

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
INSTITUTO DE CIÊNCIAS EXATAS  
DEPARTAMENTO DE ESTATÍSTICA  
PÓS-GRADUAÇÃO: ESPECIALIZAÇÃO EM ESTATÍSTICA APLICADA

ANÁLISE DE CORRESPONDÊNCIA APLICADA AO ESTUDO DA ASSOCIAÇÃO ENTRE O SEXO, ORDEM DE PARTO E TAMANHO DA PROLE EM COLÔNIAS DE RATOS WISTAR COM SUSCEPTIBILIDADE À CRISE AUDIOGÊNICA (ICB/UFMG) UTILIZADOS EM MODELOS DE PESQUISA EM EPILEPSIA

MAURA REGINA SILVA DA PÁSCOA VILELA

Monografia apresentada ao Programa de Pós-Graduação em Estatística, Instituto de Ciências Exatas, Universidade Federal de Minas Gerais como requisito parcial para obtenção do título de Especialista em Estatística.

Orientadora: Dra. Sueli Aparecida Mingoti

Co-Orientador: Dr. Márcio Flávio Dutra Moraes

Belo Horizonte/MG

2017

## AGRADECIMENTOS

À professora Dra. Sueli Aparecida Mingoti por introduzir-me aos conhecimentos da estatística multivariada, pela confiança e orientação. Sem os seus ensinamentos jamais teria realizado essa monografia.

Ao professor Dr. Márcio Flavio Dutra Moraes que mais uma vez abriu as portas do NCC e acreditou no meu projeto. Minha eterna gratidão.

Ao Dr. Noberto Garcia Cairasco e colaboradores por fazerem parte dessa equipe.

Ao corpo docente do curso de Especialização em Estatística Aplicada. Aos professores que contribuíram para minha formação: Edna Afonso Reis, Ela Mercedes M de Toscano, Gregório Saraiva Atuncar, Ilka Afonso Reis, Marcelo Azevedo Costa, Roberto da Costa Quinino e Sueli Aparecida Mingoti.

Ao monitor da Pós-Graduação Guilherme Lopes de Oliveira.

Às secretárias Rogéria Ferreira Figueiredo, Maria Cristina Morandi, Márcia de Souza Fileto, Rosiane Araújo Gonçalves e Ana Maria Rodrigues Maltez.

Às minhas amigas e companheiras de estudo: Ana Carolina de Souza Eustáquio e Kellen Roberta Rocha da Silva.

Aos professores e amigos do NNC. Em especial, à professora Juliana Carvalho Tavares e ao professor André Ricardo Masenssini que me acolhem nos momentos de tristeza.

Aos meus pais, irmãos, sogro, cunhadas(o) e sobrinhas(os). À Catharine, minha irmã querida, sempre presente na minha vida.

Aos meus filhos e a minha família agradeço pelo apoio e compreensão. Em especial ao Jefferson, alguém cujo conselho, senso de humor e crítica admiro tanto.

A todos que participaram direta ou indiretamente neste trabalho, expresso os meus sinceros agradecimentos.

## RESUMO

O objetivo desse trabalho foi utilizar a técnica de estatística multivariada denominada análise de correspondência simples para investigar como a ordem do parto interfere no tamanho da prole e na estratificação do sexo dos ratos com susceptibilidade à crise audiogênica. O rato WAR é um modelo de epilepsia reflexa, obtido por meio de consanguinidade de ratos Wistar sensíveis a estímulo de alta intensidade acústica. Trata-se um modelo que apresenta crises epiléticas recorrentes, que mimetiza a crise epilética desenvolvida em seres humanos. Foi elaborado um banco de dados do tipo relacional contendo todas as informações da colônia WAR/UFMG desde a sua implantação (1995) até os dias atuais. A primeira abordagem do banco de dados foi direcionada especificamente aos acasalamentos dos animais. O estudo analisou 809 acasalamentos e 6354 roedores com susceptibilidade à crise audiogênica, sendo 51,3% do sexo feminino e 48,7% do sexo masculino. A análise de correspondência mostrou relação de dependência para as variáveis envolvidas. Os partos de primeira e segunda ordem se associaram à prole composta por 10 a 15 filhotes, para ambos os sexos. Em contrapartida, observou-se associação fraca entre o parto de quarta ordem e as proles categorizadas como muita pequena (menos que 5 filhotes) e pequena (de 5 a 10 filhotes), para ambos os sexos. Outro resultado interessante foi associação, embora fraca, entre o parto de terceira ordem e prole com mais ou igual a 15 ratos, para ambos os sexos.

Palavras chave: Epilepsia, rato WAR, Instituto de Ciências Biológicas (ICB), Banco de dados, Análise de Correspondência.

## ABSTRACT

The aim of this study was to use the multivariate statistical technique called simple correspondence analysis to investigate how the number of crossing interferes in the size of the offspring and in the sex stratification of seizure susceptible rats. The Wistar audiogenic rat (WAR), is characterized as audiogenic reflex epilepsy animal model obtained by consanguinity of Wistar rats sensitive to high intensity acoustic stimuli. This model has spontaneously recurrent seizures which undoubtedly closer to human epilepsy. For this purpose, a relation database containing information about the WAR/UFGM strain has been developed since its implementation (1995) to the present day. The first approach to the database was directed specifically to animal mating. The study analyzed of 809 crossings and 6354 seizure susceptible rats, 51.3% female and 48,7% male. Correspondence analysis showed dependence relationship for the variables involved. The first and second births were associated with largest offspring number (from 10 to 15 rats), for both sexes. On the other hand, a weak association was observed between the fourth birth and the offspring categorized as very small (less than 5 rats) and small (5 to 10 rats), for both sexes. Another interesting result was an association, although weak, between third birth and then offspring with more than or equal 15 rats, for both sexes.

Keyword: Epilepsy, WAR rat, Institute of Biological Sciences (ICB), Database, Correspondence Analysis.

## LISTA DE ABREVIATURAS

CEUA	Comissão de Ética no Uso de Animais
FMRP	Faculdade de Medicina de Ribeirão Preto
ICB	Instituto de Ciências Biológicas
ICEx	Instituto de Ciências Exatas
NMDA	N-metil-D-aspartato
LNNE	Laboratório de Neurofisiologia e Neuroetologia Experimental
PEPS	Potencial excitatório pós-sináptico
NNC	Núcleo de Neurociências
SGBD	Sistema gerenciador de banco de dados
UFMG	Universidade Federal de Minas Gerais
USP	Universidade do Estado de São Paulo
WAR	Wistar Audiogenic Rat

## LISTA DE TABELA

Tabela 1	Banco de dados acasalamento
Tabela 2	Banco de dados identificação do animal
Tabela 3	Banco de dados referente ao comportamento da crise audiogênica
Tabela 4	Frequência das variáveis do banco acasalamento segundo a presença de registro
Tabela 5	Frequência do acasalamento em função tempo
Tabela 6	Distribuição da ordem de parto segundo o sexo
Tabela 7	Análise descritiva do número de filhote segundo o sexo e ordem de parto
Tabela 8	Distribuição do tamanho da prole
Tabela 9	Tabela de contingência
Tabela 10	Tabela de contingência das variáveis sexo, ordem de parto e tamanho da prole
Tabela 11	Matriz de correspondência das proporções para uma tabela de contingência
Tabela 12	Proporções para a matriz de correspondência
Tabela 13	Proporções para o total da linha
Tabela 14	Proporções para o total da coluna
Tabela 15	Frequências esperadas
Tabela 16	Desvio da frequência observada em relação à esperada
Tabela 17	Contribuição para a estatística Qui-Quadrado
Tabela 18	Inércias relativas para as linhas e colunas
Tabela 19	Inércia total
Tabela 20	Medidas de ajuste para o perfil de linha
Tabela 21	Medidas de ajuste para o perfil de coluna

## LISTA DE QUADROS

Quadro 1	Sequências das etapas do acasalamento
Quadro 2	Índice de gravidades das crises audiogênicas
Quadro 3	Classificação das variáveis
Quadro 4	Dados excluídos

## LISTA DE FIGURAS

- Figura 1      Distribuição dos acasalamentos desde a implantação até os dias atuais
- Figura 2      Distribuição da ordem de parto segundo a frequência do sexo
- Figura 3      Histograma do número de filhotes segundo o sexo e a ordem de parto
- Figura 4      Frequência dos filhotes em função da ordem de parto e do sexo
- Figura 5      Análise de correspondência para avaliar as relações entre o sexo, a ordem de parto e o tamanho da prole

## SUMÁRIO

1 INTRODUÇÃO .....	1
1.1. Epilepsia .....	1
1.2 Crise Audiogênica .....	2
1.3 Métodos de Estatística Multivariada .....	3
2 Justificativa .....	4
3 Objetivos .....	6
3.1 Objetivo Geral .....	6
3.2 Objetivos Específicos .....	6
4 COLÔNIA WAR .....	7
4.1 Local de Realização da Pesquisa .....	7
4.2 Amostra .....	7
4.3 Comitê de Ética .....	7
4.4 Acasalamento.....	7
4.5 Método de Avaliação Comportamental da Crise Audiogênica .....	8
5 BANCO DE DADOS .....	10
6 ANÁLISE DESCRITIVA.....	13
7 ANÁLISE DE CORRESPONDÊNCIA .....	20
7.1 Descrição da Metodologia de Análise de Correspondência .....	21
7.1.1. Tabela de Contingência .....	21
7.1.2. Matriz de Correspondência e Perfil de Linha e Coluna.....	23
7.1.3 Análise de Correspondência e Independência Entre as Variáveis.....	29
7.1.4 Relação com a estatística Qui-Quadrado de Pearson .....	29
7.1.5 Decomposição Em Coordenadas Principais .....	34
7.1.6 Medidas de Avaliação do Ajuste .....	36
7.1.7 Gráfico de Correspondência Simples .....	39

8 CONSIDERAÇÕES FINAIS .....	42
ANEXO .....	45
REFERÊNCIAS.....	49

## 1 INTRODUÇÃO

Nesta seção apresentamos uma breve revisão de literatura com a finalidade descrever aspectos sobre a epilepsia e a sua relação com a crise audiogênica e a estatística multivariada.

### 1.1. Epilepsia

A palavra epilepsia, de origem grega *epilambanien* significa ser invadido, dominado ou apossado. As explicações para os fenômenos epilépticos ao longo da história foram divididos em dois pensamentos, o senso comum e o pensamento científico. Na antiguidade, essa síndrome arraigada às crenças impunha ao paciente epiléptico a condição de um indivíduo acometido por forças malignas de um mundo sobrenatural, derivada de um castigo divino. Então, a epilepsia era vista como uma condição demoníaca cujo diagnóstico estava associado à magia negra, ciclo lunar, sexo e masturbação. Opondo-se a esse pensamento, em alguns registros médicos incluindo-se relatos hipocráticos, o termo epilepsia já era utilizado sem nenhuma conotação divina sendo uma patologia que acomete o encéfalo (LEES et al., 2000).

A epilepsia é uma síndrome encefálica caracterizada por predisposição persistente a gerar crises epiléticas (BLUME et al., 2001). As crises epiléticas são eventos clínicos que refletem disfunção temporária do encéfalo associados à ocorrência de sinais e ou sintomas decorrentes de uma atividade excessiva ou sincrônica de neurônios encefálicos (FISHER et al., 2005). A epilepsia pode ser expressa por mais de 40 tipos distintos de crises epiléticas (LEES et al., 2000) originadas por diferentes etiologias, anatomia, idade de início, gravidade, comportamento circadiano, prognóstico e características eletroencefalográficas. As crises convulsivas podem afetar a consciência, o estado emocional, a memória, a cognição, o comportamento, as funções sensoria, motora e autonômica. Nem todas as crises afetam esses fatores, mas todas influenciam em pelo menos em um deles (SCHEFFER et al., 2017; FISHER et al., 2005).

Essa síndrome afeta aproximadamente 1 a 2% da população mundial. Torna-se um importante problema de saúde pública cuja prevalência, em média, é de cinco casos a cada mil habitantes. Sua incidência é de cinquenta novos casos para uma população de cem mil indivíduos por ano. Estima-se que mais de cinquenta milhões de pessoas sejam afetadas no mundo independentemente da idade, gênero, raça e condição socioeconômica (PAHO, 2016). Estudos consideram que 60 a 80% das crises epiléticas sejam de natureza idiopática e que as

crises generalizadas ocorram em 40 a 60% das epilepsias (HAUSER et al., 1975; GRANIERI et al., 1983).

## 1.2 Crise Audiogênica

Os modelos de epilepsia em animais ainda são pré-requisitos importantes na busca pré-clínica por novos fármacos anticonvulsivantes. Sendo assim, uns dos modelos genéticos de grande importância para a compreensão dos mecanismos das epilepsias é o modelo da crise audiogênica em ratos. Nesse modelo, a crise é induzida por estímulo sonoro de alta intensidade. Poucas linhagens de ratos são capazes de exibir a crise audiogênica, entre elas destaca-se a colônia de ratos com susceptibilidade audiogênica denominada *Wistar Audiogenic Rat* (WAR) ou rato WAR (GARCIA-CAIRASCO et al., 2017 e 1983; ROSS et al., 2000).

Em ratos, a primeira observação de crise audiogênica foi realizada em 1924 por pesquisadores soviéticos que desenvolveram, a partir de ratos Wistar, a linhagem *Krushinsky-Molodkina* (POLETAEVA, 2017; KRUSHINSKY et al., 1970). Nos anos 50, outra linhagem derivada de ratos *Sprague-Dawley* foi desenvolvida na Universidade do Arizona e denominada como GEPR (JOBE, 1982; DAILEY et al., 1989). Além das linhagens soviética e americana, duas outras derivadas de ratos Wistar foram desenvolvidas. Uma delas ficou em Estrasburgo na França (KIESMANN et al., 1988) e a outra, denominada WAR, foi desenvolvida no Brasil, no laboratório de Neurofisiologia e Neuroetologia Experimental do Departamento de Fisiologia da Faculdade de Medicina de Ribeirão Preto (GARCIA-CAIRASCO et al., 2017 e 1983; DORETTO et al., 2003a). Um ramo da linhagem WAR é mantido no biotério do Departamento de Fisiologia e Biofísica do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais.

Os ratos Wistar com susceptibilidade audiogênica exibem crise tônica ou tônico-clônica em resposta à intensidade de estímulo sonoro de 125 dB (GARCIA-CAIRASCO et al., 1983). A progressão da crise audiogênica caracteriza-se por eventos procursivos (corrida, pulos e quedas atônicas) e eventos convulsivos (crise tônico-clônica). Independentemente do nível de gravidade obtido pela crise, a corrida é um componente sempre presente que pode ocorrer em um ou dois episódios distintos. O componente tônico, considerado a manifestação mais grave da crise, ocorre logo após a fase de corrida e caracteriza-se pela flexão do dorso, pescoço, membros anteriores, membros posteriores, acompanhado de espasmos por todo o corpo e movimento de balanço (ROSS et al., 2000). No final da crise, o animal apresenta

quadro denominado depressão pós-ictal, caracterizado por ausência de reflexo de endireitamento e de respostas a estímulos nociceptivos, refratariedade a nova crise (DORETTO, 2003a; GARCIA-CAIRASCO, 1992), vocalização (TERRA et al., 1992), alterações neuroendócrinas como o aumento da prolactina (DORETTO et al., 2003a) e de vasopressina (FONSECA et al., 2005).

Foi demonstrado que a crise audiogênica inicia-se com a ativação da via auditiva do tronco encefálico envolvendo o colículo inferior, camadas profundas do colículo superior, formação reticular, substância negra e a substância cinzenta periaquedutal (ROSSETTI et al., 2011; DORETTO et al., 2010; GARCIA-CAIRASCO, 2002; ROSS et al., 2000). A interação de fatores anatômicos, neuroquímicos e fisiológicos sugere que hiperexcitabilidade nos modelos audiogênicos está associada a uma transiente regulação para cima dos diversos tipos de receptores para o glutamato e com a redução da resposta inibitória mediada pelo GABA (GARCIA-CAIRASCO et al., 2017; FAINGOLD, 2004).

### 1.3 Métodos de Estatística Multivariada

As técnicas de estatísticas multivariadas permitem a síntese das informações contidas em dados multivariados. Os métodos de estatística multivariada se dividem em dois grupos (MINGOTI, 2007):

- o primeiro grupo compreende as técnicas exploratórias de sintetização da estrutura de variabilidade. Nesse grupo estão inseridas as técnicas de análise de componentes principais, análise fatorial, análise de correlações canônicas, análise de agrupamento, análise discriminante e a análise de correspondência.
- o segundo grupo contém as técnicas de inferências estatísticas. Fazem parte desse grupo os métodos de estimação de parâmetros, testes de hipóteses, análise de variância, análise de covariância e de regressões multivariadas.

Estudos envolvendo seres humanos mostram aplicabilidade das técnicas multivariadas para melhor compreensão das epilepsias entre elas destacam-se: análise de correspondência múltipla (AKIYAMA et al., 2010; KAMINSKA, et al., 1999), a análise de agrupamento (KIM et al., 2015) e a regressão logística (SCHEFFT et al., 2003).

