

UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE VETERINÁRIA DA UFMG
PROGRAMA DE PÓS GRADUAÇÃO EM ZOOTECNIA

Andresa Eva Melo de Araújo

**Distribuições de probabilidade para análise de dados de
contagem de parasitos em bovinos**

Belo Horizonte

2019

Andresa Eva Melo de Araújo

**Distribuições de probabilidade para análise de dados de
contagem de parasitos em bovinos**

Versão final

Dissertação apresentada ao programa de pós-graduação em Zootecnia da Escola de Veterinária da Universidade Federal de Minas Gerais como requisito parcial para obtenção do título de Mestre em Zootecnia.

Área de concentração: Genética e Melhoramento Animal

Orientador: Prof. Dr. Fabio Luiz Buranelo Toral

Belo Horizonte

2019

Araújo, Andresa Eva Melo de, 1991
A663d Distribuições de probabilidade para análise de dados de contagem de parasitos em bovinos/
 Andresa Eva Melo de Araújo - 2019
 52 p.: il.

Orientador: Fabio Luiz Buranelo Toral

Dissertação – Universidade Federal de Minas Gerais, Escola de Veterinária.

1 – Genética – Teses - 2 - Bovinos – Poisson - Teses – 3- Binominal negativa – Teses -
I - Toral, Fabio Luiz Buranelo – II. Universidade Federal de Minas Gerais, Veterinária.
III – Título.

CDD - 636

DISSERTAÇÃO defendida e aprovada em 27/02/2019 pela comissão examinadora composta pelos seguintes membros:



Prof. Fabio Luiz Duranço Toral (Orientador)



Prof. Angela Maria Quintão Lana



Prof. João Paulo Amaral Haddad

*“Sermos nós mesmos faz com que acabemos excluídos pelos outros.
No entanto, fazer o que os outros querem nos exila de nós mesmos.”*

Clarissa Pinkola Estés

AGRADECIMENTOS

A Deus e à Nossa Senhora Aparecida por terem acalentado meu coração nos momentos difíceis e por me darem força e sabedoria para concluir essa etapa.

Ao meu pai, Clayton, pelo amor incondicional e pelo apoio durante os anos da graduação e do mestrado.

À minha mãe, Geralda, meu exemplo de mulher batalhadora. Que sempre me instigou a ser uma mulher independente e que luta para conquistar tudo que deseja.

Ao meu irmão, Gabriel, pois a admiração que ele sente por mim, faz-me querer ser cada dia melhor.

Ao meu noivo, Paulo Emílio, pelo amor, carinho, cuidado e apoio em todos os momentos.

Ao meu primo, Michell, que acredita tão fortemente no meu potencial, que me faz acreditar também de que sou capaz de tudo.

À minha família, pelas orações e por terem se preocupado, verdadeiramente, com meu bem estar.

À minha psicóloga, Cecília, por ter me auxiliado no resgate e manutenção do equilíbrio das minhas emoções.

Aos amigos do “Grupo da Paróquia”, pelos momentos divertidos que tivemos durante esses dois anos em que estivemos juntos no mestrado.

Ao prof. Fabio Luiz Buranelo Toral pelos ensinamentos, pelas idéias, pelo entusiasmo nas discussões e por ter contribuído para que eu concluísse o mestrado me sentindo, pessoal e profissionalmente, mais forte.

Ao prof. Renan Pedra de Souza pelo exemplo de profissional, pelos ensinamentos e por toda a disponibilidade em ajudar quando solicitei.

Aos professores, Ângela Maria Quintão Lana e João Paulo Amaral Haddad, por aceitarem participar da banca examinadora, pela discussão extremamente produtiva e pelas sugestões para conclusão do trabalho.

Aos funcionários da Fazenda Mundo Novo, por permitirem e auxiliarem na coleta dos dados que foram utilizados no trabalho.

À amiga de graduação e mestrado, Mariana Mamedes, por estar ao meu lado durante todos esses anos tornando o dia-a-dia mais leve, pelo apoio e auxílio.

Às colegas de laboratório, Gabriela Canabrava, Luiza Abreu e Virgínia Ribeiro pelos ensinamentos, pelas discussões que contribuíram para a melhoria e conclusão desse trabalho e pela disposição em me ajudar sempre que precisei.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) pela concessão da bolsa de estudos.

A todos que, direta ou indiretamente, contribuíram para a execução e conclusão desse trabalho.

RESUMO

Características de resistência aos parasitos são consideradas critérios de seleção em programas de melhoramento genético de bovinos, devido ao impacto econômico do parasitismo, por meio de registros de contagens de carrapatos, ovos e oocistos analisados a partir de diferentes distribuições de probabilidade. Objetivou-se identificar se existem diferenças nas estimativas de variância, seleção e classificação dos animais para contagens de parasitos analisadas pelas distribuições Poisson (P), Poisson inflacionada de zero (ZIP), binomial negativa (NB) e binomial negativa inflacionada de zero (ZINB). Foram analisadas 7.858 contagens de carrapatos e 7.144 contagens de nematódeos gastrointestinais (NGI) e *Eimeria* spp., via inferência Bayesiana. A distribuição que melhor se ajustou aos dados foi a Poisson. Estimativas de variância de ambiente permanente e grupo de contemporâneo, para contagens de carrapatos e *Eimeria* spp., e de variância genética aditiva e de ambiente permanente, para contagem de NGI, foram semelhantes entre P e ZIP. Entre NB e ZINB, houve semelhança em todas as variâncias. Na seleção, a similaridade de animais selecionados entre P e ZIP variou de 50 a 94% para carrapatos, 67 a 85% para NGI e 29 a 50% para *Eimeria* spp. Entre NB e ZINB, essa similaridade sempre foi próxima a 100%. Na classificação dos touros, as distribuições NB e ZINB não apresentaram diferenças de ordenação enquanto, nas demais distribuições, as alterações não foram significativas. Os resultados indicam que existem diferenças nas estimativas de variância obtidas entre as distribuições P e ZIP e NB e ZINB. Na prática, a utilização das distribuições P, NB e ZINB não interfere na seleção e classificação dos animais.

Palavras chave: inferência bayesiana, inflacionadas de zero, resistência, Poisson, binomial negativa.

ABSTRACT

Traits of resistance against parasites are considered as selection criteria in genetic improvement programs of cattle due to the economic impact associated with parasitism. In general, resistance traits are studied through count records of ticks, parasite eggs and oocysts and those counts are analyzed through different probability distributions. We aimed to identify whether there would be any differences in the estimation of variance and in the selection and ranking of animals for parasite counts between Poisson (P), zero-inflated Poisson (ZIP), negative binomial (NB), and zero-inflated negative binomial (ZINB) distributions. We analyzed 7,858 tick counts and 7,144 counts of both gastrointestinal nematodes (GIN) and *Eimeria* spp. through Bayesian inference. The distribution which best fitted the data was Poisson. Estimates of permanent environmental and contemporary group variances for tick and *Eimeria* spp. counts, and also of additive genetic and permanent environmental variances for NGI counts were similar between P and ZIP distributions. Between NB and ZINB distributions, all variances showed some similarity. In the animal selection, the similarity of individuals equally selected between P and ZIP showed a variation of 50-94% for ticks, 67-85% for GIN and 29-50% for *Eimeria* spp. Between NB e ZINB, this similarity was always near 100%. Neither NB nor ZINB showed any differences in animal ordering at the bull ranking, although in the other distributions the observed changes were not significant. Our results indicate that there are differences in the estimates of variance obtained between P, ZIP, NB and ZINB distributions. For practical purposes, the utilization of P, NB and ZINB distributions will not interfere in animal selection and ranking.

