade=TRUE, legend=TRUE, residuals = tt$stdres)

#Fazendo o teste G2
G2 <- 2*sum(tt$observed*log(tt$observed/tt$expected))

#Resultado do teste
G2
#P-valor associado
pvalorG2<- pchisq(G2, df=4, lower.tail=FALSE)
pvalorG2

#Teste de Hipótese, onde Ho: Independência e H1: Não há independência
ifelse(pvalorG2<0.05, "Rejeita Ho", "Não Rejeita Ho")

```

```
#Construindo primeira sub-tabela
A1 <- A[1:2,1:2];
dimnames(A1) <- list(Unidade=c("Unidade A", "Unidade B"), Tipo=c("Típico",
"Trajeto"))
A1
t1 <- chisq.test(A1)
G2.1 <- 2*sum(t1$observed*log(t1$observed/t1$expected))
G2.1
p1 <- pchisq(G2.1, df=1, lower.tail=FALSE)
p1
```

```
#Teste de Hipótese, onde Ho: Independência e H1: Não há independência
ifelse(p1<0.05, "Rejeita Ho","Não Rejeita Ho")
```

```
#Construindo segunda sub-tabela
A2 <- cbind(apply(A[1:2,1:2], 1, sum), A[1:2, 3])
dimnames(A2) <- list(Unidade=c("Unidade A", "Unidade B"), Tipo=c("Típico +
Trajeto", "Biológico"))
A2
t2 <- chisq.test(A2)
G2.2 <- 2*sum(t2$observed*log(t2$observed/t2$expected))
G2.2
p2 <- pchisq(G2.2, df=1, lower.tail=FALSE)
p2
```

```
#Teste de Hipótese, onde Ho: Independência e H1: Não há independência
ifelse(p2<0.05, "Rejeita Ho","Não Rejeita Ho")
```

```
#Construindo terceira sub-tabela
A3 <- rbind(apply(A1, 2, sum), A[3, 1:2])
dimnames(A3) <- list(Unidade=c("Unidade A + Unidade B", "Unidade C"),
Tipo=c("Típico", "Trajeto"))
A3
t3 <- chisq.test(A3)
G2.3 <- 2*sum(t3$observed*log(t3$observed/t3$expected))
G2.3
p3 <- pchisq(G2.3, df=1, lower.tail=FALSE)
p3
```

```
#Teste de Hipótese, onde Ho: Independência e H1: Não há independência
ifelse(p3<0.05, "Rejeita Ho","Não Rejeita Ho")
```

```
#Construindo quarta sub-tabela
aux <- matrix(c(1, 1, 0, 0, 0, 1), nrow=2, byrow=TRUE)%%A[3,]
A4 <- rbind(apply(A2, 2, sum), t(aux))
dimnames(A4) <- list(Unidade=c("Unidade A + Unidade B", "Unidade C"),
Tipo=c("Típico + Trajeto", "Biológico"))
A4
t4 <- chisq.test(A4)
G2.4 <- 2*sum(t4$observed*log(t4$observed/t4$expected))
G2.4
```

```
p4 <- pchisq(G2.4, df=1, lower.tail=FALSE)
p4
#Teste de Hipótese, onde Ho: Independência e H1: Não há independência
ifelse(p4<0.05, "Rejeita Ho", "Não Rejeita Ho")

#Tabela com os valores do teste G2 aplicado a cada sub-tabela e seus valor-p
data.frame(Subtable=1:4, df=1, G2=c(G2.1, G2.2, G2.3, G2.4),
           P=c(p1, p2, p3, p4))

#Note que a soma dos testes G2 das sub-tabelas é igual ao teste G2 da tabela original
G2
G2.1+G2.2 + G2.3+ G2.4

#Etapas para análise de correspondência
#Pacote necessário para o gráfico de análise de correspondência
library(MASS)