## 2 JUSTIFICATIVA

Já existem revisões na literatura mostrando a importância de se avaliar o impacto da seleção genética sobre o desenvolvimento de diversos modelos de roedores que exibem crise audiogênica. Garcia-Cairasco (2017), Maxson (2017) e Ross et al. (2000) enfatizaram que após vários anos de seleção, a linhagem de ratos Sprague-Dowley se dividiu em duas subcolônias. Existem animais capazes de exibir crises graves (GEPR-9) e animais que exibem somente crise moderada (GEPR-3). Dentro do mesmo contexto, Maxson (2017) publicou uma revisão mostrando como as mutações espontâneas de genes alteraram o comportamento da crise audiogênica exibidas em camundongos DBA/2. Poletaeva et al. (2017) mostrou que após vários anos de seleção a linhagem Krushinsky–Molodkina é um modelo propenso as hemorragias encefálicas e enfatiza que essa linhagem pode ser utilizada como um modelo experimental para estudo das epilepsias como também para compreensão de distúrbios do fluxo sanguíneo encefálico.

O rato WAR é um modelo que mimetiza a crises epiléticas em humanos e por isso torna-se importante organizar as informações sobre essa colônia para melhor compreensão do seu desenvolvimento. Trata-se de uma colônia isogênica cujos acasalamentos são determinados pelo índice de gravidade e pela latência das crises. Além da susceptibilidade à crise convulsiva os ratos da colônia WAR apresentam alterações comportamentais, celulares, moleculares, eletroencefalográficas bem caracterizadas que potencializam a sua utilização como modelo genético de epilepsia em protocolos de crise aguda (tônico-clônico) e crônico (klindling). Recentemente, Garcia-Cairasco et al. (2017) publicou uma revisão mostrando que essa colônia, ao longo de 25 anos, já gerou 52 publicações (79% USP e 21% UFMG) para compreender os mecanismos associados às convulsões e enfatizam a importância de disponibilizar esses animais internacionalmente com a finalidade de compartilhar conhecimento e facilitar o planejamento e a execução de projetos multi-institucionais.

Estudos envolvendo seres humanos vêm mostrando a importância do armazenamento de informações no formato de banco de dados para compreensão dos mecanismos associados às epilepsias (SAJATOVIC et al., 2017; IHLE et al., 2012; SIADAT et al., 2005). Nos últimos anos, os avanços na tecnologia sugerem novas aplicações do sistema de banco de dados e de estatística multivariada. Entre os mais importantes estão os desenvolvimentos de *hardware* e *software* com pacotes computacionais amigáveis e abertos que permitem analisar grande volume de dados complexos com relativa facilidade. Entre as técnicas de estatística multivariada aplicada em banco de dados destacam-se as de caráter exploratório como a

análise de correspondência (ABDULHAY et al., 2017) e a análise de agrupamento (ABDULHAY et al., 2017; MARTINEZ-DEL-RINCON et al., 2017) e as técnicas direcionadas à inferência como a regressão multivariada (MAYNARD et al., 2017; ROLSTON et al., 2016). Doretto et al. (2003b) estudaram a relação entre a genética e epilepsia em ratos WAR por meio de técnica estatística de regressão múltipla e de análise de agrupamento com a finalidade de avaliar o comportamento da crise audiogênica. Porém, suas análises envolveram a colônia WAR em seus primórdios, ou seja, somente da terceira até a décima sétima geração.

Hoje a colônia WAR encontra-se na 56ª geração e há mais de duas décadas registros manuscritos dessas colônias são mantidos em biotérios de duas grandes instituições de ensino/pesquisa, a Faculdade de Medicina de Ribeirão Preto/USP e o Instituto de Ciências Biológicas/UFMG. Porém, ainda há necessidade de se avaliar o impacto da seleção genética dessas colônias em relação aos seus ancestrais bem com investigar a possibilidade do uso desses animais como modelo de estudo para outros tipos de comorbidades que não seja apenas a epilepsia. Dentro desse contexto, surgiu à ideia desenvolver um projeto colaborativo entre o Núcleo de Neurociências (NNC/UFMG), o Laboratório de Neurofisiologia e Neuroetologia (LNNE/USP) e o Departamento de Estatística do Instituto de Ciências Exatas (ICEx/UFMG) para construir um banco de dados de tamanho suficiente para um estudo comparativo de ratos que permiti avaliar quantitativamente ambos os ramos da colônia WAR. Sendo assim, ressalta-se a importância da aplicação de métodos estatísticos multivariados, entre elas, a técnica de análise de correspondência simples para investigar as consequências da seleção genética na colônia de ratos WAR.

### 3 OBJETIVOS

#### 3.1 Objetivo Geral

O objetivo geral desse trabalho é investigar as consequências da seleção genética na colônia de ratos Wistar com susceptibilidade à crise audiogênica por meio de técnicas de Estatística Multivariada.

#### 3.2 Objetivos Específicos

1. Elaborar banco de dados para extração de dados e entrada de novos utilizando as planilhas do Excel<sup>®</sup>.
2. Realizar análise descritiva dos dados sobre o acasalamento.
3. Estabelecer as possíveis relações entre as variáveis que determinam o desenvolvimento da prole por meio da técnica de análise de correspondência simples.

## 4 COLÔNIA WAR

### 4.1 Local de Realização da Pesquisa

A pesquisa foi realizada no Departamento de Estatística do Instituto de Ciências Exatas (ICEX) e no Núcleo de Neurociências (NNC) do Departamento de Fisiologia e Biofísica do Instituto de Ciências Biológicas (ICB), ambos pertencentes à Universidade Federal de Minas Gerais (UFMG).

### 4.2 Amostra

A amostra foi composta pelo registro 809 acasalamentos referentes a 6354 ratos da colônia WAR criados e mantidos no biotério do Departamento de Fisiologia e Biofísica ICB/UFMG.

### 4.3 Comitê de Ética

O estudo visou à construção de banco de dados e análise estatística de dados já coletados. Não houve intervenção experimental. Logo, não houve a necessidade do parecer da Comissão de Ética no Uso de Animais (CEUA) nesta monografia.

### 4.4 Acasalamento

O acasalamento consiste na união consanguínea entre um macho com até três fêmeas da mesma ninhada, ambos os sexos com índice de gravidade  $\geq 0,85$ . A sequência dos eventos que compõe o acasalamento está descrita no Quadro 1.

Quadro 1: Sequências das etapas do acasalamento

Eventos	Duração
Período de união	5 dias
Período gestacional	$\approx 21$ dias
Nascimento	Registro do número de filhotes
Desmame	30 dias após o nascimento/Marcação
Crescimento	40 dias após o desmame
Avaliação comportamental da crise	Aos 70, 74 e 78 dias de nascimento.

#### 4.5 Método de Avaliação Comportamental da Crise Audiogênica

Aos setenta, setenta e quatro e setenta e oito dias todos os ratos WAR receberam três estímulos sonoros de alta intensidade (120 dB, SPL - nível de pressão sonora) com finalidade de selecionar animais com predisposição às crises epiléticas. Essa técnica de seleção é denominada *screening*. Os animais foram submetidos ao estímulo sonoro dentro de um cilindro de acrílico transparente de 32 cm de altura, com base de 30 cm de diâmetro, feito de modo a não interferir na expressão motora da crise audiogênica e ao mesmo tempo facilitar sua visualização. Este cilindro de acrílico fica posicionado dentro de uma câmara de 45 x 45 x 40 cm, sendo esta revestida nas paredes internas por material isolante acústico. O acesso visual é possível em toda parte frontal da câmara (MORAES et al., 2000).

O sistema de estimulação sonora da caixa consiste de um amplificador de 15 watts, um alto falante (tweeter - 4Ω, TLX1e, 5KHz-20KHz) e uma fonte de alimentação (12V, 3A). Neste sistema se aplica o sinal de uma campainha elétrica, gravado em CD e executado por um CD player. A intensidade do estímulo é ajustada por meio do controle de volume áudio, sendo que se colocado no máximo da escala, a intensidade sonora equivale a 120 dB SPL. O alto-falante, instalado no teto do cilindro de acrílico, está conectado ao amplificador por meio de uma chave que possui a finalidade de iniciar e interromper o som abruptamente quando necessário (MORAES et al., 2000).

Os animais foram colocados, individualmente, dentro da caixa de estimulação e observados durante o período de 1 minuto antes da aplicação do estímulo sonoro. Em seguida, um som de intensidade máxima foi aplicado até o aparecimento do componente tônico da crise. Na ausência do componente tônico, o som foi aplicado durante um período máximo de um minuto. O comportamento foi avaliado por meio de observações e quantificado pelo índice de gravidade das crises (IS) proposto por Garcia-Cairasco et al. (1996) e modificada por Rossetti et al. (2006). O IS representa uma escala de gravidade gradual, na qual a sequência de eventos característicos das crises são categorizados e expressos em valores ordinais de 0 a 1 (Quadro 2).

Quadro 2: Índice de gravidade das crises audiogênicas

Índice	Comportamentos
0,00	Ausência de qualquer comportamento relacionado à crise
0,11	Um episódio de corrida
0,23	Um episódio de corrida + pulos e quedas atônicas
0,38	Dois episódios de corrida + pulos e quedas atônicas
0,61	Todos os comportamentos anteriores + convulsão tônica com arqueamento dorsal da cabeça
0,73	Todos os comportamentos anteriores + convulsões clônicas parciais e generalizadas
0,85	Todos comportamentos anteriores + espasmos clônicos
0,90	Todos os comportamentos anteriores + flexão ventral da cabeça
0,95	Todos os comportamentos anteriores + hiperextensão dos membros anteriores
1,00	Todos os comportamentos anteriores + hiperextensão de membros posteriores

Fonte: GARCIA-CAIRASCO et al., 1996; ROSSETTI et al., 2006.

## 5 BANCO DE DADOS

A construção de um banco de dados é um processo de armazenar dados. Banco de dados computadorizado pode ser criado e mantido por um grupo de programas de aplicação escritos especificamente para essa tarefa ou por um sistema gerenciador de banco de dados (SGBD). O SGBD é um sistema de *software* de uso geral que facilita o processo de definição, construção, manipulação e tratamento de bancos de dados entre os diversos usuários. Banco de dados digital é um componente essencial na sociedade moderna. Atualmente, um dos formatos mais populares é o banco de dados “*MySQL/Oracle database*”. Trata-se de um sistema de gerenciamento de banco de dados (SGBD) que utiliza SQL (linguagem de consulta estruturada, do inglês *Structured Query Language*) como interface. Esse formato permite que sejam cruzados os dados entre as planilhas de modo relacional (ELMASRI e NAVATHE, 2011).

Hoje, grandes centros de pesquisa ligados às universidades por todo mundo estão fazendo amplo uso desse modelo de banco de dados para criar e gerenciar conhecimento eficaz. Estudos envolvendo seres humanos vêm mostrando a importância do armazenamento de informações em banco de dados (SAJATOVIC et al., 2017), principalmente no formato “*MySQL/Oracle database*” para compreensão dos mecanismos associados às epilepsias (IHLE et al., 2012; SIADAT et al., 2005).

### 5.1 Elaboração do Banco de Dados

Dentro dessa proposta foi elaborado um banco com os dados reais referentes à colônia WAR. Foram inseridos dados desde implantação (1995) da colônia até os dias atuais (2017). Os dados são atualizados constantemente. Os dados foram inseridos na planilha eletrônica do *Microsoft Office Excel 2010*<sup>®</sup>. Nessa planilha foram criadas partições contendo tabelas de domínios que relacionam os atributos descritos entre si. O banco de dados está dividido em dois blocos. Um contendo dados sobre os animais desenvolvidos no biotério do Departamento de Fisiologia e Biofísica (ICB/UFMG) e outro, contendo dados sobre animais provenientes do Laboratório de Neurofisiologia e Neuroetologia Experimental do Departamento de Fisiologia da Faculdade de Medicina de Ribeirão Preto (FMRP/USP). As planilhas elaboradas estão descritas a seguir:

- no banco de dados “origem do animal” foi inserida os códigos (1) e (2) para designar os dados oriundos do Laboratório de Neurofisiologia e Neuroetologia Experimental do

Departamento de Fisiologia da Faculdade de Medicina de Ribeirão Preto (FMRP/USP) e os dados provenientes do Departamento de Fisiologia e Biofísica (ICB/UFMG), respectivamente.

- no banco de dados “acasalamento” foi inserido os dados das variáveis ninhada, progenitores, ordem do parto (sequência de nascimento dos filhotes oriundos da mesma progenitora), datas (acasalamento, separação e desmame), tamanhos da prole e a frequência de filhotes vivos (Tabela 1).

Tabela 1: Banco de dados acasalamento

Ninhada	Progenitora	Progenitor	Ordem Parto	Acasalamento	Separação	Desmame	Prole	Vivos
1	...	...	...	...	...	...	...	...
507	3635	3640	1	27/12/2006	11/01/2007	09/03/2007	9	8
511	3635	3640	2	30/03/2007	19/04/2007/	29/05/2007	3	3
528	3635	3640	3	29/05/2007	19/06/2007	24/07/2007	8	8
n	...	...	...	...	...	...	...	...

- no banco de dados “identificação do animal” foi inserida as informações sobre o registro do animal e da ninhada, a data de nascimento e o sexo. Foram inseridos os códigos (1) e (2) para designar os sexe masculinos fêmeas e machos, respectivamente (Tabela 2).

Tabela 2: Banco de dados identificação do animal

Registro	Ninhada	Data de nascimento	Sexo
1	...	...	...
3854	507	30/01/2007	1
3855	507	30/01/2007	1
3856	507	30/01/2007	1
3857	507	30/01/2007	1
3858	507	30/01/2007	1
3859	507	30/01/2007	1
3860	507	30/01/2007	1
3861	507	30/01/2007	2
n	...	...	...

- no banco de dados “comportamento da crise audiogênica” foi inserido os registros dos animais, as datas dos estímulos sonoros, o índice de gravidade da crise, a frequência das crises (F), as latências das primeiras e segundas corridas e o tempo inicial das crises designado como latência CVT. Todas as latências foram expressas em segundos (Tabela 3).

Tabela 3: Banco de dados referente ao comportamento da crise audiogênica

Ratos	Datas dos estímulos sonoros(ES)			Índice de gravidade			F	Latências(C1)			Latência (C2)			Latência CVT		
	ES (1)	ES(2)	ES(3)	(1)	(2)	(3)		(1)	(2)	(3)	(1)	(2)	(3)	(1)	(1)	(3)
I	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3854	10/04/2007	12/04/2007	16/04/2007	0,85	0,85	0,85	3	4	5	4			22	11	12	34
3855	10/04/2007	12/04/2007	16/04/2007	0,85	0,85	0,85	3	3	3	4				9	10	11
3856	10/04/2007	12/04/2007	16/04/2007	0,85	0,90	0,95	3	2	5	4				9	12	11
3857	10/04/2007	12/04/2007	16/04/2007	0,85	0,95	0,90	3	2	2	3	30			6	6	8
3858	10/04/2007	12/04/2007	16/04/2007	0,85	0,85	0,95	3	3	6	4				35	14	10
3859	10/04/2007	12/04/2007	16/04/2007	0,85	0,85	0,85	3	3	4	4		23		9	10	33
3860	10/04/2007	12/04/2007	16/04/2007	0,85	0,11	0,85	2	3	2	3		28		46	2	41
3861	10/04/2007	12/04/2007	16/04/2007	0,85	0,85	0,85	3	3	5	4	14	25		48	60	13
..n	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

F: Total de frequência do componente tônico-clônico

- no banco de dados “destino dos animais” foi inserido a identificação e o destino do animal. Foram inseridos os códigos (1) para acasalamento, (2) experimento, (3) descarte, (3) morte após estimulação (4) sacrifício devido ao desenvolvimento de patologia e (6) omissões.

-no banco dados referente aos “dados metabólicos” foram inseridos a identificação do animal, a data e o registro do peso corporal (g).

A construção de um banco de dados é uma tarefa árdua que exige muita paciência e conhecimento prévio sobre as variáveis que serão inseridas. De fato, a possibilidade de contato com a colônia durante o período de 2004 até 2010 colaborou para a prevenção e o controle dos potenciais vieses.

Foram 12 meses de trabalhos até o resultado final do banco de dados. Os primeiros três meses foram dedicados à parte relacional do modelo, ou seja, como as planilhas do Excel<sup>®</sup> poderiam relacionar entre si. O restante foi dedicado ao levantamento, à obtenção dos dados em registros manuais e a sua transformação em dados digitais.

Fatores como padronização dos acasalamentos, perda de registros, localização dos dados, omissão de registro, inversão de registro, duplicação ou ausência dos registros foram os principais obstáculos que dificultaram o processo de construção do banco digital. Por outro lado, a prévia estruturação do modelo relacional e a definição das variáveis que seriam inseridas no banco de dados facilitaram a extração e a interpretação dos registros manuais.

Portanto, o objetivo da organização desses registros no formato de banco de dados foi para facilitar para acessibilidade às informações aos pesquisadores do Núcleo de Neurociências (NNC/UFMG), assim como para outros grupos de pesquisa do Departamento de Fisiologia e Biofísica e Centros Institucionais Colaboradores (LNNE/USP). Além disso, esse banco de dados servirá de modelo para entrada de novos dados e o desenvolvimento de pesquisas futuras.