Keywords: Bayesian inference, negative binomial, Poisson, resistance, zero inflated.

LISTA DE TABELAS

Tabela 1 – Estatísticas descritivas ⁽¹⁾ e porcentagem de dados zerados para contagens de carrapatos, nematódeos gastrointestinais (NGI) e <i>Eimeria</i> spp. em bovinos Nelore.....	29
Tabela 2 – Médias (e limites inferiores e superiores dos intervalos de alta densidade com 90% das amostras) para as variâncias genéticas aditivas (σ_a^2), variâncias de ambiente permanente (σ_{pe}^2), variâncias de grupo de contemporâneos (σ_{cg}^2), variâncias residuais (σ_r^2) e critérios de informação da deviance (DIC) para contagens de carrapatos, nematoides gastrointestinais (NGI) e <i>Eimeria</i> spp. em bovinos Nelore obtidas por meio de análises unicaracterísticas sob diferentes distribuições.....	30
Tabela 3 – Médias (e limites inferiores e superiores dos intervalo de alta densidade com 90% das amostras) das probabilidades de zeros adicionais para as distribuições Poisson inflacionada de zero (ZIP) e binomial negativa inflacionada de zero (ZINB) para contagem de carrapatos, nematoides gastrointestinais (NGI) e <i>Eimeria</i> spp. em bovinos Nelore obtidos por meio de análises unicaracterísticas.....	31
Tabela 4 – Médias das frequências relativas (acima da diagonal) e absolutas (abaixo da diagonal) de tourinhos selecionados em comum, em análises unicaracterísticas sob diferentes distribuições, quando foram selecionados 10% daqueles com melhores valores genéticos para as contagens de carrapatos, nematódeos gastrointestinais (NGI) e <i>Eimeria</i> spp. (em seis estações de nascimento).....	32
Anexo A – Estatísticas descritivas ⁽¹⁾ para contagem de carrapatos, nematódeos gastrointestinais (NGI) e <i>Eimeria</i> spp. diferentes de zero.....	47
Anexo B – Critérios de informação da deviance (DIC) para contagem de carrapatos, nematoides gastrointestinais (NGI) e <i>Eimeria</i> spp. em bovinos Nelore obtidos no modelo completo e no modelo simples (sem efeito genético aditivo) sob diferentes distribuições.....	48
Anexo C – Frequências relativas (acima da diagonal) e absolutas (abaixo da diagonal) de tourinhos selecionados em comum, em análises unicaracterísticas sob diferentes distribuições, quando foram selecionados 10% daqueles com melhores valores genéticos para a contagem de carrapatos em cada estação de nascimento.....	49
Anexo D – Frequências relativas (acima da diagonal) e absolutas (abaixo da diagonal) de tourinhos selecionados em comum, em análises unicaracterísticas sob diferentes	

distribuições, quando foram selecionados 10% daqueles com melhores valores genéticos para a contagem de nematódeos gastrointestinais em cada estação de nascimento.....	50
Anexo E – Frequências relativas (acima da diagonal) e absolutas (abaixo da diagonal) de tourinhos selecionados em comum, em análises unicaracterísticas sob diferentes distribuições, quando foram selecionados 10% daqueles com melhores valores genéticos para a contagem de <i>Eimeria</i> spp. em cada estação de nascimento.....	51

LISTA DE FIGURAS

- Figura 1 – Distribuições dos dados de contagem de carrapatos, nematódeos gastrointestinais (NGI) e *Eimeria* spp. O pico referente à contagem 50 indica amostras que apresentaram contagem igual ou superior à 50 para cada uma das características (50 contagens de carrapatos, 15 contagens de NGI e 77 contagens de *Eimeria* spp.).....23
- Figura 2 – Classificações dos 13 touros com maior número de filhos no banco de dados de acordo com os valores genéticos para contagem de carrapatos, nematódeos gastrointestinais (NGI) e *Eimeria* spp. em bovinos Nelore segundo diferentes distribuições (N, normal; Nlog, normal com transformação logarítmica; P, Poisson; ZIP, Poisson inflacionada de zero; NB, binomial negativa; ZINB, binomial negativa inflacionada de zero).....34

SUMÁRIO

1. INTRODUÇÃO GERAL	11
2. REVISÃO DE LITERATURA	13
2.1. Avaliação genética para resistência à endo e ectoparasitas em bovinos	13
2.2. Análise de dados de contagem	14
2.2.1. Transformação logarítmica	15
2.2.2. Distribuição de Poisson	16
2.2.3. Distribuição Binomial Negativa.....	17
2.2.4. Distribuições Inflacionadas de Zero.....	17
2.3. Critérios de Ajustamento de Modelos	19
3. MATERIAL E MÉTODOS	21
3.1. Dados	21
3.2. Distribuição dos dados	23
3.3. Análises	25
4. RESULTADOS	28
5. DISCUSSÃO	34
6. CONCLUSÃO	38
REFERÊNCIAS BIBLIOGRÁFICAS	39
MATERIAL COMPLEMENTAR	46

1. INTRODUÇÃO GERAL

Doenças parasitárias em bovinos são responsáveis pela redução do desempenho dos animais, causando perdas anuais de até US\$13,96 bilhões no Brasil (Grisi et al., 2014). As parasitoses são responsáveis por perdas econômicas e produtivas em distintos sistemas de produção (Kornás et al., 2015; Mapholi et al., 2017; Aguerre et al., 2018) e seu controle ainda é feito com o uso, muitas vezes indiscriminado, de carrapaticidas e anti helmínticos, estimulando o desenvolvimento de parasitas resistentes e reduzindo a eficácia dos produtos (Le Gall et al., 2018). Diante disso, a resistência dos hospedeiros a esses parasitas se tornou característica de interesse em programas de melhoramento genético animal já que estudos comprovaram que ela pode responder à seleção (Cardoso et al., 2012; Passafaro et al., 2015; Ayres et al., 2015).

As estimativas dos parâmetros genéticos para resistência a parasitas são feitas utilizando-se dados de contagem de carrapatos, ovos e/ou oocistos por grama de fezes. Por serem dados quantitativos discretos e que possuem superdispersão, os dados de contagens não se adequam à distribuição normal (Ridout et al., 1998). Apesar disso, na maioria dos trabalhos, esses dados são submetidos à transformação logarítmica para aproximá-los da distribuição normal (May et al., 2017; Mapholi et al., 2016) ou são tratados como categóricos (animais susceptíveis ou resistentes) para serem analisados (Passafaro et al., 2015). Dessa forma, a real distribuição dos dados não é levada em consideração nas análises, o que pode comprometer as estimativas dos parâmetros genéticos e, posteriormente, a seleção dos animais mais resistentes aos parasitas. Entretanto, por meio da análise da distribuição dos dados e das propriedades estatísticas da própria característica, é possível sugerir algumas distribuições que sejam mais apropriadas à análise desses dados para a obtenção de estimativas mais adequadas dos parâmetros genéticos.