#Análise de Correspondência
(cp <- corresp(A))
biplot(corresp(A, nf = 2), xlim=c(-0.6, 0.6), ylim=c(-0.2,0.2))
```

#### ANEXO 4 - Conceito básico de análise de correspondência

Vamos usar com base o material online de Fogo (2017) e dissertação mestrado de Costa (2016), que nós dará uma ótima introdução à análise de correspondência.

Se considerarmos uma tabela com  $l$  linhas e  $c$  colunas (tabela  $l \times c$ ), podemos imaginar cada valor de cada linha como sendo uma coordenada de  $l$  pontos num espaço  $c$ -dimensional, desta forma podemos calcular as distâncias entre esses pontos. Essas distâncias irão resumir as informações a respeito das similaridades entre as linhas da tabela.

"Suponha agora que encontremos um espaço com menos dimensões, sobre o qual possamos representar as posições dos  $l$  pontos de forma que toda, ou quase toda, informação das diferenças entre as linhas seja nele retida." (FOGO, 2017)

Podemos representar informações a respeito das similaridades entre as linhas num gráfico de 1, 2 ou 3 dimensões.

Para construir um gráfico de análise de correspondência, primeiro devemos obter os escores (coordenadas) associados a cada célula da tabela de contingência, e para obter os escores utilizaremos uma matriz de frequências relativas, conhecida como matriz de correspondência  $\mathbf{P}$ .

$$\mathbf{P} = p_{lc} = \frac{n_{lc}}{n}$$

em que  $p_{lc}$  representa uma proporção de cada frequência  $n_{lc}$  da tabela de contingência em relação ao total das frequências,  $n$ .

Sabemos que a diferença dos perfis linhas ou colunas aos seus respectivos centróides é contemplada por meio da distância  $X^2$ , pois a mesma satisfaz a propriedade da equivalência distribucional, com isso em mente, considere a matriz de frequências relativas  $\mathbf{P}$ , reproduz a matriz  $\mathbf{W}$ , definida por:

$$\mathbf{W} = \left[ \frac{p_{lc} - p_{l+}p_{+c}}{\sqrt{p_{l+}p_{+c}}} \right]$$

$$\text{em que } p_{l+} = \sum_{c=1}^C p_{lc} \text{ e } p_{+c} = \sum_{l=1}^L p_{lc}.$$

Após encontrada a matriz  $W$ , aplicaremos a decomposição dos valores singulares, para finalmente obter os escores de linhas ou colunas, de tal forma que a operação  $W'W$  corresponderá à matriz de covariância associada às linhas e a operação  $WW'$  corresponderá à matriz de covariância associada às colunas.

"O objetivo final da análise de correspondência é reduzir a dimensionalidade da nuvem de dados com o mínimo possível de perda de informação, ou seja, retendo o máximo possível da inércia (variabilidade)." (COSTA, 2016)

Inércia em análise de correspondência, é definida como uma medida de dispersão entre as variáveis da tabela dada pelo teste Qui-quadrado, dividido pelo total das frequências,  $n$ .

As coordenadas(escore) das linhas serão dadas por:

$$Y = D_r^{-1} W^t I$$

em que  $D_r$  é uma matriz diagonal dada por:

$$D_r = \begin{bmatrix} p_{1+} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & p_{L+} \end{bmatrix}$$

e  $I$  representa a normalização usada para distribuir a inércia pelas linhas, que será dada por:

$$I = D_r^{-\frac{1}{2}} U, \text{ em que } U \text{ é o autovetor de } W'W$$

De maneira análoga, as coordenadas(escore) das colunas serão dadas por:

$$Z = D_c^{-1} W^t J$$

em que  $D_c$  é uma matriz diagonal dada por:

$$D_c = \begin{bmatrix} p_{+1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & p_{+C} \end{bmatrix}$$

e  $J$  representa a normalização usada para distribuir a inércia pelas colunas, que será dada por:

$$J = D_c^{-2} V, \text{ em que } V \text{ é o autovetor de } \mathbf{W}\mathbf{W}'$$