## 6 ANÁLISE DESCRITIVA

A análise descritiva foi utilizada com o objetivo de resumir e compreender as características importantes do conjunto de dados referentes aos bancos de dados “acasalamento” e “identificação dos animais”. Foram avaliadas as medidas de tendência central (média e mediana), dispersão, distribuição, valores discrepantes (*outliers*) e a frequência dos dados em relação ao tempo. A análise descritiva foi realizada utilizando-se o *software* estatístico Minitab *for Windows*<sup>®</sup>, versão 14.1, 2003.

As variáveis do banco de dados “acasalamento” e “identificação dos animais” foram divididas em quantitativas (discretas e contínuas) e variáveis qualitativas (nominais e ordinais) descritas no Quadro 3. A variável discreta “frequência da crise” está relacionada à quantidade de crise com índice  $\geq 0,61$  após as três estimulações sonoras. A variável latência, expressa em segundos, está relacionada ao tempo de surgimento da crise audiogênica. A variável ordem de parto está relacionada com a sequência de nascimento dos filhotes oriundos da mesma progenitora.

Quadro 3: Classificação das variáveis

Variáveis Quantitativas		Variáveis Qualitativas	
Discreta	Contínua	Dicotômica	Ordinal
Frequência da crise	Peso corporal (g)	Sexo	Índice de gravidade
Número de filhotes	Latência (s)	Presença da crise	Ordem de Parto
Frequência de mortos			
Idade (dias)			

Atualmente, existem registros de 812 ninhadas referentes a 6364 ratos no ICB/UFMG e 971 ninhadas referentes 7910 na FMRP/USP. Os resultados analisados dessa monografia referem-se apenas ao banco de dados da colônia do Departamento de Fisiologia e Biofísica (ICB/UFMG) com ênfase ao banco de dados “acasalamento” e “identificação dos animais”.

Devido à “omissão de dados, foram excluídos três acasalamentos do “banco acasalamento” e dez animais do “banco identificação do animal”. Os dados excluídos estão descritos no Quadro 4. Após a exclusão, a amostra foi composta por 809 acasalamentos e 6354 animais. Em ambos os bancos a exclusão foi inferior a 1%.

Quadro 4: Dados excluídos

Banco acasalamento	Banco identificação do animal	
Registro do Acasalamento	Registro do animal	Registro da ninhada
19	36 - 43	Sem registro
20	4457	588
166	6127	773

A Tabela 4 mostra a frequência da variável ordem de parto, as datas de acasalamento, separação e desmame que apresentaram ou não seus registros. No “banco acasalamento” as omissões dos registros se concentraram para as datas de separação e desmame dos animais.

Tabela 4: Frequência das variáveis do banco acasalamento segundo a presença de registro

Variável	Frequência com registro (%)	Frequência sem registro (%)
Ordem de Parto	697 (86,2)	112 (13,8)
Data do acasalamento	699 (86,4)	110 (13,6)
Data da separação	322 (39,8)	487 (60,2)
Data do desmame	457 (56,6%)	351 (43,4%)

A frequência dos acasalamentos com a presença de registro ao longo tempo está descrito na Tabela 5 e na Figura 1. Observa-se maior frequência no ano de 2006 e a menor no ano de 2011. Provavelmente, a grande amplitude entre os anos de 2006 e 2011 esteja associada à demanda de animais solicitada pelos alunos de Iniciação Científica e Pós-Graduação para seus experimentos. Para o ano de 1995, a baixa frequência está associada ao período de implantação da colônia WAR no ICB/UFMG. Desde a sua implantação, colônia apresentou duas supervisões. A primeira sob a responsabilidade da professora Dra. Maria Carolina Doretto e atualmente sob a responsabilidade do professor Dr. Márcio Flávio Dutra Moraes. Ao longo desses anos a colônia foi administrada por cinco pessoas. A primeira administração compreendeu os anos 1995 - 2002, a segunda 2003 – 2009, a terceira 2010 – 2014, a quarta 2014 – 2015 e a última a partir do ano de 2016. Observa-se que durante o período de transição (2010/2011 e 2015/2016) também houve queda na frequência dos acasalamentos.

Tabela 5: Frequência do acasalamento em função do tempo

Ano	Frequência	Ano	Frequência
1995	6	2006	81
1996	18	2007	60
1997	31	2008	35
1998	41	2009	21
1999	22	2010	25
2000	37	2011	4
2001	38	2012	9
2002	24	2013	18
2003	47	2014	36
2004	35	2015	37
2005	47	2016	20
Total	346	Total	353

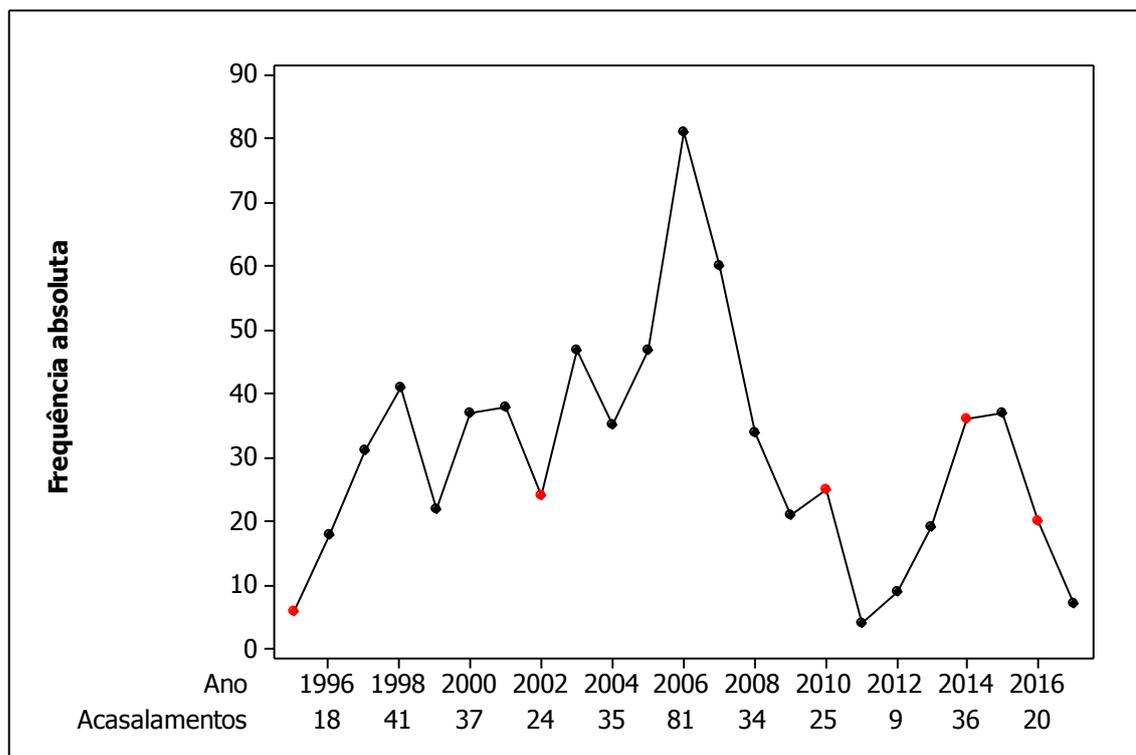


FIGURA 1: Distribuição dos acasalamentos desde a implantação até os dias atuais. Os pontos em vermelhos representam os períodos de transição entre os cuidadores da colônia.

Em função da ordem de parto os animais foram estratificados em fêmeas e machos. Observa-se que 3093 (48,7%) foram registrados como fêmeas e 3261 (51,3%) como machos. Nota-se uma similaridade entre os nascimentos de fêmeas e machos nessa colônia WAR. A Tabela 6 e a Figura 2 mostram a distribuição da ordem de parto segundo o sexo. A maioria

dos acasalamentos se concentrou no parto de ordem 1 (41,9%) seguido do parto de ordem 2 (27,2%), para ambos os sexos. O parto de ordem 1 gerou 1260 (19,8%) fêmeas e 1407 (22,1%) machos. O parto de ordem 2 gerou 864 (13,6%) fêmeas e 866 (13,6%) machos. O menor percentual foi observado para o parto de ordem 4 (4,2%) sendo 122 (1,9%) fêmeas e 146 (2,3%) machos.

Tabela 6: Distribuição da ordem de parto segundo sexo

Ordem do parto	Fêmeas (%)	Machos (%)
1	1260 (19,8)	1407 (22,1)
2	864 (13,6)	866 (13,6)
3	370 (5,8)	400 (6,3)
4	122 (1,9)	146 (2,3)
Indefinido	477 (7,5)	442 (7,0)

Indefinido: acasalamentos com omissão da ordem de parto.

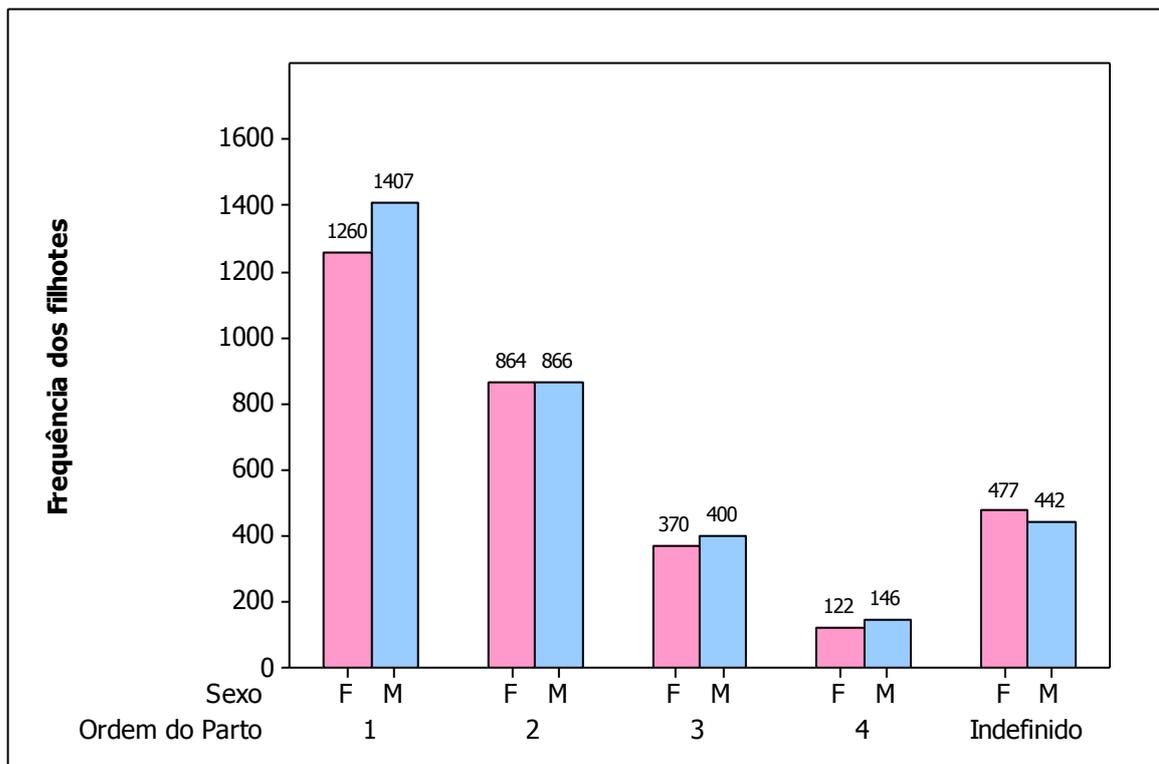


Figura 2: Distribuição da ordem de parto segundo a frequência do sexo. Indefinido: acasalamento com omissão da ordem de parto.

A Tabela 7 mostra a análise descritiva do número de filhotes segundo o sexo e a ordem de parto. Apresentam-se os valores absolutos e relativos, a média seguida do desvio padrão, a mediana seguida do primeiro e terceiro quartis e valores máximo e mínimo. Para ambos os sexos, observa-se similaridade entre os valores da estatística-resumo do número de filhotes em relação à ordem de parto. Observa-se que termos medianos as proles foram compostas por 11 filhotes, para a maioria dos partos. Os menores valores para a estatística-

resumo foi para o parto de ordem 4, para ambos os sexos. A menor prole foi de 1 filhote (categoria fêmea, parto de ordem 3) e a maior prole foi de 18 filhotes (categoria macho, ordem de parto 2). A saída dos resultados da análise descritiva por meio computacional encontra-se em Anexo.

Tabela 7: Análise descritiva do número de filhotes segundo o sexo e a ordem de parto

Variável	N	%	Média	SD	Min	Q <sub>1</sub>	Mediana	Q <sub>3</sub>	Max
Fêmea Parto 1	1260	19,8	10,3	2,9	2	8	11	13	16
Fêmea Parto 2	864	13,6	10,6	3,0	2	9	11	13	18
Fêmea Parto 3	370	5,8	10,1	3,2	1	8	10	12	17
Fêmea Parto 4	122	1,9	8,1	2,2	2	7	9	10	11
Fêmea P Indefinido	477	7,5	10,3	3,1	2	8	11	13	16
Macho Parto 1	1407	22,1	10,4	3,0	2	8	11	13	16
Macho Parto 2	866	13,6	10,4	3,0	2	9	11	13	18
Macho Parto 3	400	6,3	9,8	3,1	2	8	10	12	17
Macho Parto 4	146	2,3	8,3	2,0	2	7	9	10	12
Macho P Indefinido	442	7,0	10,3	3,0	2	8	11	12	16

Indefinido: acasalamentos com omissão da ordem de parto.

A Figura 3 mostra a distribuição número de filhotes segundo o sexo e a ordem de parto. A distribuição apresentou média seguida do desvio padrão de  $10,221 \pm 2,977$  filhotes e valor mediano igual a 11 filhotes seguido do primeiro quartil ( $P_{25} = 8$  filhotes) e do terceiro quartil ( $P_{75} = 12$  filhotes). A distribuição apresentou simetria negativa. Isso indica que ela é assimétrica à esquerda com valor mediano superior ao da média.

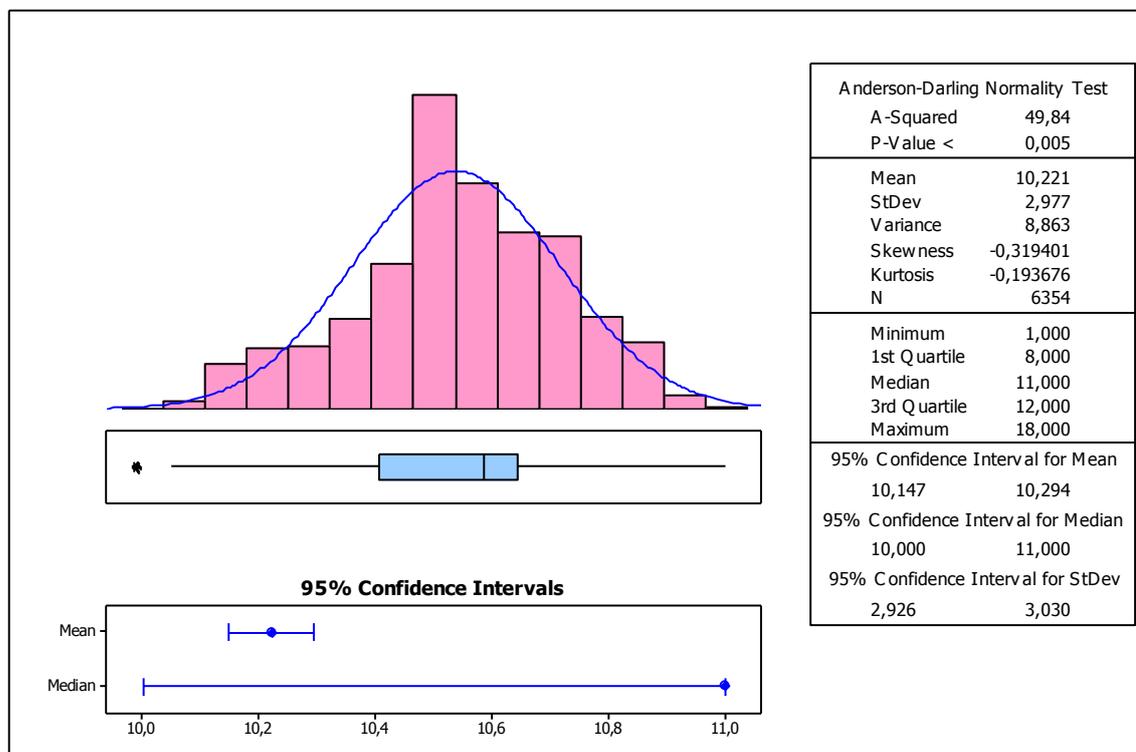


Figura 3: Histograma do número de filhotes.

A Figura 4 mostra a distribuição de filhotes em função da ordem de parto e do sexo. De acordo com o *Box-plot*, para cada ordem de parto foi possível visualizar os alinhamentos das medidas de tendência central (média e mediana) em função do sexo e da frequência de animais. O alinhamento sugere que há uma similaridade entre fêmeas e machos para mesma categoria de parto. O menor valor da média e da mediana foi para o parto de ordem 4, para ambos os sexos. A distribuição apresenta pontos fora da curva (*outliers*). Isso aconteceu para ambos os sexos, nos partos de ordem 2 e 4 e somente para a categoria fêmeas, parto de ordem 3.