A distribuição de Poisson, por exemplo, é caracterizada por média e variância iguais, sendo utilizada em dados assimétricos e capaz de modelar a ocorrência de um evento em intervalos de tempo, volume ou área independentes (Resende et al., 2014). Com maior flexibilidade para a média e variância dos dados, a distribuição binomial negativa é indicada para dados que possuem superdispersão (Fang et al., 2014). Já quando os dados possuem excesso de zeros, as distribuições Poisson inflacionada de zero (ZIP) e binomial negativa inflacionada de zero (ZINB) podem se ajustar melhor aos dados (Ridout et al., 1998) ao

combinar a distribuição de Bernoulli à Poisson ou à binomial negativa (Costa et al., 2016), respectivamente. O excesso de zeros presente nos dados de contagens dos parasitos pode ocorrer porque a presença dos parasitos oscila em função de fatores diários como o turno em que os registros são obtidos, a eliminação dos ovos e oocistos nas fezes, a distribuição dos carrapatos no corpo do animal e a carga parasitária a qual os animais são expostos (Spedding, 1954; Hansen and Perry, 1994). Sendo assim, os zeros presentes nos dados são provenientes da baixa frequência dos parasitos no momento das contagens e de erros de amostragem, sendo subdivididos em zeros verdadeiros e falsos zeros, respectivamente (Costa et al., 2016)

A utilização das distribuições Poisson, ZIP, binomial negativa e ZINB na análise de dados de contagem já foi descrita por alguns autores (Rodrigues-Motta et al., 2017; Naya et al., 2008; Chipeta et al., 2014; Sae-Lim et al., 2017; Wang et al., 2018). Entretanto, não existe consenso a respeito de qual distribuição melhor se adequa à estrutura de dados de contagem de endo e ectoparasitas. Além disso, não se sabe qual o impacto causado pela escolha da distribuição a ser utilizada na estimativa dos valores genéticos dos animais e, conseqüentemente, na seleção dos animais superiores em programas de melhoramento genético.

O objetivo nessa dissertação foi identificar qual distribuição é mais adequada para a análise de registros de contagem de carrapatos, nematódeos gastrointestinais e *Eimeria* spp. analisados em escala real, pelas distribuições de Poisson, Poisson inflacionada de zero, binomial negativa, binomial negativa inflacionada de zero, e em escala logarítmica. A partir desses resultados, será possível definir, também, se há alterações no ordenamento dos animais segundo seus valores genéticos estimados a partir das diferentes distribuições.

2. REVISÃO DE LITERATURA

2.1. Avaliação genética para resistência à endo e ectoparasitas em bovinos

O parasitismo causado por endo e ectoparasitas é um entrave para a produção de carne bovina em todo o mundo, sendo responsável por perdas econômicas e produtivas em diferentes sistemas de produção (Grisi et al., 2014). Parasitas endêmicos são a principal causa de perdas econômicas na pecuária (Lopes et al., 2015) principalmente em países como o Brasil, cuja ocorrência dos parasitas é favorecida pelos climas tropical e subtropical predominantes (Grisi et al., 2014).

As perdas econômicas à bovinocultura geradas diretamente pelos parasitas, e indiretamente pelos patógenos por eles vinculados, são estimadas anualmente em US\$7.1 bilhões para nematódeos gastrointestinais e US\$3.2 bilhões para carrapatos no Brasil (Grisi et al., 2014) e US\$23.78 milhões para *Eimeria* spp., no México (Rodríguez-Vivas et al., 2017). Esse déficit pode chegar a, aproximadamente, US\$13,96 bilhões ao ano (Grisi et al., 2014) considerando os principais parasitas bovinos do Brasil.

O controle de parasitas nos bovinos resulta em uma série de benefícios relacionados à produtividade, incluindo aumento do ganho de peso, melhora na conversão alimentar e no sistema imunológico, aumento da produção de leite e da capacidade reprodutiva e redução da morbidade e mortalidade (Hawkins, 1993). Mas esse controle ainda é feito majoritariamente pelo uso de antiparasitários, como a Ivermectina, e a alta frequência dos tratamentos promoveu o desenvolvimento de resistência pelos nematódeos gastrointestinais e, simultaneamente, pelos carrapatos (Alegría-López et al., 2015), tornando os químicos ineficazes no combate aos parasitas (Le Gall et al., 2018). Dessa forma, a resistência a endo e ectoparasitas se tornou critério de seleção em programas de melhoramento genético animal, associado a outras características de interesse econômico (Cardoso et al., 2012), a fim de permitir a identificação e seleção de animais geneticamente superiores (resistentes), diminuir as perdas econômicas geradas pelo parasitismo e diminuir a quantidade de resíduos medicinais nos produtos de origem animal.

Foram relatadas herdabilidades médias de 0,12 e 0,11 para a resistência a carrapatos em cruzamentos Nelore x Hereford em duas diferentes regiões do Brasil (Ayres et al., 2015),

0,25 em Nelore (Passafaro et al., 2015), 0,17 nas raças Hereford e Braford (Biegelmeyer et al., 2017) e estimativas que variaram de 0 a 0,38 na raça Nguni, na África do Sul (Mapholi et al., 2017). Estimativas de herdabilidade obtidas para a resistência a nematoides gastrointestinais também apresentaram grande variabilidade. Zinsstag et al. (2000) relataram herdabilidade média de 0,18 em bovinos africanos. Outros autores relataram herdabilidade média de 0,21 em bovinos Holandeses (Coppieters et al., 2009), 0,28 e 0,11 em Angus e Hereford (Morris & Amyes, 2012) e 0,24 para a resistência a nematódeos gastrointestinais em Nelore (Passafaro et al., 2015). Para a resistência à *Eimeria* spp., a média da herdabilidade obtida por Passafaro et al. (2015) foi 0,23.

As variações de baixa a alta magnitude relatadas para as herdabilidades das características de resistência indicam a existência de variabilidade genética para as mesmas, sugerindo que podem ser utilizadas em programas de melhoramento genético a fim de promover mudanças genéticas e fenotípicas ao longo dos anos (Passafaro et al., 2015).

2.2. Análise de dados de contagem

Dados de contagem são tipicamente assimétricos e apresentam excesso de zeros (Chipeta et al., 2014). Nesse sentido, dados dessa natureza não possuem distribuição normal e não podem ser analisados considerando esse modelo de distribuição de probabilidade.

Na intenção de contornar esse problema, muitos pesquisadores realizam transformação logarítmica em dados de contagem para aproximá-los da distribuição normal e, assim, analisá-los como normais (Cardoso et al., 2015; Kornás et al., 2015). Outra alternativa é a análise desses dados na escala original, utilizando-se distribuições que mais se adequam à estrutura original dos dados. Alguns trabalhos que analisaram dados de contagem de diversas características consideraram que as distribuições mais adequadas na análise dos mesmos em escala real são: Poisson e Poisson inflacionado de zero (Rodrigues-Motta et al., 2007; Naya et al., 2008) e binomial negativa e binomial negativa inflacionada de zero (Sae-Lim et al., 2017).

2.2.1. Transformação logarítmica

Dados que não possuem distribuição normal, frequentemente, passam por uma transformação a fim de torná-los normais e aumentar a validade das análises estatísticas realizadas (Feng et al., 2014). De acordo com a estrutura dos dados que serão transformados, podem ser realizadas as transformações de Box-Cox, cúbica, quadrática, linear, raiz quadrada e logarítmica (Pino, 2014), entre outras. A transformação logarítmica é a mais popular delas e é dada por:

$$Y = \log_{10}(\text{trait} + 1,001),$$

em que: Y = variável na escala logarítmica; trait = variável na escala original e 1,001 é uma constante que evita dificuldades na transformação de observações iguais a zero (Barlett, 1947).

Apesar de ser utilizada para aproximar os dados da distribuição normal e reduzir a assimetria presente no banco de dados original, a transformação logarítmica pode, ao contrário do esperado, tornar esses dados mais assimétricos que os originais quando os mesmos seguem uma distribuição muito diferente da normal (Feng et al., 2014). Dessa forma, a variabilidade dos dados aumenta, tornam-se distantes dos dados originais e compromete as estimativas obtidas pelas análises (Feng et al., 2014).