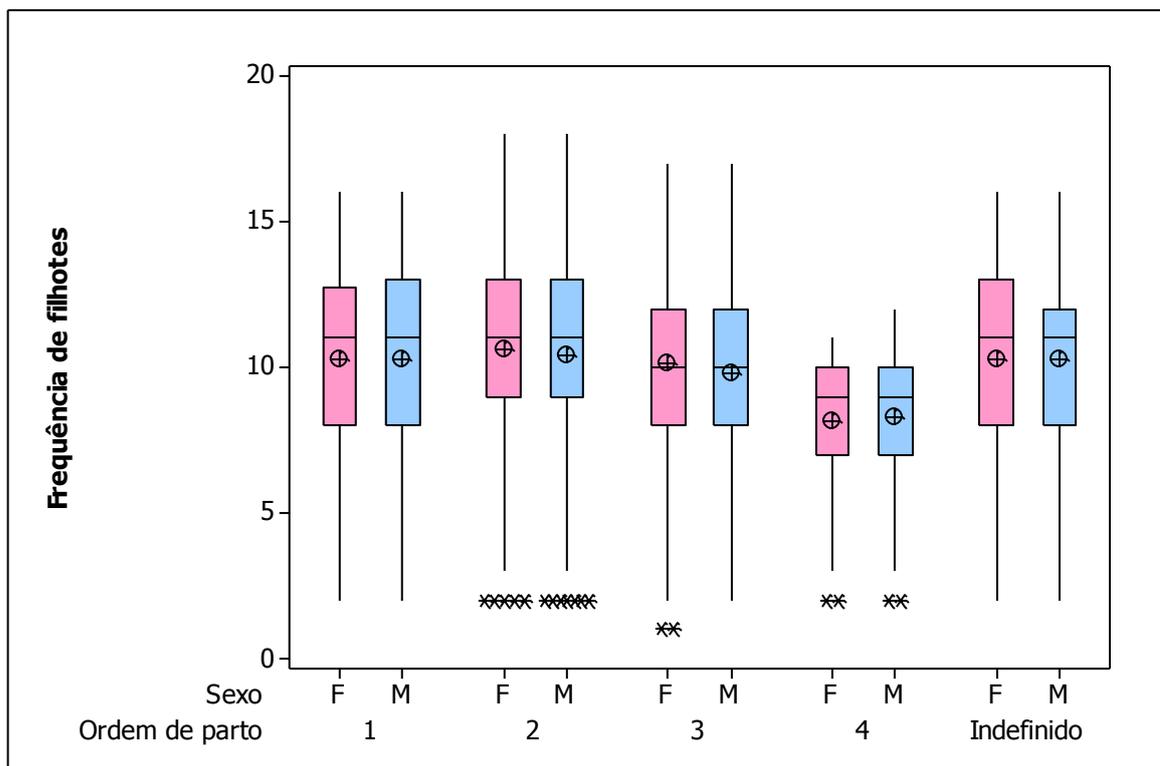


Figura 4: Frequência dos filhotes em função da ordem do parto e do sexo. Indefinido: acasalamento com omissão da ordem de parto.

A variável tamanho da prole foi obtida por meio da estratificação do número de filhotes como prole muito pequena ( $< 5$  filhotes), prole pequena ( $5 \leq 10$ ), prole grande ( $10 \leq 15$ ) e prole muito grande ( $\geq 15$  filhotes). O intervalo de classe foi definido pela regra de Sturge (VIEIRA, 2008):

$$k = 1 + 3,222(\log m)$$

$$k = 1 + 3,222(18)$$

$$k = 5,044$$

onde:  $k$  é o número de classes e  $m$  é maior frequência de animais observada na prole.

A Tabela 8 mostra a distribuição do tamanho da prole. Maiores frequências se concentraram nas proles categorizadas como prole de tamanho pequeno e grande.

Tabela 8: Distribuição do tamanho da prole

Tamanho da prole	Número de filhotes	Frequência	(%)
Muito pequena	$< 5$	233	3,7
Pequena	$5 \text{ † } 10$	2121	33,4
Grande	$10 \text{ † } 15$	3644	57,3
Muito Grande	$\geq 15$	356	5,6

## 7 ANÁLISE DE CORRESPONDÊNCIA

No início da década de 1960, um grupo de pesquisadores franceses, liderado pelo filósofo Jean-Paul Benzécri, desenvolveu técnicas para estruturar e interpretar conjuntos de dados complexos. Uma dessas técnicas visava transformar uma tabela de dados composta por contagens de frequências em um mapa visual que exibe linhas e colunas de tabela em relação às subjacentes contínuas. A contribuição de Benzécri para a análise de dados em geral e para a técnica de análise de correspondência em particular não foi tanto para a teoria matemática, mas sim para a atenção dada à interpretação gráfica dos resultados e pela ampla aplicabilidade dos métodos aos dados em muitos contextos. Seu interesse inicial era analisar grandes matrizes contendo contagens de palavras. Mas logo se percebeu o poder da técnica em campos tão diversos como a biologia, a arqueologia, a física e a música. Este grupo consolidou-se em Paris nos anos 1970 com pensamentos influentes e controversos depois da França de 1968. Em 1973, eles publicaram duas importantes bibliografias: a primeira “*L’Analyses de Données*” cujo o contexto se direcionava para classificação ou *cluster* e a segunda “*L’Analyses des Correspondances*” ou análise de correspondência (GREENACRE, 2015).

A análise de correspondência é um método de estatística multivariada, de caráter exploratório, utilizado para analisar dados categóricos. Esta técnica visa verificar se há uma relação entre as linhas e as colunas de tabelas de contingência (GREENACRE, 2015, 1988 e 1984; ALBERTI, 2015; MINGOTI, 2007; JOBSON, 1992) sendo empregada em diversas áreas do conhecimento (ALBERTI, 2013, BETH et al., 2011; FREUDENTHAL, 2009; BLANCO, 2007; GREENACRE, 2002).

A técnica de Análise de Correspondência Simples foi aplicada nesta monografia para compreender como as relações entre o sexo e a ordem de parto interfere no tamanho da prole. A amostra analisada foi composta de 6354 animais da colônia WAR (ICB/UFMG) oriundos de 809 acasalamentos. A análise foi realizada utilizando-se o *software* estatístico Minitab *for Windows*<sup>®</sup>, versão 14.1, 2003.

## 7.1 Descrição da Metodologia de Análise de Correspondência

Para facilitar o entendimento do leitor, nesta seção apresentamos alguns detalhes relativos à metodologia de Análise de Correspondência Simples com aplicação direta aos dados que estão sendo analisados nessa monografia. A saída dos resultados da Análise de Correspondência por meio computacional encontra-se em Anexo.

### 7.1.1. Tabela de Contingência

A distribuição bivariada dos dados categóricos é expressa por meio de tabela de contingência do tipo ( $r \times c$ ). Trata-se de uma tabulação cruzada a qual a frequência das observações é simultaneamente cruzada em relação às duas variáveis aleatórias categóricas, sendo a variável  $X$  utilizada para categorizar as linhas ( $r$ ) e a variável  $Y$  utilizada para categorizar a coluna ( $c$ ).

Genericamente, uma tabela de contingência entre duas variáveis  $X$  e  $Y$  com  $r$  linhas e  $c$  colunas é dada como mostrado na Tabela 9, sendo  $n_{ij}$ ,  $i = 1, 2, \dots, r$ ,  $j = 1, 2, \dots, c$ , a frequência pertencente à categoria  $i$  da variável  $X$  e à categoria  $j$  da variável  $Y$ .

Tabela 9: Tabela de Contingência

		Variável Y				Total
		1	2	...	c	
Variável X	1	$n_{11}$	$n_{12}$	...	$n_{1c}$	$n_{1\cdot}$
	2	$n_{21}$	$n_{22}$	...	$n_{2c}$	$n_{2\cdot}$
	3	$n_{31}$	$n_{32}$	...	$n_{3c}$	$n_{3\cdot}$
	.	.	.	...	.	.
	.	.	.	...	.	.
	r	$n_{r1}$	$n_{r2}$	...	$n_{rc}$	$n_{r\cdot}$
Total		$n_{\cdot 1}$	$n_{\cdot 2}$	...	$n_{\cdot c}$	$n_{\cdot\cdot} = N$

Na Tabela 10 apresenta-se a tabela de contingência relativa aos dados tratados nessa monografia. As linhas ( $r$ ) são descritas como a combinação dos dados sexo e ordem de parto. O sexo foi estratificado em fêmeas e machos. O parto foi categorizado em parto de primeira, segunda, terceira, quarta ordem e em parto de ordem indefinida. As colunas ( $c$ ) foram categorizadas pela variável tamanho da prole estratificado como prole muito pequena ( $<$

5 filhotes), prole pequena (5 | 10), prole grande (10 | 15) e prole muito grande ( $\geq 15$  filhotes).

Tabela 10: Tabela de contingência das variáveis sexo, ordem de parto e tamanho da prole

Sexo	Ordem do Parto	Tamanho da Prole				Total
		< 5	5   10	10   15	$\geq 15$	
Fêmeas	1	51	384	772	53	1260
	2	29	241	550	44	864
	3	17	144	165	44	370
	4	8	76	38	0	122
	Indefinido	13	159	274	31	477
Machos	1	47	457	834	69	1407
	2	36	247	554	29	866
	3	14	180	161	45	400
	4	6	95	45	0	146
	Indefinido	12	138	251	41	442
Total		233	2121	3644	356	<b>6354</b>

Tamanho da prole: Muito pequena (< 5 filhotes); Pequena (5 | 10 filhotes); Grande (10 | 15 filhotes) e Muito grande ( $\geq 15$  filhotes). Indefinido: acasalamentos com omissão da ordem de parto.

De acordo com notação com a notação usada na Tabela 9:

$n_{11} = 51$  é a frequência de fêmeas oriundas do parto de ordem 1 cuja progenitora teve uma prole inferior a 5 filhotes;

$n_{1.} = 1260$  é a frequência de fêmeas oriundas do parto de ordem 1;

$n_{.1} = 233$  é a frequência de filhotes oriundos de prole inferior a 5 filhotes;

O mesmo é aplicado às demais notações.

### 7.1.2. Matriz de Correspondência e Perfil de Linha e Coluna

A partir da tabela de contingência (Tabela 10) são construídas as matrizes de correspondência e os perfis de linhas e colunas como descrito a seguir.

A matriz de correspondência denotada por  $\mathbf{P}$  foi constituída pelas proporções das frequências observadas na tabela de contingência. Trata-se de uma matriz  $\mathbf{P}$  de dimensão  $(r \times c)$  constituída pelas proporções  $p_{ij} = \frac{n_{ij}}{n}$  (Tabela 10), conforme mostrado na Tabela 11,  $i = 1, 2, \dots, r; j = 1, 2, \dots, c$ . Os valores  $p_{i.}$  e  $p_{.j}$  representam as proporções marginais da linha  $i$  e das colunas  $j$ , respectivamente, sendo  $p_{i.} = \sum_{j=1}^c p_{ij}$  e  $p_{.j} = \sum_{i=1}^r p_{ij}$ . Assim,  $p_{i.}$  representa a proporção total de elementos classificados na categoria  $i$  da variável  $X$  e  $p_{.j}$  a proporção total de elementos classificados na categoria  $j$  da variável  $Y$ .

Tabela 11: Matriz de correspondência das proporções para uma tabela de contingência

		Colunas				
		1	2	...	c	Proporção de linha
Linhas	1	$p_{11}$	$p_{12}$	...	$p_{1c}$	$p_{1.}$
	2	$p_{21}$	$p_{22}$	...	$p_{2c}$	$p_{2.}$
	3	$p$	$p_{32}$	...	$p_{3c}$	$p_{3.}$
	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
	R	$p_{r1}$	$p_{r2}$	...	$p_{rc}$	$p_{r.}$
	Proporção de coluna	$p_{.1}$	$p_{.2}$	...	$p_{.c}$	1

A matriz perfil de linhas é denotada por  $D_r$  e tem a forma:

$$D_r = \begin{bmatrix} p_{1.} & 0 & \dots & 0 \\ 0 & p_{2.} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p_{r.} \end{bmatrix}$$

A matriz de perfil de colunas é denotada por  $D_c$  e tem a forma:

$$D_c = \begin{bmatrix} p_{.1} & 0 & \dots & 0 \\ 0 & p_{.2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p_{.c} \end{bmatrix}$$

Ambas as matrizes são diagonais e tem dimensões iguais à  $r \times r$  e  $c \times c$ , respectivamente.

Em análise de correspondência as proporções  $p_i$  e  $p_j$  são, em geral chamadas de massas das linha  $i$  e da coluna  $j$ , respectivamente, e  $p_{ij}$  chamada de densidade da linha  $i$  e da coluna  $j$ . Os vetores transpostos  $R = [p_1 \ p_2 \ \dots \ p_r]$  e  $C = [p_{.1} \ p_{.2} \ \dots \ p_{.c}]$  são as massas de linha e coluna, respectivamente. A massa é uma medida relativa de frequência utilizada para descrever o tamanho da categoria em uma tabulação cruzada. Logo, a massa total ao longo das linhas e colunas sempre será expressa pelo valor numérico 1.

A Tabela 12 mostra as proporções da matriz de correspondência e as densidades marginais de linhas e de colunas. Os valores foram expressos como  $i = 1, 2 \dots, 10; j = 1, 2 \dots, 4$ .

Como exemplo para entendimento da Tabela 12, o valor 0,008 para fêmeas, parto de ordem 1 e prole muito pequena ( $< 5$  filhotes) foi calculado como  $(51/6354)$ .

Tabela 12: Proporções para a matriz de correspondência

Sexo	Ordem do Parto	Tamanho da Prole				Massa linha
		$< 5$	5   10	10   15	$\geq 15$	
Fêmeas	1	0,008	0,060	0,121	0,008	<b>0,198</b>
	2	0,005	0,038	0,087	0,007	<b>0,136</b>
	3	0,003	0,023	0,026	0,007	<b>0,058</b>
	4	0,001	0,012	0,006	0,000	<b>0,019</b>
	Indefinido	0,002	0,025	0,043	0,005	<b>0,075</b>
Machos	1	0,007	0,072	0,131	0,011	<b>0,221</b>
	2	0,006	0,039	0,087	0,005	<b>0,136</b>
	3	0,002	0,028	0,025	0,007	<b>0,063</b>
	4	0,001	0,015	0,007	0,000	<b>0,023</b>
	Indefinido	0,002	0,022	0,040	0,006	<b>0,070</b>
Massa coluna		<b>0,037</b>	<b>0,334</b>	<b>0,573</b>	<b>0,056</b>	<b>1</b>

Tamanho da prole: Muito pequena ( $< 5$  filhotes); Pequena (5 | 10 filhotes); Grande (10 | 15 filhotes) e Muito grande ( $\geq 15$  filhotes). Indefinido: acasalamentos com omissão da ordem de parto.

A matriz com perfil de linhas é dada por:

$$D_r = \begin{bmatrix} 0,198 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0,136 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,058 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0,019 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0,075 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,221 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0,136 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,063 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,023 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,070 \end{bmatrix}$$

e matriz perfil de coluna é:

$$D_c = \begin{bmatrix} 0,037 & 0 & 0 & 0 \\ 0 & 0,334 & 0 & 0 \\ 0 & 0 & 0,573 & 0 \\ 0 & 0 & 0 & 0,056 \end{bmatrix}$$

Na Tabela 13 apresenta-se a proporções de frequências das caselas ( $ij$ ) relativas à frequência total da linha  $i$ ,  $i = 1, 2 \dots 10; j = 1, 2 \dots 4$ . A título de ilustração o valor para o sexo fêmeas, parto de ordem 1 e prole de tamanho muito pequeno ( $< 5$  filhotes) foi calculado como  $51/1260 = 0,040$  (Tabela 10).

Na Tabela 13, observa-se que para as categorias fêmeas, parto de ordem 1 e 2, as maiores proporções se concentraram nas proles de tamanho grande (10 ┆ 15 filhotes) e pequeno (5 ┆ 10 filhotes). Nota-se que os valores das proporções obtidas na prole de tamanho grande (10 ┆ 15 filhotes) foram superiores ao valor da prole de tamanho pequeno (5 ┆ 10 filhotes). Observa-se que os menores valores se concentraram nas proles de tamanho muito pequeno ( $< 5$  filhotes) e muito grande ( $\geq 15$  filhotes).

A categoria fêmeas, parto de ordem 3, também, apresentou maiores valores para as proles de tamanho grande (10 ┆ 15 filhotes) e pequeno (5 ┆ 10 filhotes). Diferentemente das outras categorias, o valor proporcional da prole de tamanho muito grande ( $\geq 15$  filhotes) foi superior ao valor da prole de tamanho muito pequeno ( $< 5$  filhotes).

A categoria fêmeas, parto de ordem 4, apresentou maiores valores proporcionais para a prole de tamanho pequeno (5 ┆ 10 filhotes) e grande (10 ┆ 15 filhotes). Ao contrário das outras categorias, o valor proporcional da prole de tamanho pequeno (5 ┆ 10 filhotes) foi superior ao valor obtido pela prole de tamanho grande (10 ┆ 15 filhotes). Além disso, foi a

única categoria que apresentou valor nulo para a prole de tamanho muito grande ( $\geq 15$  filhotes).

O parto categorizado como indefinido apresentou maiores valores para a prole de tamanho grande (10 | 15 filhotes) seguido da prole de tamanho pequeno (5 | 10 filhotes). Em contrapartida, os menores valores se concentraram nas proles de tamanho muito pequeno ( $< 5$  filhotes) e muito grande ( $\geq 15$  filhotes). Ressalta-se que o parto de ordem indefinida é caracterizado pela omissão da exata ordem de parto. Conseqüentemente, sugere-se que as maiores proporções observadas nessa categoria estejam atreladas à prevalência, principalmente, dos partos de ordem 1 e 2 em relação à prevalência dos partos de ordem 3 e 4 como descrito anteriormente.

Tabela 13: Proporções para o total da linha

Sexo	Ordem do Parto	Tamanho da Prole				Total
		$< 5$	5   10	10   15	$\geq 15$	
Fêmeas	1	0,040	<b>0,305</b>	<b>0,613</b>	0,042	1
	2	0,034	<b>0,279</b>	<b>0,637</b>	0,051	1
	3	0,046	<b>0,389</b>	<b>0,446</b>	<b>0,119</b>	1
	4	0,066	<b>0,623</b>	0,311	0,000	1
	Indefinido	0,027	<b>0,333</b>	<b>0,574</b>	0,065	1
Machos	1	0,033	<b>0,325</b>	<b>0,593</b>	0,049	1
	2	0,042	<b>0,285</b>	<b>0,640</b>	0,033	1
	3	0,035	<b>0,450</b>	<b>0,403</b>	<b>0,113</b>	1
	4	0,041	<b>0,651</b>	0,308	0,000	1
	Indefinido	0,027	<b>0,312</b>	<b>0,568</b>	0,093	1
Massa da coluna		<b>0,037</b>	<b>0,334</b>	<b>0,573</b>	<b>0,056</b>	1

Tamanho da prole: Muito pequena ( $< 5$  filhotes); Pequena (5 | 10 filhotes); Grande (10 | 15 filhotes) e Muito grande ( $\geq 15$  filhotes). Indefinido: acasalamentos com omissão da ordem de parto.

Na Tabela 13, resultados similares foram observados para os machos, parto de ordem 1, 2, 4 e indefinido. As maiores proporções se concentraram nas proles de tamanho grande (10 ┆ 15 filhotes) e pequeno (5 ┆ 10 filhotes), para os partos de ordem 1, 2 e indefinido. Em contrapartida, os menores valores se concentraram nas proles de tamanho muito pequeno (< 5 filhotes) e muito grande ( $\geq$  15 filhotes). Nessas categorias, também, observa-se superioridade na proporção da prole de tamanho grande (10 ┆ 15 filhotes) em relação à prole pequena (5 ┆ 10 filhotes). Como na categoria fêmeas, as maiores proporções encontradas para a categoria macho, parto de ordem indefinida, provavelmente, são decorrentes da maior prevalência dos partos de ordem 1 e 2 comparados a prevalência dos partos de ordem 3 e 4.

A categoria machos, parto de ordem 4, apresentou maiores valores proporcionais para a prole de tamanho pequeno (5 ┆ 10 filhotes) e grande (10 ┆ 15 filhotes). Ao contrário das outras categorias, o valor proporcional da prole de tamanho pequeno (5 ┆ 10 filhotes) foi superior ao valor obtido pela prole de tamanho grande (10 ┆ 15 filhotes). Além disso, foi a única categoria que apresentou valor nulo para a prole de tamanho muito grande ( $\geq$  15 filhotes).

A categoria machos, parto de ordem 3 apresentou maiores valores para as proles de tamanho pequeno (5 ┆ 10 filhotes) e grande (10 ┆ 15 filhotes). Ao contrário das fêmeas, parto de ordem 3, o maior valor se concentrou na prole de tamanho pequeno (5 ┆ 10 filhotes). Nessa categoria, o valor proporcional da prole de tamanho muito grande ( $\geq$  15 filhotes) foi superior ao valor da prole de tamanho muito pequeno (< 5 filhotes).

Para os valores da massa de coluna foram observados maiores proporções para a prole de tamanho grande (0,573) e para a prole de tamanho pequeno (0,334). As menores proporções foram para a prole de tamanho muito pequeno (0,037) e para a prole de tamanho muito grande (0,056).

A Tabela 14 apresenta-se a proporções das frequências das caselas ( $ij$ ) relativas à frequência total da coluna  $j$ ,  $j = 1, 2 \dots 4$ ;  $i = 1, 2 \dots 10$ . A título de ilustração o valor para o sexo fêmeas, parto de ordem 1 e prole de tamanho muito pequeno (< 5 filhotes) foi calculado como  $51/233 = 0,219$  (ver Tabela 10).

Na Tabela 14 observa-se que as maiores proporções se concentraram no parto de ordem 1, para ambos os sexos, em todos os tamanhos de prole. A prole de tamanho muito pequeno (< 5 filhotes) apresentou maiores proporções de frequência para a categoria fêmeas, parto de ordem 1 (0,219) seguida da categoria machos (0,202). Em contrapartida, a prole de tamanho pequeno (5 ┆ 10 filhotes), apresentou maiores proporções de frequência para a categoria machos, parto de ordem 1 (0,215) seguido da categoria fêmeas (0,181). Para prole

de tamanho grande (10 | 15 filhotes), parto de ordem 3, observa-se proporções similares entre fêmeas (0,045) e machos (0,044). A prole de tamanho muito grande (> 15 filhotes), parto de ordem 4, foi a única que apresentou valores nulos, para ambos sexos. Para a massa da linha observa-se que a maior proporção ficou concentrada nos machos com ordem de parto 1 (0,221) seguida das fêmeas (0,198). O parto de ordem 2 apresentou proporções similares (0,136), para ambos os sexos. Menores proporções foram observadas para as fêmeas, parto de ordem 4 (0,019) seguido do machos, parto de ordem 4 (0,023).

Tabela 14: Proporções para o total de coluna

Sexo	Ordem do Parto	Tamanho da Prole				Massa Linha
		< 5	5   10	10   15	≥ 15	
Fêmeas	1	<b>0,219</b>	<b>0,181</b>	<b>0,212</b>	<b>0,149</b>	<b>0,198</b>
	2	0,124	0,114	0,151	0,124	<b>0,136</b>
	3	0,073	0,068	<b>0,045</b>	0,124	<b>0,058</b>
	4	0,034	0,036	0,010	0,000	<b>0,019</b>
	Indefinido	0,056	0,075	0,075	0,087	<b>0,075</b>
Machos	1	<b>0,202</b>	<b>0,215</b>	<b>0,229</b>	<b>0,194</b>	<b>0,221</b>
	2	0,155	0,116	0,152	0,081	<b>0,136</b>
	3	0,060	0,085	<b>0,044</b>	0,126	<b>0,063</b>
	4	0,026	0,045	0,012	0,000	<b>0,023</b>
	Indefinido	0,052	0,065	0,069	0,115	<b>0,070</b>
Total		1	1	1	1	1

Tamanho da prole: Muito pequena (< 5 filhotes); Pequena (5 | 10 filhotes); Grande (10 | 15 filhotes) e Muito grande (≥ 15 filhotes). Indefinido: acasalamentos com omissão da ordem de parto.

### 7.1.3 Análise de Correspondência e Independência Entre as Variáveis

No teste estatístico de Qui-Quadrado a frequência observada em cada casela  $(i, j)$ , da tabela de contingência, ou seja os valores de  $n_{ij}$ , são comparados com os de frequência esperada sob a hipótese de que as variáveis  $X$  e  $Y$  são independentes (hipótese nula). Assim como é feito no teste estatístico Qui-Quadrado para avaliar a associação entre as variáveis, o objetivo da análise de correspondência no estudo das tabelas de contingência é avaliar a relação entre as frequências observadas em relação às frequências esperadas sob a condição de independência. No modelo de independência, a probabilidade de uma observação ser classificada na linha  $i$  e na coluna  $j$  ( $p_{ij}$ ) é igual ao produto das probabilidades marginais, isto é,  $p_{ij} = p_i \cdot p_j$ , sendo  $p_i$  a probabilidade da observação ser classificada na linha  $i$  e  $p_j$  de ser classificada na coluna  $j$ . Para avaliar essa premissa comparam-se as estimativas  $p_{ij}$  sob a suposição de independência como aquela estimada pelos dados amostrais.

Por meio da tabela de contingência,  $p_{ij}$  foi estimada como  $n_{ij}/n$ . Como exemplo, vamos considerar, o sexo fêmea, ordem de parto 1 e tamanho da prole muito pequena ( $< 5$  filhotes). A probabilidade estimada pelos dados amostrais é  $\hat{p}_{11} = 51/6354 = 0,008$  (Tabela 10) e a probabilidade estimada pela suposição de independência é  $\hat{p}_{11} = \hat{p}_i \cdot \hat{p}_{.1} = \frac{1260}{6354} \times \frac{233}{6354} = 0,007$ .

### 7.1.4 Relação com a estatística Qui-Quadrado de Pearson

Um teste comum de independência entre as duas variáveis aleatórias categóricas  $X$  e  $Y$  em tabelas de contingência é o teste Qui-Quadrado de *Pearson* ( $\chi^2$ ). Se  $X$  e  $Y$  forem independentes, espera-se que as densidades ( $n_{ij}/n$ ) sejam semelhantes ao produto das densidades marginais  $(n_{i.}/n)(n_{.j}/n)$  e, portanto, as frequências estimadas esperadas ( $e_{ij}$ ) sob a condição de independência são dadas por  $e_{ij} = n p_i p_j = n \frac{n_{i.}}{n} \frac{n_{.j}}{n} = \frac{n_{i.} n_{.j}}{n}$ .

Se a hipótese de independência é verdadeira, em grandes amostras, a estatística do teste de *Pearson* é:

$$\chi_{ob}^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - n_{i.} n_{.j}/n)^2}{n_{i.} n_{.j}/n}$$

que sob a hipótese nula tem uma distribuição Qui-quadrado  $gl = (r - 1)(c - 1)$  graus de liberdade. Grandes valores de  $\chi_{ob}^2$  indicam a rejeição da hipótese de independência entre as variáveis. Para um nível de significância  $\alpha$  fixo,  $0 < \alpha < 1$ , a hipótese nula é rejeitada se o valor de  $\chi_{ob}^2$  for maior ou igual ao valor crítico do teste  $\chi_c^2$ , obtido através da relação  $P[\chi_{gl}^2 \geq \chi_c^2] = \alpha$ , ou seja,  $\chi_c^2$  é o percentil  $(1 - \alpha)$  da distribuição Qui-Quadrado com  $gl$  graus de liberdade.

O teste Qui-Quadrado foi aplicado considerando os dados da Tabela 10. Na Tabela 15 têm-se as frequências esperadas ( $e_{ij}$ ) sob a suposição de independência. Como exemplo, para sexo fêmea, ordem de parto 1 e tamanho da prole muito pequena ( $< 5$  filhotes), a frequência esperada foi calculada como:  $6354 \times \frac{1260}{6354} \times \frac{233}{6354} = 46,204$ .

Tabela 15: Frequências esperadas

Sexo	Ordem do Parto	Tamanho da Prole			
		$< 5$	5   10	10   15	$\geq 15$
Fêmeas	1	46,204	420,595	722,606	70,595
	2	31,683	288,408	495,501	48,404
	3	13,568	123,508	212,194	20,730
	4	4,474	40,724	69,967	6,835
	Indefinido	17,492	159,225	273,558	26,725
Machos	1	51,594	469,664	806,910	78,831
	2	31,756	289,076	496,648	48,520
	3	14,668	133,522	229,399	22,411
	4	5,354	48,736	83,731	8,180
	Indefinido	16,208	147,542	253,486	24,764

Tamanho da prole: Muito pequena ( $< 5$  filhotes); Pequena (5 | 10 filhotes); Grande (10 | 15 filhotes) e Muito grande ( $\geq 15$  filhotes). Indefinido: acasalamentos com omissão da ordem de parto.

A Tabela 16 mostra os desvios (diferenças) entre os valores observados ( $n_{ij}$ ) e os esperados ( $e_{ij}$ ). A título de ilustração o desvio para sexo fêmeas, parto de ordem 1 e prole de tamanho muito pequeno ( $< 5$  filhotes) foi calculada como  $(51 - 46,204) = 4,796$ .

Tabela 16: Desvios da frequência observada em relação à esperada

Sexo	Ordem do Parto	Tamanho da Prole			
		$< 5$	$5 \mid 10$	$10 \mid 15$	$\geq 15$
Fêmeas	1	4,796	-36,595	49,394	-17,595
	2	-2,683	-47,408	54,499	-4,408
	3	3,432	20,492	-47,194	23,270
	4	3,526	35,276	-31,967	-6,835
	Indefinido	-4,492	-0,225	0,442	4,275
Machos	1	-4,594	-12,664	27,090	-9,831
	2	4,244	-42,076	57,352	-19,520
	3	-0,668	46,478	-68,399	22,586
	4	0,646	46,264	-38,731	-8,180
	Indefinido	-4,208	-9,542	-2,486	16,236

Tamanho da prole: Muito pequena ( $< 5$  filhotes); Pequena ( $5 \mid 10$  filhotes); Grande ( $10 \mid 15$  filhotes) e Muito grande ( $\geq 15$  filhotes). Indefinido: acasalamentos com omissão da ordem de parto.

As contribuições de cada frequência para o valor final da estatística Qui-Quadrado estão descritas na Tabela 17. A título de ilustração a contribuição da categoria fêmeas, parto de ordem 1 e prole de tamanho muito pequena ( $< 5$  filhotes) para o valor final da estatística

$$\text{Qui-Quadrado foi calculado como } \chi^2 = \frac{(n_{ij} - e_{ij})^2}{e_{ij}} = \frac{(4,796)^2}{46,204} = 0,498.$$

Tabela 17: Contribuição para a estatística Qui-Quadrado

Sexo	Ordem do Parto	Tamanho da Prole				Total
		< 5	5   10	10   15	≥ 15	
Fêmeas	1	0,498	3,184	3,376	4,385	11,443
	2	0,227	7,793	5,994	0,401	14,415
	3	0,868	3,400	10,496	26,120	<b>40,885</b>
	4	2,779	30,556	14,605	6,835	<b>54,776</b>
	Indefinido	1,153	0,000	0,001	0,684	1,838
Machos	1	0,409	0,341	0,919	1,226	2,886
	2	0,567	6,124	6,623	7,853	21,167
	3	0,030	16,178	20,394	22,768	<b>59,371</b>
	4	0,078	43,919	17,915	8,180	<b>70,092</b>
	Indefinido	1,093	0,617	0,024	10,644	12,378
Total		7,703	<b>112,113</b>	<b>80,339</b>	<b>89,098</b>	289,253

Tamanho da prole: Muito pequena (< 5 filhotes); Pequena (5 | 10 filhotes); Grande (10 | 15 filhotes) e Muito grande (≥ 15 filhotes). Indefinido: acasalamentos com omissão da ordem de parto.

A partir da Tabela 17 pode-se verificar que para as linhas relacionadas aos partos de ordem 3 e 4, as contribuições para a estatística Qui-quadrado foram maiores, para ambos os sexos. Para as colunas, observa-se que as maiores contribuições foram da prole de tamanho pequeno (5 | 10 filhotes), muito grande (≥ 5 filhotes) e grande (10 | 15) filhotes. Entretanto, a menor contribuição foi para prole de tamanho muito pequeno (< 5 filhotes).

O valor observado da estatística Qui-Quadrado ( $\chi^2 = 289,3$ ) foi elevado e altamente significativo a 5% de significância. O valor crítico do teste  $\chi_c^2$  igual a 40,11 para 27 graus de liberdade (Ver Anexo).

A inércia relativa é calculada dividindo-se a contribuição de cada frequência para o valor final da estatística Qui-quadrado pelo valor observado da estatística do Qui-Quadrado (Tabela 18). A título de ilustração a contribuição do sexo fêmeas, parto de ordem 1 e prole de tamanho muito pequeno (< 5 filhotes) o valor da inercia relativa foi  $0,498/289,253 = 0,002$ .

Tabela 18: Inércias relativas para as linhas e colunas

Sexo	Ordem do Parto	Tamanho da Prole				Total
		< 5	5   10	10   15	≥ 15	
Fêmeas	1	0,002	0,011	0,012	0,015	0,040
	2	0,001	0,027	0,021	0,001	0,050
	3	0,003	0,012	0,036	<b>0,090</b>	<b>0,141</b>
	4	0,010	<b>0,106</b>	0,050	0,024	<b>0,189</b>
	Indefinido	0,004	0,000	0,000	0,002	0,006
Machos	1	0,001	0,001	0,003	0,004	0,010
	2	0,002	0,021	0,023	0,027	0,073
	3	0,000	0,056	0,071	<b>0,079</b>	<b>0,205</b>
	4	0,000	<b>0,152</b>	0,062	0,028	<b>0,242</b>
	Indefinido	0,004	0,002	0,000	0,037	0,043
Total		0,027	<b>0,338</b>	0,278	<b>0,308</b>	1,00

Tamanho da prole: Muito pequena (< 5 filhotes); Pequena (5 | 10 filhotes); Grande (10 | 15 filhotes) e Muito grande (≥ 15 filhotes). Indefinido: acasalamentos com omissão da ordem de parto.