Devido à sua popularidade, a transformação logarítmica é amplamente aplicada em trabalhos que utilizam dados de contagem em diversas áreas de pesquisa. Na área do melhoramento genético, por exemplo, a transformação foi utilizada recentemente por Cardoso et al. (2015) para a seleção genômica de bovinos resistentes aos carrapatos com base em registros de contagens dos parasitos nos corpos dos animais, por Gaddis et al. (2017) na avaliação dos componentes genéticos para o número de embriões viáveis obtidos a partir de diferentes técnicas reprodutivas e por Bobbo et al. (2018) na contagem de células somáticas para a seleção de bovinos resistentes à mastite.

2.2.2. Distribuição de Poisson

A distribuição de Poisson é usada para descrever a ocorrência de um evento raro, como por exemplo a mutação genética, em uma grande população em determinada unidade de tempo, volume ou área (Resende et al., 2014). Dados de contagem são, frequentemente, modelados pela distribuição de Poisson (Ridout et al., 1998) já que ela desempenha em dados categorizados o mesmo papel da distribuição normal em dados contínuos (Cordeiro e Demétrio, 2013).

Para que a Poisson consiga modelar o número de eventos raros, caracterizados por possuírem alta probabilidade nas baixas frequências e baixa probabilidade nas altas frequências de ocorrência, a distribuição dos dados deve ter número de eventos e intervalos de tempo independentes (Resende et al., 2014). Além disso, a probabilidade de ocorrência de um evento em um intervalo de tempo específico deve ser a mesma probabilidade em outro intervalo de tempo e a probabilidade de dois eventos ocorrerem simultaneamente deve ser sempre igual a zero (Martina et al., 2015).

Dados que assumem a distribuição de Poisson devem ser assimétricos e não apresentar valores negativos, como os dados de contagens, e devem possuir poucos valores distintos determinando média e variância iguais. Sendo $Y_i \sim Poisson(\lambda)$:

$$P(Y_i) = \frac{1}{Y_i!} \lambda^{Y_i} \exp(-\lambda) \quad ,$$

em que: λ é o parâmetro da distribuição de Poisson, $\lambda > 0$ e $E_P(Y_i) = var_P(Y_i) = \lambda$ (Gelman et al., 2013).

Visando a utilização dos dados na escala observada, Perez et al. (2017) realizaram a avaliação genética para a produção in vitro de embriões na raça Guzerá a partir de registros do número de oócitos e embriões produzidos na técnica. Em bovinos leiteiros, Hertl et al. (2018) avaliaram a relação existente entre o número observado de episódios de mastite clínica nos primeiros 100 dias de lactação de vacas de diferentes rebanhos e a ocorrência de mastite clínica ao longo da vida desses mesmos animais, também por meio da distribuição de Poisson.

2.2.3. Distribuição Binomial Negativa

Dados de contagens também podem ser analisados a partir da distribuição binomial negativa, considerada uma extensão da distribuição de Poisson (Fang et al., 2014). Pelo fato de ter variância maior que a média, a distribuição é mais adequada para a análise de dados que possuem uma superdispersão não causada pelo excesso de zeros. Sendo $Y_i \sim NB(k, \mu)$:

$$P(Y_i) = \frac{\Gamma(Y_i+k)}{\Gamma(k)\Gamma(Y_i+1)} \cdot \left(\frac{k}{\mu+k}\right)^k \cdot \left(1 - \frac{k}{\mu+k}\right)^{Y_i},$$

em que: k controla o nível de superdispersão do conjunto de dados e $E_{BN}(Y_i) = \mu$ e $var_{BN}(Y_i) = \mu + \frac{\mu^2}{k}$ (Harrison, 2014).

Devido à característica de superdispersão dos dados de contagem de parasitas, a distribuição binomial negativa muitas vezes é utilizada nas análises dos dados (Fang et al., 2014). A fim de obter maiores informações a respeito das infecções causadas por parasitas gastrointestinais em bovinos de corte, Jelinski et al. (2017) utilizaram a distribuição binomial negativa em dados de contagem de ovos de *Nematodirus* spp. e *Trichuris* spp. em amostras de fezes bovinas. Pensando em possíveis estratégias de controle de ectoparasitas em bovinos, Morel et al. (2017) também utilizaram essa distribuição para análise de registros de contagem de carrapatos realizados em bovinos na Argentina.

2.2.4. Distribuições Inflacionadas de Zero

Na prática, dados de contagem apresentam uma superdispersão causada pelo excesso de zeros que pode tornar a utilização das distribuições de Poisson e binomial negativa inadequadas para as análises (Ridout et al., 1998). Sendo assim, as distribuições Poisson inflacionada de zero (ZIP) e binomial negativa inflacionada de zero (ZINB) podem se ajustar melhor aos dados, sendo capazes de modelar melhor o excesso de zeros ao considerar que os

valores iguais a zero são mais esperados que os valores diferentes de zero (Lambert et al., 1992).

As distribuições inflacionadas de zero levam em consideração o surgimento de um excesso de dados zerados causados por dois processos distintos: os zeros verdadeiros ou estruturais que advém da baixa frequência do evento e ocorrem segundo um processo de Bernoulli com probabilidade π ; e os falsos zeros ou amostrais que resultam de erros de amostragem ou vícios visuais e são produzidos pelas distribuições de Poisson ou binomial negativa, com probabilidade $(1 - \pi)$ (Costa et al., 2016). A probabilidade total de zeros é a combinação das probabilidades dos dois processos, sendo impossível fazer a distinção entre os dois tipos.

Dessa maneira, para a distribuição ZIP temos que se $Y_i \sim ZIP(\lambda, \pi)$:

$$P(Y_i = 0) = \pi_i + (1 - \pi_i) e^{-\lambda_i}$$

$$P(Y_i = y) = \pi_i + (1 - \pi_i) \cdot \frac{\lambda^{Y_i} e^{-\lambda_i}}{Y_i!}$$

(Harrison, 2014).

Já para a distribuição ZINB, temos que se $Y_i \sim ZINB(k, \mu, \pi)$:

$$P(Y_i = 0) = \pi_i + (1 - \pi_i) (1 + k\mu)^{-1/k}$$

$$P(Y_i = y) = \frac{(1-\pi)\Gamma(Y_i + \frac{1}{k})(k\mu)^{Y_i}}{\Gamma(Y_i + 1)\Gamma(\frac{1}{k})(1+k\mu)^{Y_i + 1/k}}$$

(Fang et al., 2014).

As médias e variâncias para ambos os modelos inflacionados de zero são dadas por:

$$E_{ZI}(Y_i) = (1 - \pi_i)E(Y_i)$$

$$var_{ZI}(Y_i) = E_{ZI}(Y_i) + \pi_i(1 - \pi_i)[E_{ZI}(Y_i)]^2$$

(Sae-Lim et al., 2017).

Em ambas as distribuições, $P(Y_i = 0)$ representa a função de probabilidade da distribuição degenerada pelo excesso de zeros e $P(Y_i = y)$, da distribuição que se ajustaria aos dados estudados caso não houvesse esse excesso (Rodrigues, 2003). O autor também afirma que a mistura das distribuições de Bernoulli e Poisson ou binomial negativa, permite

que o excesso de zeros presente nos dados seja incorporado na distribuição e que a proporção de zeros que excede o predito seja dada por π .