As maiores contribuições da inércia relativa foram representadas pelas proles categorizadas como pequena (5 | 10 filhotes) associada ao parto de ordem 4 e muito grande (≥ 15 filhotes) associada ao parto de ordem 3, para ambos os sexos. Para a prole categorizada como pequena a contribuição relativa da inércia dos machos, partos de ordem 4 (0,152) foi superior aos das fêmeas, parto de ordem 4 (0,106). Já, para a prole categorizada como muito grande (≥ 15 filhotes), parto de ordem 3 a contribuição relativa da inércia das fêmeas (0,090) foi superior aos machos (0,079).

Devido ao arredondamento (Ver Anexo), observa-se a presença de frequência relativas da inércia com valor próximo de zero para a prole de tamanho pequeno (< 5 filhotes), machos parto de ordem 3 (0,0001) e parto de ordem 4 (0,0003) e para prole de tamanho grande (10 | 15 filhotes), machos parto de ordem 4 (0,008). Em relação ao tamanho da prole a menor contribuição se concentrou na prole categorizada como muito pequena (< 5 filhotes) com valor de 0,027.

### 7.1.5 Decomposição Em Coordenadas Principais

Na análise de correspondência, a distância Qui-Quadrado é utilizada para construir novas variáveis chamadas de coordenadas principais, cujos valores representam o comportamento das linhas e colunas da tabela de contingência (Tabela 10). A partir desse valor um gráfico é construído e permite a visualização das agregações e as possíveis dependências entre elas.

Considere a matriz  $P_{r \times c}^*$  definida como:

$$P_{r \times c}^* = D_r^{-1/2} (P - RC') D_c^{-1/2}$$

sendo:  $\mathbf{R} = (R_1 R_2 \dots R_r)'$ ,  $\mathbf{C} = (C_1 C_2 \dots C_c)'$ ;  $R_i = \frac{n_i}{n}$ ;  $C_j = \frac{n_j}{n}$ ;  $i = 1, 2 \dots r$ ;  
 $j = 1, 2 \dots c$ ;  $D_r = \text{diag}(R)$ ;  $D_c = \text{diag}(C)$ ;  $D_r^{-1/2} = \text{diag}((R^{-1})^{1/2})$  e  
 $D_c^{-1/2} = \text{diag}((C^{-1})^{1/2})$ .

O termo genérico da matriz  $P^*$  é dado por  $P_{ij}^* = \frac{(p_{ij} - R_i C_j)}{\sqrt{R_i} \sqrt{C_j}}$ , sendo  $p_{ij} = \frac{n_{ij}}{n}$

Observe que o termo  $p_{ij}^*$  está relacionado ao termo genérico usado no teste Qui-Quadrado para testar a independência entre as variáveis. Se multiplicarmos o numerador por  $n$  e o denominador  $\sqrt{n}$ , ou seja, multiplicarmos toda a fração por  $\sqrt{n}$  teremos o termo:

$$\frac{(n_{ij} - n_i n_j)}{\sqrt{\frac{n_i}{n}} n_j}$$

que elevado ao quadrado resulta em  $\frac{(n_{ij} - \frac{n_i}{n} n_j)^2}{\frac{n_i}{n} n_j}$ , que é o termo genérico do total do teste

Qui-Quadrado.

Na Análise de Correspondência, as coordenadas que representam as linhas e colunas da tabela de contingência são constituídas a partir da matriz  $P_{r \times c}^*$ , utilizando-se o teorema de decomposição de matrizes em seus valores singulares (GREENACRE, 2015 e 1983; MINGOTI, 2007; JOBSON, 1992).

Na aplicação do teorema, as coordenadas das linhas são obtidas por meio da extração dos autovalores (e autovetores) da matriz  $P^* P^{*t}$  e as coordenadas das colunas por meio da extração dos autovalores (e autovetores) da matriz  $P^{*t} P^*$ . Os autovalores não nulos são iguais, mas os autovetores são diferentes.

A diagonal principal das matrizes  $P^*P^{*'} e P^{*'}P^*$  contém os valores dos termos  $(P_{ij}^*)^2$  e o traço dessas matrizes são iguais ao valor da estatística Qui-quadrado dividido por  $n$ . A soma dos autovalores ao quadrado dessas matrizes é chamada de Inércia Total expressa por:

$$\sum_{i=1}^k \lambda_i^2 = \frac{\chi_{obs}^2}{n} = \frac{1}{n} \sum_{(ij)} \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$

onde  $k = \min(r - 1, c - 1)$ ,  $n_{ij}$  representa o número de observações na casela  $(i, j)$  da tabela de contingência e  $e_{ij} = np_{ij}$  representa o número esperado de observações na casela  $(i, j)$  calculado sob a hipótese de que as duas variáveis  $X$  e  $Y$  são independentes. Sob a hipótese de independência,  $e_{ij}$  é dado por:

$$e_{ij} = np_{i.}p_{.j} = \frac{n_{i.}}{n} \frac{n_{.j}}{n} = \frac{n_{i.}n_{.j}}{n}$$

As coordenadas principais das linhas e colunas normalizadas são respectivamente dadas por:

$$Y_{(r \times k)}^* = D_r^{-1/2} U_{r \times k} \Lambda_{k \times k}$$

$$Z_{(c \times k)}^* = D_c^{-1/2} V_{c \times k} \Lambda_{k \times k}$$

sendo  $U_{r \times k}$  a matriz que contém os autovetores da matriz  $(P^*P^{*'})$  e  $V_{c \times k}$  a matriz que contém os autovetores da matriz  $(P^{*'}P^*)$  e  $\Lambda_{k \times k}$  é uma matriz diagonal que contém a raiz

quadrada dos autovalores isto é:  $\Lambda_{k \times k} = \begin{bmatrix} \sqrt{\lambda_1^2} & 0 & \dots & 0 \\ 0 & \sqrt{\lambda_2^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{\lambda_k^2} \end{bmatrix}$ , sendo  $\lambda_1^2 \geq \lambda_2^2 \geq \dots \geq \lambda_k^2$ .

A inércia total é igual a soma dos  $k$   $\lambda_i^2$ , isto é,  $\sum_{i=1}^K \lambda_i^2$ . A inércia total do sistema é alocada em  $K$  dimensões principais onde  $k = \min(r - 1; c - 1)$ . Para os dados tratados nessa monografia tem-se  $k = \min(10 - 1; 4 - 1) = 3$ , ou seja, podemos obter 3 coordenadas principais de linhas e colunas. Na Tabela 19 tem-se o percentual de explicação simples e acumulada em relação à inércia total, de cada dimensão. As duas primeiras coordenadas explicam quase a totalidade da inércia total do sistema (98%), sendo a primeira coordenada de maior importância (67,11%).

Tabela19: Inércia total

k	Inércia $\lambda_i^2$	Proporção (%)	Acumulativo (%)
1	0,0306	67,11	67,11
2	0,0141	30,93	98,04
3	0,0009	1,96	1,0
Total	0,0455		

Valores aproximados para 4 casas decimais – *Software Minitab for Windows*® (saída em anexo).

Saída computacional completa no Anexo.

### 7.1.6 Medidas de Avaliação do Ajuste

As Tabelas 20 e 21 contêm as medidas de avaliação do ajuste da Análise de Correspondência para as três dimensões das linhas e colunas, respectivamente. As medidas de ajuste da análise de correspondência são representadas pela contribuição, correlação e qualidade.

Tabela 20: Medidas de ajuste para o perfil de linhas

Coordenadas das linhas (l)													
		1ª Dimensão ( $l = 1$ )						2ª Dimensão ( $l = 2$ )			3ª Dimensão ( $l = 3$ )		
		$\lambda_1^2 = 0,0306$ (67,11%)						$\lambda_2^2 = 0,0141$ (30,93%)			$\lambda_3^2 = 0,0009$ (1,96%)		
Sexo	OP	Q	Massa	Inércia	Coord	Corr	Contr	Coord	Corr	Contr	Coord	Corr	Conrt
F	1	1	0,198	0,040	-0,082	0,735	0,043	-0,045	0,222	0,028	0,020	0,043	0,086
F	2	1	0,136	0,050	-0,129	<b>0,993</b>	0,074	0,008	0,003	0,001	-0,007	0,003	0,008
F	3	1	<b>0,058</b>	0,141	<b>0,240</b>	0,520	<b>0,110</b>	0,221	0,442	<b>0,202</b>	0,064	0,037	0,270
F	4	1	0,019	0,189	0,546	0,664	<b>0,187</b>	-0,383	0,326	<b>0,200</b>	0,067	0,010	0,098
F	Ind	1	0,075	0,006	0,004	0,004	0,000	0,042	0,468	0,010	-0,045	0,528	0,171
M	1	1	0,221	0,010	-0,035	0,607	0,009	-0,022	0,231	0,007	-0,018	0,162	0,082
M	2	1	0,136	0,073	-0,137	0,767	0,084	-0,071	0,206	0,049	0,026	0,027	0,101
M	3	1	0,063	0,205	0,344	0,797	<b>0,244</b>	0,173	0,202	<b>0,134</b>	-0,008	0,000	0,004
M	4	1	0,023	0,242	0,576	0,690	<b>0,249</b>	-0,380	0,301	<b>0,235</b>	-0,066	0,009	0,112
M	Ind	1	0,070	0,043	0,008	0,002	0,000	0,165	<b>0,967</b>	0,134	-0,029	0,031	0,067

F: fêmeas; M:machos; OP: ordem de parto; Ind: indefinido (acasalamentos com omissão da ordem); Q: qualidade; Coord: coordenada; Corr: correlação; Contr: contribuição; massa da linha ( $p_i$ ) e coluna ( $p_j$ ).

Tabela 21: Medidas de ajuste para o perfil de colunas

Coordenadas das colunas												
			1ª Dimensão ( $l=1$ )				2ª Dimensão ( $l=2$ )			3ª Dimensão ( $l=3$ )		
			$\lambda_1^2 = 0,0306$ (67,11%)				$\lambda_2^2 = 0,0141$ (30,93%)			$\lambda_3^2 = 0,0009$ (1,96%)		
Prole	Q	Massa	Inércia	Coord	Corr	Contr	Coord	Corr	Contr	Coord	Corr	Conrt
< 5 filhotes	1	0,037	0,027	0,058	0,103	0,004	-0,082	0,204	0,018	0,151	0,693	<b>0,942</b>
5   10 filhotes	1	0,334	0,388	0,221	<b>0,926</b>	<b>0,535</b>	-0,062	0,072	0,090	-0,011	0,002	0,041
10   15 filhotes	1	<b>0,573</b>	0,278	<b>-0,148</b>	<b>0,998</b>	<b>0,413</b>	-0,005	0,001	0,001	-0,004	0,001	0,012
≥ 15 filhotes	1	0,056	0,308	0,162	0,105	0,048	0,473	<b>0,895</b>	<b>0,891</b>	0,009	0,000	0,005

Tamanho da prole: Muito pequena (< 5 filhotes); Pequena (5 | 10 filhotes); Grande (10 | 15 filhotes) e Muito grande (≥ 15 filhotes). Q: qualidade; Coord: coordenada; Corr: correlação; Contr: Contribuição; massa da linha ( $p_i$ ) e coluna ( $p_j$ ).

- **Contribuição**

A contribuição de cada ponto no gráfico da análise de correspondência apresenta uma importância diferente a determinado eixo. Essa importância é medida pela contribuição que representa o quanto da variância (inércia) do eixo é devido ao ponto. A contribuição leva em consideração o valor do perfil (linhas e colunas), da coordenada principal e o da inércia total como descrito no exemplo a seguir considerando-se a Tabela 20:

*Contribuição para cada linha*

$$= (\text{massa da linha})(\text{valor da coordenada})^2 / \text{autovalor}$$

$$\text{Contribuição fêmeas parto de ordem 3; } l(1) = (0,058)(0,240)^2 / 0,0306 = 0,110$$

Desse modo, quanto maior o valor maior é a contribuição. Para as linhas, na Tabela 20, 1ª dimensão ( $l = 1$ ), os maiores valores das contribuições se concentraram nos machos, parto de ordem 4 (0,249) e ordem 3 (0,244) seguidos das contribuições das fêmeas, parto de ordem 4 (0,187) e de ordem 3 (0,110). Contribuições similares também foram verificadas para 2ª coordenada principal. Observa-se que os maiores valores se concentraram nos partos de ordem 3 e 4, para ambos os sexos.

- **Correlação**

A correlação mede a contribuição da coordenada para a variação da categoria. Desse modo, ela é útil para identificar quais coordenadas melhor representam a categoria. Em seu cálculo é considerado o valor de cada coordenada principal em relação aos valores de todas as coordenadas principais como ilustrado no exemplo a seguir, para fêmeas parto de ordem 3, dimensão  $l = 1$  (1ª coordenada principal de linhas).

$$\text{Correlação fêmeas parto de ordem 3; } l(1) = \frac{(0,240)^2}{(0,240)^2 + (0,221)^2 + (0,064)^2} = 0,520$$

Para a primeira dimensão de linhas, a maior correlação foi para a categoria fêmeas, parto de ordem 2 (0,993). Para a segunda dimensão, a maior correlação foi para os machos de parto indefinido (0,967). Observa-se que os maiores valores de correlações se concentraram na primeira dimensão ( $l = 1$ ).

- Qualidade

A qualidade representa a soma das correlações de acordo com número de coordenadas principais. Sempre a soma de todas as dimensões que representam o sistema (linha/coluna) é igual a 1 como descrito no exemplo a seguir:

$$\text{Qualidade Fêmea parto de ordem 3} = (0,520 + 0,442 + 0,037) = 1$$

Nesse exemplo, observa-se que a primeira e a segunda dimensões juntas contabilizam a maior parte do valor referente à qualidade do parto de ordem ( $0,520 + 0,442 = 0,962$ ).

A seguir apresentamos as medidas de ajuste para as três dimensões da coluna (Tabela 21). Como exemplo ilustrativo foi utilizado os dados da prole categorizada como grande (5 ┆ 10 filhotes) para expressar os valores da contribuição, correlação e qualidade.

*Contribuição para cada coluna*

$$= (\text{massa da coluna})(\text{valor da coordenada})^2 / \text{autovalor}$$

$$\text{Contribuição Prole tamanho grande ; } l(1) = (0,573)(-0,148)^2 / 0,0306 = 0,413$$

Para a primeira dimensão de colunas ( $l = 1$ ) os maiores valores das contribuições se concentraram na prole de tamanho pequeno (5 ┆ 10 filhotes) e grande (10 ┆ 15 filhotes). Para a segunda coordenada ( $l = 2$ ) a maior contribuição se concentrou na prole categorizada como muito grande ( $\geq 15$  filhotes) e para a terceira dimensão ( $l = 3$ ) a prole categorizada como muito pequena ( $< 5$  filhotes).

A correlação da prole de tamanho grande com a primeira coordenada principal das colunas é dado por:

$$\text{Correlação Prole tamanho grande; } l(1) = \frac{(-0,148)^2}{(-0,148)^2 + (-0,005)^2 + (-0,004)^2} = 0,99$$

Para a primeira dimensão, as maiores correlações foram para as proles categorizadas como pequena (0,926) e grande (0,998). Para a segunda dimensão a maior correlação foi para a prole de tamanho muito grande (0,895) e para a terceira dimensão foi para a prole muito pequena (0,942). Observa-se que os maiores valores de correlações se concentraram na primeira dimensão.

Considerando as 3 dimensões tem-se o valor da qualidade para prole de tamanho grande.

$$\text{Qualidade Prole tamanho grande} = (0,998 + 0,001 + 0,001) = 1$$

Nesse exemplo, observa-se que primeira dimensão sozinha contabilizou a maior parte do valor referente à qualidade (0,998) das coordenadas principais de colunas.

#### 7.1.7 Gráfico de Correspondência Simples

A análise de correspondência é uma técnica composicional que fornece uma representação multivariada de interdependência para dados categóricos. Nessa monografia, a Análise de Correspondência Simples foi aplicada com a finalidade de criar um mapa perceptual capaz de posicionar as categorias sexo, ordem de parto e tamanho da prole em um único mapa. Para isso, utiliza-se a distância Qui-Quadrado entre as proporções observadas na tabela de contingência sob a suposição de independência entre linhas e colunas da tabela.

A partir da decomposição da matriz ( $P^*$ , ver p. 34) obtêm-se os valores numéricos das coordenadas principais que descrevem o comportamento das linhas e colunas da tabela de contingência. Cada linha da tabela de contingência passa ser representada por  $k$  valores, ou seja,  $k$  dimensões de linhas e  $k$  dimensões de colunas.

O propósito é construir um gráfico que nos permita visualizar o comportamento das linhas e colunas simultaneamente usando os valores dessas dimensões. Em geral, é utilizado o gráfico de dispersão considerando-se as duas primeiras coordenadas de linhas e colunas

grafadas simultaneamente, pois essas estão relacionadas aos maiores autovalores (maiores inércias) e, portanto são os mais representativos da inércia total do sistema.

Para os dados analisados nessa monografia obteve-se  $k = 3$  coordenadas, sendo que as duas primeiras representam 98% da inércia total, indicando que um gráfico com essas coordenadas é bem representativo do sistema como todo, ou seja, das informações de associação entre linhas e colunas da tabela de contingência.

A qualidade do gráfico para duas coordenadas principais é igual a 96% para representação das linhas e 99,8% para representação das colunas que são valores bem elevados.

Na Figura 5 tem-se o gráfico de Correspondência utilizando-se as duas principais coordenadas principais de linhas e colunas. Como exemplo, a informação da linha fêmeas, parto de ordem 1 é representada no gráfico pelos valores  $(-0,082; -0,045)$  enquanto que a informação da prole de tamanho pequeno ( $5 \mid 10$  filhotes) é representado pelos valores  $(0,221; -0,062)$ .

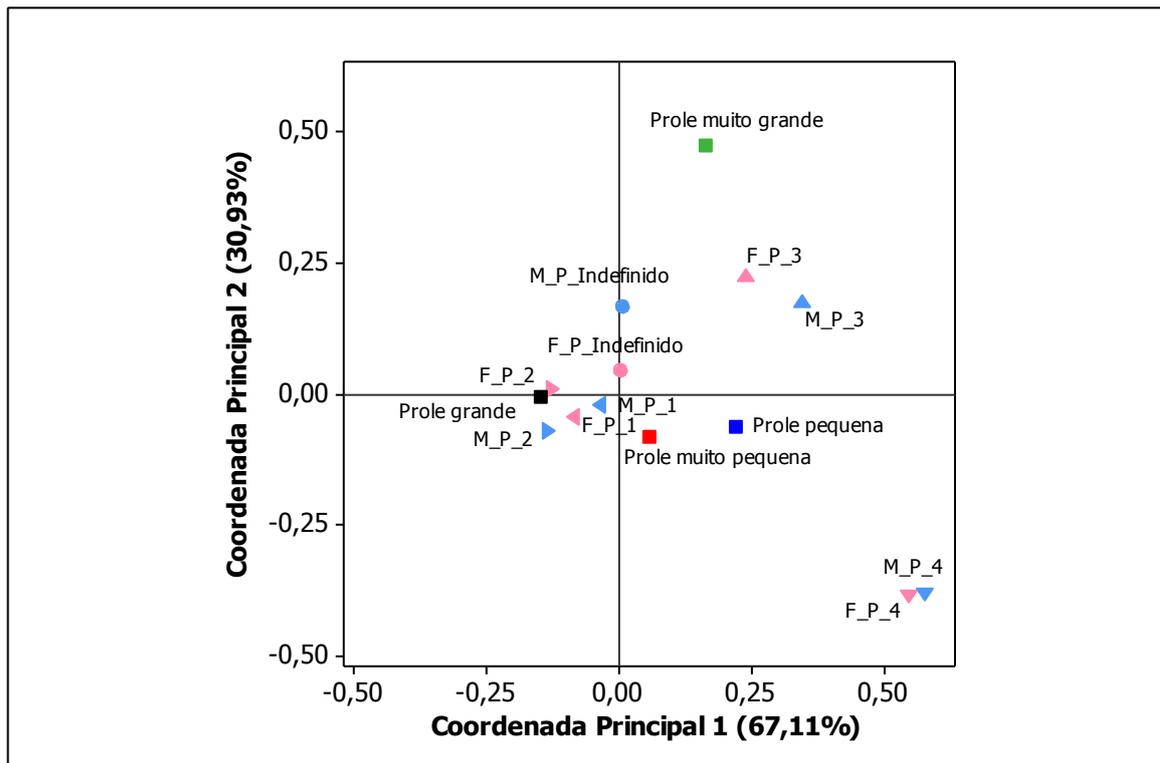


FIGURA 5: Análise de correspondência para avaliar as relações entre o sexo, a ordem de parto e tamanho da prole.

As linhas foram expressas pela variável sexo categorizado como F: fêmeas, M: machos e pela variável parto de ordem (P: 1, 2, 3, 4 e indefinido: acasalamentos com omissão da ordem). As colunas representam o tamanho da prole expressas como: prole muito pequena ( $< 5$  filhotes); pequena ( $5 \mid 10$  filhotes); grande ( $10 \mid 15$  filhotes) e muito grande ( $\geq 15$  filhotes).

De acordo com a Figura 5 é possível visualizar as seguintes associações:

- As proles categorizadas como muito pequena ( $< 5$  filhotes) e pequena ( $5 \text{ } \vdash \text{ } 10$  filhotes) se associaram ao parto de ordem 1, para ambos os sexos. Entretanto, em função da distância sugere-se uma fraca associação dessas proles com o parto de ordem 4, para ambos os sexos.
- A prole de tamanho grande ( $10 \text{ } \vdash \text{ } 15$  filhotes) se associou aos partos de ordem 2, para ambos os sexos. Essa categoria, também, se associou ao parto de ordem 1, para ambos os sexos.
- A prole categorizada como muito grande ( $\geq 15$  filhotes) se associou ao parto de ordem 3, para ambos os sexos. Entretanto, é uma associação mais fraca comparada à anterior.
- Os partos categorizados como parto de ordem indefinida (acasalamentos com omissão da ordem) não se associaram a nenhuma das categorias do tamanho da prole. Porém, fêmeas e machos dessa categoria associaram-se entre si.

## 8 CONSIDERAÇÕES FINAIS

Esse estudo surgiu da ideia de se conhecer as consequências da seleção genética na colônia de ratos Wistar com susceptibilidade à crise audiogênica em institutos de ensino/pesquisa (UFMG e USP) que utilizam esse modelo para melhor compreensão das epilepsias. Desse modo, o primeiro passo do estudo foi organizar os dados manuscritos da colônia e transcrevê-los para o formato digital.

Para isso foi elaborado um banco de dados no formato relacional utilizando as planilhas do Excel<sup>®</sup>. O modelo relacional permitiu que os dados do “banco acasalamento” e do banco “identificação do animal” fossem relacionados entre si. Porém, como desvantagem, as planilhas do programa Excel<sup>®</sup> não permitiram que os cruzamentos gerassem relatórios, implicando em maior gasto de tempo para organizar os dados no formato estatístico.

Ao longo de 22 anos de existência da colônia WAR/UFMG foram avaliadas 809 acasalamentos e 6354 animais. Desses registros, foi possível extrair dados sobre o sexo, a ordem de parto e o tamanho da prole. Esses dados foram estudados inicialmente com técnicas simples de estatística. A análise descritiva mostrou que não houve uma perda significativa de registros. Portanto, nesse momento, não foi necessário aplicar técnicas avançadas para avaliar o impacto sobre dados perdidos, tais como os métodos MAR para avaliar dados perdidos ao acaso (não aleatório) e MCAR para avaliar dados completamente perdidos ao acaso (aleatório) como sugerido pela literatura (HAIR-JR et al., 2009).

Num segundo passo do estudo, avaliou-se a associação entre sexo, ordem de parto e tamanho da prole. Esse tipo de estudo demandou o uso da técnica de Análise Correspondência Simples. Essa técnica é fundamentada na medida métrica de distância Qui-Quadrado e na transformação da matriz de distâncias em coordenadas principais por meio do uso do teorema de decomposição de matrizes em seus valores singulares. A técnica de Análise de Correspondência Simples mostrou que há associação entre as variáveis que fazem parte do acasalamento, ou seja, para ambos os sexos o tamanho da prole está associado à ordem de parto.

Para ambos os sexos, os partos de ordem 1 e 2 se associaram ao tamanho da prole grande (10 ┆ 15 filhotes). Fisiologicamente, esse tamanho de prole está condizente com a frequência esperada para roedores. A monografia mostrou que há fraca associação entre o parto de ordem 4 com as proles de tamanho muito pequeno (< 5 filhotes) e pequena (5 ┆ 10 filhotes), para ambos os sexos.

A prole categorizada como muito grande ( $\geq 15$  filhotes) foi mais rara e apresentou associação fraca com ao parto de ordem 3, para ambos os sexos. O aumento no tamanho da prole sugere que alterações hormonais estejam contribuindo para essa resposta.

A omissão de informação sobre a ordem do parto no banco de dados “acasalamento” mostrou que o parto categorizado como “indefinido” não apresentou qualquer associação com os tamanhos da prole propostos nessa monografia. Essa informação indica que há necessidade de padronização dos registros visando à homogeneidade dados. Inserir dados sobre acasalamento que não completaram o termo, progenitoras com prole nula ou com prole que foram disseminadas em função de canibalismo são informações importantes que ajudam compreender o desenvolvimento da colônia.

Esse foi o primeiro estudo que mostrou as possíveis relações entre o sexo, ordem de parto e o tamanho da prole em ratos da colônia WAR/UFMG por meio de técnica de caráter exploratório. A elaboração do Banco de Dados da colônia WAR/UFMG permitiu que simples dados se transformassem em informações e a Análise de Correspondência Simples possibilitou que essas informações gerassem conhecimentos para melhor compreensão do processo de acasalamento envolvendo três variáveis simultaneamente.

A Análise de Correspondência Simples pode ser estendida para a avaliação de associação entre múltiplas variáveis categóricas. Essa abordagem é denominada Análise de Correspondência Múltipla (GREENACRE, 2015, 1984; JOBSON, 1992) e está disponível no *software* estatístico R, dentre outros. Essa técnica será utilizada futuramente com a finalidade de explorar simultaneamente as variáveis que interferem no desenvolvimento da prole de ambos os ramos da colônia WAR mantidas em instituições de ensino/pesquisa, a faculdade de Medicina de Ribeirão Preto/ e o Instituto de Ciências Biológicas/UFMG.

Esse trabalho sugere que, no futuro, sejam investigados os mecanismos que estão envolvidos na relação entre o processo de envelhecimento e a geração de proles muito grandes.

Essa monografia gerou novos objetivos que serão desenvolvidos em Residência Pós-Doutoral com a finalidade de:

- 1 Avaliar a prevalência de morte por meio da técnica de regressão logística;
- 2 Avaliar como a seleção dos progenitores com susceptibilidade audiogênica interfere no comportamento da crise audiogênica por meio do método da regressão em árvore;
- 3 Estudar a relação entre o sexo e a susceptibilidade à crise audiogênica por meio da técnica de análise de agrupamento;
- 4 Comparar as colônias utilizando técnicas de inferência multivariadas;

- 5 Avaliar o desenvolvimento da colônia por meio de técnicas de séries temporais;
- 6 Transcrever os dados para uma base específica capaz de interagir todos os atributos, denominada “*MySQL* da “*Oracle Database*”.

## ANEXO

Saída do Software Minitab para análise descritiva e análise de correspondência

**Descriptive Statistics: FP1; FP2; FP3; FP4; FPI; MP1; MP2; MP3; MP4; MPI**

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3
FP1	1260	0	10,293	0,0816	2,895	2,000	8,000	11,000	12,750
FP2	864	0	10,598	0,102	2,986	2,000	9,000	11,000	13,000
FP3	370	0	10,146	0,169	3,248	1,000	8,000	10,000	12,000
FP4	122	0	8,131	0,197	2,174	2,000	7,000	9,000	10,000
FPI	477	0	10,254	0,140	3,067	2,000	8,000	11,000	13,000
MP1	1407	0	10,284	0,0779	2,920	2,000	8,000	11,000	13,000
MP2	866	0	10,432	0,0996	2,932	2,000	9,000	11,000	13,000
MP3	400	0	9,815	0,154	3,084	2,000	8,000	10,000	12,000
MP4	146	0	8,301	0,168	2,029	2,000	7,000	9,000	10,000
MPI	442	0	10,267	0,143	3,006	2,000	8,000	11,000	12,000

Variable	Maximum
FP1	16,000
FP2	18,000
FP3	17,000
FP4	11,000
FPI	16,000
MP1	16,000
MP2	18,000
MP3	17,000
MP4	12,000
MPI	16,000

## Contingency Table

	Column1	Column2	Column3	Column4	Total
Row1	51,000	384,000	772,000	53,000	1260,000
Row2	29,000	241,000	550,000	44,000	864,000
Row3	17,000	144,000	165,000	44,000	370,000
Row4	8,000	76,000	38,000	0,000	122,000
Row5	13,000	159,000	274,000	31,000	477,000
Row6	47,000	457,000	834,000	69,000	1407,000
Row7	36,000	247,000	554,000	29,000	866,000
Row8	14,000	180,000	161,000	45,000	400,000
Row9	6,000	95,000	45,000	0,000	146,000
Row10	12,000	138,000	251,000	41,000	442,000
Total	233,000	2121,000	3644,000	356,000	6354,000

## Row Profiles

	Column1	Column2	Column3	Column4	Mass
Row1	0,040	0,305	0,613	0,042	0,198
Row2	0,034	0,279	0,637	0,051	0,136
Row3	0,046	0,389	0,446	0,119	0,058
Row4	0,066	0,623	0,311	0,000	0,019
Row5	0,027	0,333	0,574	0,065	0,075
Row6	0,033	0,325	0,593	0,049	0,221
Row7	0,042	0,285	0,640	0,033	0,136
Row8	0,035	0,450	0,403	0,113	0,063
Row9	0,041	0,651	0,308	0,000	0,023
Row10	0,027	0,312	0,568	0,093	0,070
Mass	0,037	0,334	0,573	0,056	

## Column Profiles

	Column1	Column2	Column3	Column4	Mass
Row1	0,219	0,181	0,212	0,149	0,198
Row2	0,124	0,114	0,151	0,124	0,136
Row3	0,073	0,068	0,045	0,124	0,058
Row4	0,034	0,036	0,010	0,000	0,019
Row5	0,056	0,075	0,075	0,087	0,075
Row6	0,202	0,215	0,229	0,194	0,221
Row7	0,155	0,116	0,152	0,081	0,136
Row8	0,060	0,085	0,044	0,126	0,063
Row9	0,026	0,045	0,012	0,000	0,023
Row10	0,052	0,065	0,069	0,115	0,070
Mass	0,037	0,334	0,573	0,056	

## Expected Frequencies

	Column1	Column2	Column3	Column4
Row1	46,204	420,595	722,606	70,595
Row2	31,683	288,408	495,501	48,408
Row3	13,568	123,508	212,194	20,730
Row4	4,474	40,724	69,967	6,835
Row5	17,492	159,225	273,558	26,725
Row6	51,594	469,664	806,910	78,831
Row7	31,756	289,076	496,648	48,520
Row8	14,668	133,522	229,399	22,411
Row9	5,354	48,736	83,731	8,180
Row10	16,208	147,542	253,486	24,764

## Observed - Expected Frequencies

	Column1	Column2	Column3	Column4
Row1	4,796	-36,595	49,394	-17,595
Row2	-2,683	-47,408	54,499	-4,408
Row3	3,432	20,492	-47,194	23,270
Row4	3,526	35,276	-31,967	-6,835
Row5	-4,492	-0,225	0,442	4,275
Row6	-4,594	-12,664	27,090	-9,831
Row7	4,244	-42,076	57,352	-19,520
Row8	-0,668	46,478	-68,399	22,589
Row9	0,646	46,264	-38,731	-8,180
Row10	-4,208	-9,542	-2,486	16,236

## Chi-Square Distances

	Column1	Column2	Column3	Column4	Total
Row1	0,498	3,184	3,376	4,385	11,443
Row2	0,227	7,793	5,994	0,401	14,415
Row3	0,868	3,400	10,496	26,120	40,885
Row4	2,779	30,556	14,605	6,835	54,776
Row5	1,153	0,000	0,001	0,684	1,838
Row6	0,409	0,341	0,909	1,226	2,886
Row7	0,567	6,124	6,623	7,853	21,167
Row8	0,030	16,178	20,394	22,768	59,371
Row9	0,078	43,919	17,915	8,180	70,092
Row10	1,093	0,617	0,024	10,644	12,378
Total	7,703	112,113	80,339	89,098	289,253

## Relative Inertias

	Column1	Column2	Column3	Column4	Total
Row1	0,002	0,011	0,012	0,015	0,040
Row2	0,001	0,027	0,021	0,001	0,050
Row3	0,003	0,012	0,036	0,090	0,141

Row4	0,010	0,106	0,050	0,024	0,189
Row5	0,004	0,000	0,000	0,002	0,006
Row6	0,001	0,001	0,003	0,004	0,010
Row7	0,002	0,021	0,023	0,027	0,073
Row8	0,000	0,056	0,071	0,079	0,205
Row9	0,000	0,152	0,062	0,028	0,242
Row10	0,004	0,002	0,000	0,037	0,043
Total	0,027	0,388	0,278	0,308	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0306	0,6711	0,6711	*****
2	0,0141	0,3093	0,9804	*****
3	0,0009	0,0196	1,0000	
Total	0,0455			

Row Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	Row1	1,000	0,198	0,040	-0,082	0,735	0,043	-0,045	0,222	0,028
2	Row2	1,000	0,136	0,050	-0,129	0,993	0,074	0,008	0,003	0,001
3	Row3	1,000	0,058	0,141	0,240	0,520	0,110	0,221	0,442	0,202
4	Row4	1,000	0,019	0,189	0,546	0,664	0,187	-0,383	0,326	0,200
5	Row5	1,000	0,075	0,006	0,004	0,004	0,000	0,042	0,468	0,010
6	Row6	1,000	0,221	0,010	-0,035	0,607	0,009	-0,022	0,231	0,007
7	Row7	1,000	0,136	0,073	-0,137	0,767	0,084	-0,071	0,206	0,049
8	Row8	1,000	0,063	0,205	0,344	0,797	0,244	0,173	0,202	0,134
9	Row9	1,000	0,023	0,242	0,576	0,690	0,249	-0,380	0,301	0,235
10	Row10	1,000	0,070	0,043	0,008	0,002	0,000	0,165	0,967	0,134