Diante disso, a comparação dos resultados obtidos em análises que utilizaram os modelos inflacionados de zero das distribuições de Poisson e binomial negativa com os resultados obtidos a partir da utilização das distribuições Poisson e binomial negativa clássicas tem sido explorada por diversos autores. Perez et al. (2017) compararam os resultados obtidos com a utilização de ZIP e Poisson em dados de produção embrionária in vitro e concluíram que ambas poderiam ser utilizadas. Sae-Lim et al. (2017) compararam as distribuições Poisson, ZIP, binomial negativa e ZINB em dados de contagens de carrapatos e concluíram que a ZIP se ajustou melhor aos dados. Por outro lado, Adachi & Makita (2017), também comparando as quatro distribuições, concluíram que a ZINB apresentou melhor qualidade de ajuste a dados referentes ao número de parasitas *Echinococcus multilocularis* presentes na carcaça de suínos.

2.3. Critérios de Ajustamento de Modelos

A modelagem de dados é amplamente utilizada em diversas áreas de pesquisa, sendo que, determinar qual modelo melhor se ajusta aos dados é o principal obstáculo para os pesquisadores. Entretanto mais de um modelo pode ser capaz de descrever um mesmo fenômeno, já que cada pesquisador tem a liberdade de modelar um fenômeno de acordo com a metodologia que considera mais adequada, o que reforça a necessidade de definição de um modelo mais apropriado aos dados que serão analisados (Emiliano, 2013). Devido à complexidade existente na seleção do modelo mais apropriado, são utilizados critérios de ajustamento que consideram a complexidade do modelo e utilizam o número de variáveis inclusas no mesmo e o tamanho da amostra nessa seleção (Emiliano, 2009). Dessa forma, os critérios mais comumente utilizados em análises de diferentes modelos são o critério de informação de Akaike (AIC) (Akaike, 1974), critério de informação Bayesiano (BIC) (Schwarz, 1978) e o critério de informação da Deviance (DIC) (Spiegelhalter et al., 2002).

O critério de informação de Akaike (AIC) (Akaike, 1974) oferece uma medida relativa das informações perdidas quando determinado modelo é utilizado para descrever a realidade (Emiliano, 2009), tendo sido desenvolvido a partir da distância de Kullback-Leibler que

representa a distância entre o modelo verdadeiro e o modelo candidato (Burnham & Anderson, 2002). Segundo esses autores, os modelos em comparação devem ser classificados de acordo com o AIC e, aqueles com menores valores, são os selecionados.

Por outro lado, o critério de informação Bayesiano (BIC) (Schwarz, 1978) é um critério de avaliação de modelos de acordo com a probabilidade a posteriori e comprovado pela metodologia bayesiana, o que determinou o nome do critério (Emiliano, 2009). Da mesma forma que o AIC, os modelos que apresentam menores valores de BIC são os selecionados (Burnham & Anderson, 2002).

Modelos comparados através de análises Bayesianas são frequentemente selecionados pelo critério de informação da Deviance (DIC) (Spiegelhalter et al., 2002), considerado uma generalização do critério AIC (Ntzoufras, 2009). Além de possibilitar a seleção do melhor modelo, o DIC também pode ser utilizado para a seleção das (co)variáveis que constituirão o modelo utilizado (Oliveira et al., 2009). Spiegelhalter et al. (2002) sugerem que deve ser utilizado o módulo da diferença entre os valores de DIC de dois modelos distintos analisados para indicar se há ou não diferença entre os mesmos e que a seleção do modelo mais adequado também dependerá de um menor valor de DIC obtido. A verificação de existência de diferença entre dois modelos distintos é dada por:

$$D = |DIC_A - DIC_b|$$

em que: - $D < 5$ indica diferença não significativa;

- $5 \leq D \leq 10$ indica diferença significativa;

- $D > 10$ indica diferença altamente significativa.

Por ser mais adequado para a comparação de modelos baseando-se nas distribuições a posteriori dos desvios (Spiegelhalter, 2002), o DIC é mais utilizado em problemas bayesianos e seu valor é facilmente calculado a partir das amostras geradas pela metodologia bayesiana utilizada enquanto o cálculo de AIC e BIC requerem cálculo de verossimilhança para serem obtidos (Campos, 2011).

3. MATERIAL E MÉTODOS

3.1. Dados

Foram utilizados dados de bovinos machos da raça Nelore, nascidos em seis estações (2010/2011, 2011/2012, 2012/2013, 2013/2014, 2014/2015, 2015/2016) na Fazenda Mundo Novo localizada em Uberaba – Minas Gerais. A pluviosidade média anual da região é de 1.400 mm e as temperaturas anuais variam entre 5°C e 39°C, condizendo com o clima do cerrado brasileiro. Os animais foram criados a pasto com gramíneas do gênero *Urochloa*, em piquetes de aproximadamente 30 hectares com reservatórios naturais de água e receberam suplementação mineral à vontade durante todo o ano. Foi feita vermifugação ao nascimento e à desmama (aos sete meses de idade) com ivermectina 4% (Master LP, Ouro Fino Saúde Animal, Cravinhos, SP) e não houve uso de carrapaticidas durante o período de coleta de dados. Todos os animais de um mesmo lote foram, ou não, vermifugados no início das provas de ganho de peso. Além disso, os lotes eram rotacionados em diferentes piquetes a fim de garantir a mesma exposição dos animais aos parasitos.

Após a desmama, parte dos bezerros machos nascidos em cada ano (cerca de 50%) foram reagrupados em oito lotes, cuja amplitude máxima de idade dentro de cada lote era de 90 dias, enquanto a outra parte foi descartada por apresentarem peso abaixo da média. Estes animais participaram de provas de ganho de peso com duração de 294 dias, sendo 70 dias de adaptação e 224 dias de teste, nos quais os animais de um mesmo grupo foram mantidos no mesmo pasto, sob as mesmas condições de manejo. Os tourinhos foram pesados no início da prova (dia 0), ao final do período de adaptação (dia 70) e a cada 56 dias até o término da avaliação (dias 126, 182, 238 e 294)

As coletas dos dados coincidiram com as datas de pesagem das provas de ganho de peso e terminaram ao final da prova de cada safra. Foram realizadas contagens de carrapatos e coleta de amostras de fezes para contagens de ovos de nematódeos gastrointestinais (NGI) e oocistos de *Eimeria* spp. em animais naturalmente infestados pelos parasitas. As contagens de carrapatos seguiram a técnica proposta por Wharton & Utech (1970) em que se consideram as fêmeas ingurgitadas de *Rhipicephalus (Boophilus) microplus*, maiores ou iguais a 4,5 mm, no lado direito do corpo dos animais. Nesse trabalho, o número de carrapatos encontrados não foi

multiplicado por dois, portanto, os valores utilizados nas análises representam apenas a contagem de carrapatos em um lado do corpo dos animais. As amostras de fezes foram retiradas diretamente do reto dos tourinhos com uso de sacos plásticos lubrificados, identificados e mantidos resfriados até o momento da contagem de ovos e oocistos, de acordo com a técnica de McMaster Modificada (Ueno & Gonçalves, 1998), no Laboratório de Doenças Parasitárias da Escola de Veterinária da Universidade Federal de Minas Gerais. Para realização da contagem, foram diluídos 2 g de fezes em 28 ml de água potável e, após essa mistura ser peneirada, 2 ml da mesma foram homogeneizados com 2 ml de solução saturada de Sheater (500 g de açúcar, 6,5 ml de fenol e 360 ml de água). A câmara de McMaster, composta por duas lâminas separadas por um espaço de 1,5 mm entre os quais localiza-se duas células de contagem com 1 cm² (Castilho et al., 1984), foi preenchida com 0,15 ml dessa solução final e as contagens dos ovos foram realizadas em microscópios com lentes de aumento de 10 x. As informações das contagens também não foram multiplicadas por qualquer constante e os resultados utilizados nas análises referem-se à quantidade de ovos e oocistos presentes nas duas células de contagem.

Foram registradas 8.793 informações de contagens de parasitos durante seis safras de coletas de dados. Na edição do banco de dados, foram consideradas apenas as informações de tourinhos que apresentavam de três a cinco registros de contagens de carrapatos ou de ovos e oocistos, totalizando 7.858 registros de contagem de carrapatos e 7.144 contagens de ovos e oocistos provenientes de 1.685 e 1.609 animais, respectivamente. Para a contagem de carrapatos foram utilizados dados de progênie de 129 touros que possuíam 13,06 filhos em média (mínimo = 1; máximo = 66), enquanto para a contagem de ovos de nematódeos gastrointestinais e *Eimeria* spp. foram utilizados dados de progênie de 127 touros com 12,67 filhos em média (mínimo = 1; máximo = 61).

Para as três características, os registros aconteceram em 57 datas diferentes e os animais foram subdivididos em 228 grupos de contemporâneos formados pelo lote e data de pesagem do indivíduo. A matriz de parentesco foi composta por 5.883 animais, sendo 3.602 vacas e 618 touros.

3.2. Distribuição dos dados

As distribuições dos dados de contagem e o caráter longitudinal desse tipo de dados foram utilizados para avaliar possíveis distribuições que melhor se adequassem ao arquivo de dados (Figura 1).

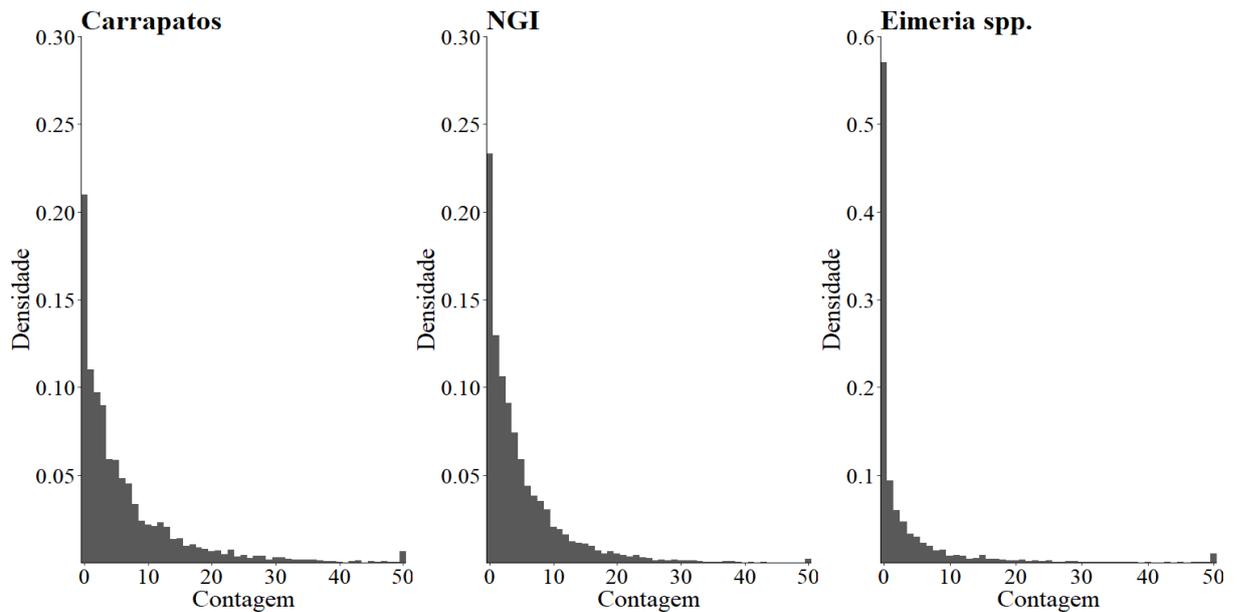


Figura 1. Distribuições dos dados de contagem de carrapatos, nematódeos gastrointestinais (NGI) e *Eimeria* spp. O pico referente à contagem 50 indica amostras que apresentaram contagem igual ou superior a 50 para cada uma das características (50 contagens de carrapatos, 15 contagens de NGI e 77 contagens de *Eimeria* spp.)

Dessa maneira, nesse trabalho os dados foram analisados considerando que os mesmos possuíam, a priori, as distribuições de Poisson, binomial negativa e suas modificações zero inflacionadas. Além disso, a título de comparação, analisamos os dados originais considerando que possuíam distribuição normal e realizamos a transformação logarítmica ($\log_{10}(\text{trait} + 1,001)$) dos dados, conforme é rotina nas análises de contagens de parasitos (Cardoso et al., 2015; Kornás et al., 2015).

Modelos que consideram a distribuição de Poisson devem ser capazes de modelar o número de eventos raros que ocorre em uma unidade de tempo, volume ou área (Resende et al., 2014) sendo o número de eventos e o intervalo de tempo, independentes. Dados que

assumem essa distribuição são caracterizados por média e variância iguais e pela assimetria causada pela inexistência de dados negativos. Sendo $Y_i \sim \text{Poisson}(\lambda)$:

$$P(Y_i) = \frac{1}{Y_i!} \lambda^{Y_i} \exp(-\lambda) \quad ,$$

em que: λ é o parâmetro da distribuição de Poisson, $\lambda > 0$ e $E_P(Y_i) = \text{var}_P(Y_i) = \lambda$ (Gelman et al., 2013).

A distribuição de Poisson inflacionada de zero (ZIP) também modela respostas com medidas independentes, mas os valores iguais a zero são mais frequentes que os diferentes de zero (Lambert et al., 1992), podendo ser mais adequada na análise de dados com excesso de zeros. Sendo $Y_i \sim \text{ZIP}(\lambda, \pi)$:

$$\begin{aligned} P(Y_i = 0) &= \pi_i + (1 - \pi_i) e^{-\lambda_i} \\ P(Y_i = y) &= \pi_i + (1 - \pi_i) \cdot \frac{\lambda^{Y_i} e^{-\lambda_i}}{Y_i!} \quad , \end{aligned}$$

(Harrison, 2014)

em que: $E_{ZIP}(Y_i) = (1 - \pi_i)\lambda$ e $\text{var}_{ZIP}(Y_i) = E_{ZIP}(Y_i) + \pi_i(1 - \pi_i)[E_{ZIP}(Y_i)]^2$ (Sae-Lim et al., 2017) e π é a probabilidade de zeros verdadeiros segundo a distribuição de Bernoulli (Costa et al., 2016).

Outra alternativa para analisar dados com superdispersão é a distribuição binomial negativa (NB). A NB assume que os dados apresentam a distribuição de Poisson mas a média (μ) é proveniente da distribuição Gamma consistindo numa mescla das duas distribuições (Harrison, 2014). Sendo $Y_i \sim \text{NB}(k, \mu)$:

$$P(Y_i) = \frac{\Gamma(Y_i + k)}{\Gamma(k)\Gamma(Y_i + 1)} \cdot \left(\frac{k}{\mu + k}\right)^k \cdot \left(1 - \frac{k}{\mu + k}\right)^{Y_i} \quad ,$$

em que: k controla o nível de superdispersão do conjunto de dados e $E_{BN}(Y_i) = \mu$ e $\text{var}_{BN}(Y_i) = \mu + \frac{\mu^2}{k}$ (Harrison, 2014).