ID	Name	Component 3		
		Coord	Corr	Contr
1	Row1	0,020	0,043	0,086
2	Row2	-0,007	0,003	0,008
3	Row3	0,064	0,037	0,270
4	Row4	0,067	0,010	0,098
5	Row5	-0,045	0,528	0,171
6	Row6	-0,018	0,162	0,082
7	Row7	0,026	0,027	0,101
8	Row8	-0,008	0,000	0,004
9	Row9	-0,066	0,009	0,112
10	Row10	-0,029	0,031	0,067

Column Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	Column1	1,000	0,037	0,027	0,058	0,103	0,004	-0,082	0,204	0,018
2	Column2	1,000	0,334	0,388	0,221	0,926	0,535	-0,062	0,072	0,090
3	Column3	1,000	0,573	0,278	-0,148	0,998	0,413	-0,005	0,001	0,001
4	Column4	1,000	0,056	0,308	0,162	0,105	0,048	0,473	0,895	0,891

ID	Name	Component 3		
		Coord	Corr	Contr
1	Column1	0,151	0,693	0,942
2	Column2	-0,011	0,002	0,041
3	Column3	-0,004	0,001	0,012
4	Column4	0,009	0,000	0,005

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	Row1	1,000	0,198	0,040	-0,082	0,735	0,043	-0,045	0,222	0,028
2	Row2	1,000	0,136	0,050	-0,129	0,993	0,074	0,008	0,003	0,001

3	Row3	1,000	0,058	0,141	0,240	0,520	0,110	0,221	0,442	0,202
4	Row4	1,000	0,019	0,189	0,546	0,664	0,187	-0,383	0,326	0,200
5	Row5	1,000	0,075	0,006	0,004	0,004	0,000	0,042	0,468	0,010
6	Row6	1,000	0,221	0,010	-0,035	0,607	0,009	-0,022	0,231	0,007
7	Row7	1,000	0,136	0,073	-0,137	0,767	0,084	-0,071	0,206	0,049
8	Row8	1,000	0,063	0,205	0,344	0,797	0,244	0,173	0,202	0,134
9	Row9	1,000	0,023	0,242	0,576	0,690	0,249	-0,380	0,301	0,235
10	Row10	1,000	0,070	0,043	0,008	0,002	0,000	0,165	0,967	0,134

### Inverse Cumulative Distribution Function

Chi-Square with 27 DF

P( X <= x )                    x  
0,95    40,1133

## REFERÊNCIAS

- 1 ABDULHAY, E.; ELAMARAN, V.; CHANDRASEKAR, M.; BALAJI, V. S.; NARASIMHAN, K. Automated diagnosis of epilepsy from EEG signals using ensemble learning approach. *Pattern Recognition Letters*, 2017, (in press).
- 2 AKIYAMA, T.; KOBAYASHI, K.; OHTSUKA, Y. Electroclinical characterization and classification of symptomatic epilepsies with very early onset by multiple correspondence analysis. *Epilepsy Research*, v. 91, p.232-239, 2010.
- 3 ALBERTI, G. Making sense of contingency tables in archaeology: the aid of correspondence analysis in intra-site activity areas research. *Journal of Data Science*, v.11, 2013, p. 479-99.
- 4 BETH, E. J.; LOMBARDO R.; SIMONETH, B. A european perception of food using two methods of correspondence analysis. *Food Quality Preference*, v. 22, 2011, p. 266-31.
- 5 BLANCO, A. M. Textbooks on differential calculus in eighteenth century Europe: a comparative stylistic analysis. *Journal of Data Science*, v.5, 2007 p.597-612.
- 6 BLUME, W.; LUDERS, H. O.; MIZRAHI, E.; TASSINARI, C.; BOAS, W. E; ENGEL-JR, J. Glossary of descriptive terminology for ictal semiology: report of the ILAE task force on classification and terminology. *Epilepsia*, v. 42, n. 9, p. 1212-1218, 2001.
- 7 DAILEY, J. W.; REIGEL, C. E.; MISHRA, P. K.; JOBE, P. C. Neurobiology of seizures predisposition in genetically epilepsy-prone rat. *Epilepsy Research*, v. 3, n. 1, p. 3-17, 1989.
- 8 DORETTO, M. C.; OLIVEIRA E SILVA, M.; FERREIRA-ALVES, D. L.; GARCIA-CAIRASCO, N.; REIS, A. M. Effect of lactation on the expression of audiogenic seizure: association with plasma prolactin profiles. *Epilepsy Research*, v. 54, p. 109 -121, 2003a.
- 9 DORETTO, M. C.; FONSECA, C. G.; LÔBO, R. B.; TERRA, V. C.; OLIVEIRA, J. A. C.; GARCIA-CAIRASCO, N. Quantitative study of the response to genetic selection of the Wistar audiogenic rat strain (WAR). *Behavior Genetics*. v. 33, n. 1, p. 33-42, 2003 b.

10 DORETTO, M. C.; CORTE-DE-OLIVEIRA, J. A.; ROSSETTI, F.; GARCIA-CAIRASCO, N. Role of the superior colliculus in the expression. of acute and kindled audiogenic seizures in Wistar audiogenic rats. *Epilepsia*, v. 50, n. 12, p. 2563-2574, 2010.

11 ELMASRI R.; NAVATHE, S. B. *Sistemas de banco de dados*. 6ed., São Paulo: Pearson, 2011, 788p.

12 FAINGOLD, C. L. Emergent properties of CNS neuronal network as target for pharmacology: application to anticonvulsant drug action. *Progress in Neurobiology*, v. 72, p.55-85, 2004.

13 FISHER, R. S.; BOAS, W. E; BLUME, W.; ELGER, C.; GENTON, P.; LEE, P.; ENGEL-JR, J. Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and International Bureau for Epilepsy (IBE). *Epilepsia*, v. 46, n. 4, p. 470-472, 2005.

14 FONSECA, A. G. A. R.; SANTOS, R. A. S.; MORAES, M. F. D; LEITE, M. F.; DORETTO, M. C. Vasopressinergic hypothalamic neurons are recruited during the audiogenic seizure of WARs. *Brain Research*, v. 1038, v. 1, n 15, p. 32-40, 2005.

15 FREUDENTHAL, M.; MARTÍN-SUÁREZ, E.; GALLARDO, J. A.; DAROCA, A. G; MINWER-BARAKAT, R. The application of correspondence analysis in palaeontology. *Compts Rendus Palevol*, v. 8, 2009, p.1-8.

16 GARCIA-CAIRASCO, N.; DORRETO, M. C.; PRADO, R. P.; JORGE, B. P. D.; TERRA, V. C.; OLIVEIRA, J. A. C. New insights into behavioral evaluation of audiogenic seizure. A comparison of two ethological methods. *Behavioural Brain Research*, v. 48, p. 49–56, 1992.

17 GARCIA-CAIRASCO, N.; SABBATINI, R. M. Role of the substantia nigra in audiogenic seizure: a neuroethological analysis in the rat. *Brazilian Journal Medical and Biological Research*, v. 16, n. 2, p. 171-183, 1983.

18 GARCIA-CAIRASCO, N.; WAKAMATSU, H.; OLIVEIRA, J. A. C.; GOMES, E. L. T.; DEL BEL, E. A.; MELLO, L. E. A. M. Neuro-ethological and morphological (neo-timm

staining) correlates of limbic recruitment during the development of audiogenic kindling in seizure susceptible Wistar rats. *Epilepsy Research*, v. 26, n. 1, p. 391-392; 1996.

19 GARCIA-CAIRASCO, N. A critical review on the participation of inferior colliculus in acoustic-motor and acoustic-limbic networks involved in the expression of acute and kindled audiogenic seizures. *Hearing Research*, v. 168, p. 208-222, 2002.

20 GARCIA-CAIRASCO, N; UMEOKA, E. H. L.; OLIVEIRA, J. A. C. The wistar audiogenic rat (WAR) strain and its contributions to epileptology and related comorbidities: History and perspectives. *Epilepsy & Behavior*, n. 71, p. 250-273, 2017.

21 GRANIERI, E.; ROSATI, G.; TOLA, R. A.; PAVONI, M.; PAOLINO, E.; PINNA, L.; MONETTI, V. A descriptive study of epilepsy in the district of Copparo, Italy, 1964-1978. *Epilepsia*, v. 24, p.502-514, 1983.

22 GREENACRE, M. J. *Theory and applications of correspondence analysis*. USA: Academic Press, 1984, 363p.

23 GREENACRE, M. Correspondence analysis of multivariate categorical data by weighted least-square. *Biometrika* , v.75, n.3, p. 457-467, 1988.

24 GREENACRE, M. Correspondence analysis of the Spanish National Health Survey. *Gaceta Sanitaria*, v.16, n.2, p. 160-170, 2002.

25 GREENACRE, M. Correspondence analysis. *International Encyclopedia of The Social & Behavioral Sciences*, v.5, p. 1-5, 2015.

26 HAIR-JR, J. F; BLACK, W. C.; BABIN, B. J. ANDERSON, R. E.;TATHAM, R. L. *Análise multivariada de dados*. 6ed., São Paulo: Bookman, 2009, 656p.

27 HAUSER, W. A; KURLAND, L. D. The epidemiology of epilepsy in Rochester, Minnesota, 1935 through 1967. *Epilepsia*, v. 16, p. 1-166, 1975.

- 28 IHLE, M.; FELDWISCH-DRENTROP, H.; TEIXEIRA, C. A.; WITON, A.; SCHELTER, B.; TIMMER, J; SCHULZE-BONHAGE, A. Epilepsiae – A European epilepsy database. *Computer Methods and Programs in Biomedicine*, v. 106, p. 127-38, 2012.
- 29 JOBE, P.C.; LAIRD, H. E; KO, K. H.; RAY, T.; DAILEY, J. W. Abnormalities in monoamine levels in the central nervous system of the genetically epilepsy-prone rat. *Epilepsia*, v. 23, p. 359-66, 1982.
- 30 JOBSON, J. D. *Applied multivariate data analysis.*, New York: Springer, V. 2, 1992, 731p.
- 31 KAMINSKA, A.; ICKOWICZ, A.; PLOUIN, P.; BRU, M. F.; DELLATOLAS, G.; DULAC, O. Delineation of cryptogenic Lennox-Gastau syndrome and myoclonic astatic epilepsy using multiple correspondence analysis. *Epilepsy Research*, v. 36, p. 15-29, 1999.
- 32 KIESMANN, M.; MARESCAUX, C.; VERGNES, M.; MICHELETTI, G.; DEPAULIS, A.; WARTER, J. M. Audiogenic seizures in Wistar rats before and after repeated auditory stimuli: clinical, pharmacological and electroencephalographic. *Journal of Neural Transmission*, v. 72, p. 235-244, 1988.
- 33 KIM, J. B.; SUH-IL, S.; KIM, J. H. Volumetric and shape analysis of hippocampal subfields in unilateral mesial temporal lobe epilepsy with hippocampal atrophy. *Epilepsy Research*, v. 117, p. 74 - 81, 2015.
- 34 KRUSHINSKY, L. V.; MOLODKINA, L. N.; FLESS, D. A.; DOBROKHOTOVA, L. P.; STESHENKO, A. P.; SEMIOKHINA, A. F.; ZORINA, Z. A.; ROMANOVA, L. G. The functional state of the brain during sonic stimulation. In: WELCH, B.; WELCH, A. *Physiological effects of noise*. New York: Plenum Press, p. 151-148, 1970.
- 35 LEES, G. V.; JONES, E, G.; McNAMARA, J. Epilepsy. *Neurobiology of Disease*, v. 7, p. 549-551, 2000.

- 36 MARTINEZ-DEL-RINCON, J.; SANTOFIMIA, M. J.; TORO, X.; BARBA, J.; ROMERO, F.; NAVAS, P.; LOPES, J. C. Non-linear classifiers applied to EEG analysis for epilepsy seizure detection. *Expert Systems With Applications*, v. 86, p. 99-112, 2017.
- 37 MAXSON, S. C. A genetic context for the study of audiogenic seizures. *Epilepsy & Behavior*, v. 71, p.154-159, 2017.
- 38 MAYNARD, L. M.; LEACH, J. M.; HORN, P. S.; SPAETH, C. G.; MANGANO, F. T.; HOLLAND, K. D.; MILES, L.; FAIST, R.; GREINER, H. M. Epilepsy prevalence and severity predictors in MRI-identified focal cortical dysplasia. *Epilepsy Research*, v.132, p. 41-49, 2017.
- 39 MINGOTI, S. A. *Análise de dados através de métodos de estatística multivariada*. Belo Horizonte: UFMG, 2007, 295p.
- 40 MINITAB 14.1 USER'S GUIDE. Pennsylvania. Minitab Inc., 2003.
- 41 MORAES, M. F. D.; GALVIS-ALONSO, O. Y.; GARCIA-CAIRASCO, N. Audiogenic kindling in the Wistar rat: a potential model for recruitment of limbic structures. *Epilepsy Research*. v.39, p. 251-259, 2000.
- 42 PAHO. Epilepsy in Latin America: technical document based on presentations at the international workshop held in Santiago, Chile, in August 2013, plus subsequent contributions, 2016. Disponível em: < <http://www.ilae.org/Visitors/policy/documents/PAHO-report2016-English.pdf> >. Acesso em: 17 Apr. 2017.
- 43 POLETAEVA, I. I.; SURINA, N. M; KOSTINA, Z. A.; PEREPELKINA, O. V. The krushinsky-Molodkina rat strain: The study of audiogenic epilepsy for 65 years. *Epilepsy & Behavior*, v.71, Part B, p. 130-141, 2017.
- 44 ROLSTON, J. D.; ENGLLOT, D. J.; KNOWLTON, R. C.; CHANG, E. F. Rate and complications of adult epilepsy surgery in North America: analysis of multiple databases. *Epilepsy Research*, v. 124, p. 55-62, 2016.

45 ROSS, K. C.; COLEMAN, J. R. Developmental and genetic audiogenic seizure models: behavior and biological substrates. *Neuroscience and Biobehavioral Reviews*, v. 24, p. 639-653, 2000.

46 ROSSETTI, F.; RODRIGUES, M. C. A.; de OLIVEIRA, J. A. C.; GARCIA-CAIRASCO, N. EEG wavelet analyses of the striatum-substantia nigra pars reticulata-superior colliculus circuitry: audiogenic seizures and anticonvulsant drug administration in Wistar audiogenic rats *Epilepsy Research*, v. 72, p. 192-208, 2006.

47 ROSSETTI, F.; RODRIGUES, M. C. A.; OLIVEIRA, J. A. C.; GARCIA-CAIRASCO, N. Behavioral and EEG effects of GABAergic manipulation of nigrothalamic pathway in the Wistar audiogenic rat strain. *Epilepsy & Behavior*, v. 22, p. 191-199, 2011.

48 SAJATOVIC, M.; TATSUOKA, C.; WELTER, E.; FRIEDMAN, D.; SPRUILL, T. M.; STOLL, S.; SAHOO, S. S.; BUKACH, A.; BAMPS, Y. A.; VALDEZ, J.; JOBST, B. C. Correlates of quality of life among individuals with epilepsy enrolled in self-management research: From the US Centers of Disease Control and Prevention Managing Epilepsy Well Network. *Epilepsy & Behavior*, v. 69, p. 177-180, 2017.

49 SIADAT, M-R.; SOLTANIAN-ZADEH, H.; FOTUHI, F.; ELISEVICH, K. Content-based image database system for epilepsy. *Computer Methods and Programs in Biomedicine*, v. 79, p. 209-26, 2005.

50 SCHEFFER, I. E.; BERKOVIC, S.; CAPOVILLA, G.; CONNOLLY, M. B.; FRENCH, J.; GUILHOTO, L.; HIRSCH, E.; JAIN, S.; MATHERN, G. W.; MOSHÉ, S. L.; NORDLI, D. R.; PERUCCA, E.; TOMSON, T.; WIEBE, S.; ZHANG, Y-H.; ZUBERI, S. M. ILAE classification of the epilepsies: Position paper of the ILAE Commission for classification and terminology. *Epilepsia*, v. 58, n.4, p. 512-521, 2017.

51 SCHEFFT, K. S.; TESTA, S. M.; DULAY, M. F.; PRIVITERA, M. D.; YEH, H-S. Preoperative assessment of confrontation naming ability and ictal paraphasia production in unilateral temporal lobe epilepsy. *Epilepsy & Behavior*, v. 4, p. 161-168, 2003.

52 TERRA, V. C.; GARCIA-CAIRASCO, N. Neuroethological evaluation of audiogenic seizures audiogenic-like seizure induced by microinjection of bicuculline into the inferior colliculus. II. Effects of nigral clobazam microinjections. *Behavioural Brain Research*, v. 52, p. 19 – 26, 1992.

53 VIEIRA, S. *Introdução à bioestatística*. 4ed., Rio de Janeiro: Elsevier, 2008. 345p.