Da mesma forma que o modelo inflacionado de zero da distribuição de Poisson, o modelo inflacionado de zero da binomial negativa também pode explicar melhor a superdispersão causada pelo excesso de zeros ao combinar as distribuições de Bernoulli e a NB. Sendo $Y_i \sim \text{ZINB}(k, \mu, \pi)$:

$$P(Y_i = 0) = \pi_i + (1 - \pi_i) (1 + k\mu)^{-1/k}$$

$$P(Y_i = y) = \frac{(1-\pi)\Gamma(Y_i+\frac{1}{k})(k\mu)^{Y_i}}{\Gamma(Y_i+1)\Gamma(\frac{1}{k})(1+k\mu)^{Y_i+1/k}},$$

(Fang et al., 2014)

em que: $E_{ZINB}(Y_i) = (1 - \pi_i)\mu$ e $var_{ZINB}(Y_i) = E_{ZINB}(Y_i) + \pi_i(1 - \pi_i)[E_{ZINB}(Y_i)]^2$ (Sae-Lim et al., 2017).

3.3. Análises

As estimativas dos componentes de variância para as três características de contagem foram obtidas por meio de análises unicaracterísticas realizadas via inferência Bayesiana onde as distribuições a posteriori para os parâmetros foram obtidas por aproximação de Laplace. As análises foram realizadas no pacote INLA (*Integrated Nested Laplace Approximation*; Rue e Martino, 2009) do programa R (The R Foundation for Statistical Computing, 2018) que possibilita a obtenção de resultados de forma mais rápida e acurada (Rue e Martino, 2007), sem a necessidade de verificar convergência de cadeias (Holand et al., 2013), em relação aos métodos convencionais frequentemente utilizados como o método de Monte Carlo via cadeia de Markov (MCMC). Os modelos consideraram efeito genético aditivo, de ambiente permanente, de grupo de contemporâneos e resíduo como efeitos aleatórios e idade como covariável e a distribuição que mais se adequou à estrutura dos dados foi determinada pela avaliação do Critério de Informação da Deviance (DIC) proposto por Spiegelhalter et al. (2002). Entretanto, dados log transformados são apresentados em uma escala logarítmica que impossibilita a comparação do DIC associado ao mesmo ao DIC dos demais modelos (Ayres et al., 2013). Nesse sentido, a comparação dos DICs só foi realizada entre as distribuições ZIP, P, ZINB e NB que seguem a escala observada dos dados.

O modelo estatístico utilizado nas análises dos dados com transformação logarítmica e considerando a distribuição normal, pode ser descrito por:

$$y = Z_1a + Z_2pe + Z_3cg + bi_m + e ,$$

em que: y = vetor com as observações de contagens de carrapatos, NGI ou *Eimeria* spp.; a = vetor dos efeitos genéticos aditivos associados à característica; pe = vetor dos efeitos de ambiente permanente associados à característica; cg = vetor dos efeitos de grupo de contemporâneo associados à característica; Z_1 , Z_2 e Z_3 = matrizes de incidência dos efeitos aleatórios genético aditivo, de ambiente permanente e grupo de contemporâneo, respectivamente; b = coeficiente de regressão linear associado à idade do animal; i_m = idade m do animal no dia da contagem; e = vetor dos resíduos associados a cada observação. O efeito de grupo de contemporâneo foi definido pela data da coleta dos dados e lote do indivíduo.

Para a distribuição de Poisson, definimos que $ln\lambda = \{ln\lambda_i\}_{i=1}^n$, sendo n o número de observações (Sae-Lim et al., 2017). O modelo linear para a distribuição é descrito por:

$$ln\lambda = Z_1a_\lambda + Z_2pe_\lambda + Z_3cg_\lambda + bi_m + e_\lambda,$$

em que: $ln\lambda$ = vetor de fenótipos dos dados de contagem considerando a distribuição de Poisson; a_λ = vetor dos efeitos genéticos aditivos associados ao fenótipo com distribuição de Poisson; pe_λ = vetor dos efeitos de ambiente permanente associados ao fenótipo com distribuição de Poisson; cg_λ = vetor dos efeitos de grupo de contemporâneo associados ao fenótipo com distribuição de Poisson; Z_1 , Z_2 e Z_3 = matrizes de incidência dos efeitos aleatórios genético aditivo, de ambiente permanente e grupo de contemporâneo, respectivamente; b = coeficiente de regressão linear associado à idade do animal; i_m = idade m do animal no dia da contagem; e_λ = vetor dos resíduos associados a cada observação com distribuição de Poisson.

Considerando a distribuição binomial negativa, $ln\omega = \{ln\omega_i\}_{i=1}^n$ (Sae-Lim et al., 2017), sendo o modelo linear descrito por:

$$ln\omega = Z_1a_\omega + Z_2pe_\omega + Z_3cg_\omega + bi_m + e_\omega.$$

em que: $ln\omega$ = vetor de fenótipos dos dados de contagem considerando a distribuição binomial negativa; a_ω = vetor dos efeitos genéticos aditivos associados ao fenótipo com distribuição binomial negativa; pe_ω = vetor dos efeitos de ambiente permanente associados ao fenótipo com distribuição binomial negativa; cg_ω = vetor dos efeitos de grupo de contemporâneo associados ao fenótipo com distribuição binomial negativa; Z_1 , Z_2 e Z_3 =

matrizes de incidência dos efeitos aleatórios genético aditivo, de ambiente permanente e grupo de contemporâneo, respectivamente; b = coeficiente de regressão linear associado à idade do animal; i_m = idade m do animal no dia da contagem; e_ω = vetor dos resíduos associados a cada observação com distribuição binomial negativa.

Por fim, os modelos das distribuições inflacionadas de zero diferem dos modelos das distribuições tradicionais apenas na função de ligação que depende da distribuição assumida para os parâmetros (Sae-Lim et al., 2017). Distribuições tradicionais possuem apenas uma função de ligação enquanto que as distribuições inflacionadas de zero, que são uma combinação da distribuição de Bernoulli e a distribuição inflacionada, apresentam duas funções de ligação (Lambert, 1992).

Por se tratar de uma inferência Bayesiana, foi preciso atribuir uma distribuição a priori para os parâmetros desconhecidos (Mathew et al., 2015). Para os hiperparâmetros (σ_a^2 , σ_{pe}^2 , σ_{cg}^2 e σ_e^2) assumimos a distribuição gama invertida (0,5; 0,5). Para os efeitos genéticos aditivos, de ambiente permanente, grupo de contemporâneo e para o resíduo foram assumidas distribuições normais multivariadas (MVN):

$$\begin{aligned} a | \sigma_a^2 &\sim MVN(0, A\sigma_a^2), \\ pe | \sigma_{pe}^2 &\sim MVN(0, I\sigma_{pe}^2), \\ cg | \sigma_{cg}^2 &\sim MVN(0, I\sigma_{cg}^2), \\ e | \sigma_e^2 &\sim MVN(0, I\sigma_e^2), \end{aligned}$$

em que: A = matriz de parentesco calculada a partir do pedigree; I = matriz identidade; σ_a^2 = variância genética aditiva; σ_{pe}^2 = variância de ambiente permanente; σ_{cg}^2 = variância do grupo de contemporâneo; σ_e^2 = variância residual.

Adicionalmente, foram ajustados modelos mais simples que desconsideravam o efeito genético aditivo direto para verificar sua importância nos modelos, segundo os valores de DIC, e comprovar que existe variabilidade genética para as características.

As estimativas de variância obtidas para as distribuições N e Nlog não foram comparadas com as demais distribuições pois as primeiras são obtidas para os registros de contagem enquanto que, para as demais distribuições, as variâncias referem-se aos parâmetros específicos de cada distribuição, como foi descrito nos modelos estatísticos das mesmas. Além disso, por haver uma transformação em que os dados são analisados em escala logarítmica, as variâncias obtidas para Nlog não são comparáveis nem com as estimativas

obtidas na distribuição normal já que essa analisa os dados em escala original. Sendo assim, as comparações realizadas foram restritas às distribuições P e ZIP e BN e ZINB. A mesma conduta foi realizada para as comparações do DIC.

O cálculo da herdabilidade para as características em cada uma das distribuições não foi realizado, pois a maneira como deve ser feita varia de uma distribuição para a outra (Holand et al., 2013) o que tornaria as comparações inadequadas.

A influência da distribuição dos dados na seleção dos melhores animais para cada característica foi verificada por meio do ordenamento dos valores genéticos e posterior seleção de 10% dos tourinhos (TOP 10) com melhores valores genéticos (VG) para verificar a proporção de animais selecionados em comum entre as diferentes distribuições estudadas. A seleção dos TOP10 foi realizada, primeiramente, em cada uma das safras para cada característica, seguindo-se do cálculo das médias de tourinhos selecionado em comum pelas diferentes distribuições nas seis safras para cada característica. Por fim, 10% dos touros que possuíam maior número de filhos no banco de dados foram selecionados e ordenados de acordo com seus VG. Essa ordenação foi feita dentro de cada distribuição e de cada característica a fim de observar se existe mudança na classificação desses animais de uma distribuição para a outra. Para a contagem de carrapatos os 13 touros com maior número de filhos apresentam 41,46 filhos em média (mínimo = 32; máximo = 66), enquanto para a contagem de ovos de nematódeos gastrointestinais e *Eimeria* spp. os touros avaliados apresentam 40,23 filhos em média (mínimo = 31; máximo = 61).

4. RESULTADOS

Considerando as 7.858 informações de contagens de carrapatos, o valor médio de carrapatos contados por animal foi 6,54 (mínimo = 0; máximo = 131) (Tabela 1). Para nematoides gastrointestinais (NGI) e *Eimeria* spp., que possuíam 7.144 informações de contagens, as médias apresentadas foram 5,01 (mínimo = 0; máximo = 80) e 3,84 (mínimo = 0; máximo = 328), respectivamente (Tabela 1). O número de contagens iguais a zero varia de acordo com o parasito, representando 20,97% dos registros de contagens de carrapatos, 23,35% das contagens de NGI e 57,10% das constagens de *Eimeria* spp. (Tabela 1). Diante disso, o número de contagens maiores que zero são 6.210 com média 8,28 para carrapatos, 5.476 com média 6,54 para NGI e 3.065 com média 8,95 para *Eimeria* spp. (Anexo A).

Tabela 1. Estatísticas descritivas⁽¹⁾ e porcentagem de dados zerados para contagens de carrapatos, nematódeos gastrointestinais (NGI) e *Eimeria* spp. em bovinos Nelore

⁽¹⁾*n*, número de dados; RGN, número de animais com registros de nascimento distintos; sd, desvio padrão; Mín., mínimo; Máx., máximo.

	<i>n</i>	RGN	Média	Sd	Mediana	Mín.	Máx.	Dados zerados
Carrapato	7.858	1.685	6,54	9,06	3	0	131	20,97%
NGI	7.144	1.609	5,01	6,74	3	0	80	23,35%
<i>Eimeria</i> spp.	7.144	1.609	3,84	13,17	0	0	328	57,10%

A distribuição que mais se adequou à estrutura dos dados de contagens de carrapatos, (NGI) e *Eimeria* spp. foi a Poisson (P), segundo o critério de ajuste DIC (Tabela 2). As análises adicionais realizadas confirmam que o efeito genético aditivo explica uma parte significativa da variância observada e que, portanto, há variabilidade genética em todas as características (Anexo B). Em relação às estimativas de variância, houve diferença nos valores obtidos para a variância genética aditiva para contagem de carrapatos e *Eimeria* spp. e para as variâncias de ambiente permanente e de grupo de contemporâneo para contagem de NGI entre as distribuições P e Poisson inflacionada de zero (ZIP). Além disso, as estimativas de variância residual foram sempre superiores na distribuição P, em relação à ZIP, e as demais estimativas obtidas dos parâmetros foram semelhantes (Tabela 2). Em relação às distribuições binomial negativa (BN) e binomial negativa inflacionada de zero (ZINB), os valores médios e de dispersão foram semelhantes para todas as variâncias estimadas para as três características (Tabela 2). Por fim, considerando-se as médias das estimativas para as probabilidade de zeros adicionais nos modelos inflacionados de zero, a distribuição ZIP apresentou maior probabilidade de zeros quando comparada com a distribuição ZINB nos três parasitos analisados (Tabela 3).

Tabela 2. Médias (e limites inferiores e superiores dos intervalos de alta densidade com 90% das amostras) para as variâncias genéticas aditivas (σ_a^2), variâncias de ambiente permanente (σ_{pe}^2), variâncias de grupo de contemporâneos (σ_{cg}^2), variâncias residuais (σ_r^2) e critérios de informação da

deviance (DIC) para contagens de carrapatos, nematoides gastrointestinais (NGI) e *Eimeria* spp. em bovinos Nelore obtidas por meio de análises unicaracterísticas sob diferentes distribuições

⁽¹⁾N, normal; Nlog, normal com transformação logarítmica dos dados; P, Poisson; ZIP, Poisson inflacionada de zero; NB, binomial negativa; ZINB, binomial negativa inflacionada de zero

Tabela 3. Média (e limites inferiores e superiores dos intervalo de alta densidade com 90%

Distribuições ⁽¹⁾	σ_a^2	σ_{pe}^2	σ_{cg}^2	σ_e^2	DIC
Carrapatos					
N	4,13 (2,32 ; 5,99)	7,36 (5,62 ; 9,08)	36,07 (30,04 ; 42,06)	18,50 (16,91 ; 20,23)	49.698,59
Nlog	0,02 (0,02 ; 0,02)	0,02 (0,01 ; 0,01)	0,12 (0,10 ; 0,14)	0,07 (0,06 ; 0,07)	-9.859,46
P	0,17 (0,12 ; 0,22)	0,14 (0,10 ; 0,18)	1,20 (1,00 ; 1,39)	0,35 (0,33 ; 0,37)	35.198,44
ZIP	0,08 (0,08 ; 0,09)	0,14 (0,14 ; 0,14)	1,05 (0,87 ; 1,23)	0,20 (0,19 ; 0,21)	36.592,52
NB	0,17 (0,12 ; 0,22)	0,14 (0,11 ; 0,18)	1,20 (1,00 ; 1,40)	0,09 (0,06 ; 0,12)	38.847,57
ZINB	0,16 (0,15 ; 0,17)	0,13 (0,12 ; 0,14)	1,15 (1,11 ; 1,18)	0,05 (0,05 ; 0,06)	39.018,86
NGI					
N	8,29 (6,29 ; 10,43)	3,62 (2,32 ; 4,95)	6,56 (5,23 ; 7,77)	11,97 (9,89 ; 13,79)	43.625,58
Nlog	0,03 (0,02 ; 0,03)	0,02 (0,02 ; 0,02)	0,05 (0,04 ; 0,06)	0,09 (0,08 ; 0,09)	-7.671,79
P	0,29 (0,21 ; 0,37)	0,16 (0,11 ; 0,22)	0,57 (0,46 ; 0,66)	0,58 (0,55 ; 0,61)	30.879,84
ZIP	0,21 (0,20 ; 0,22)	0,19 (0,18 ; 0,19)	0,34 (0,34 ; 0,35)	0,32 (0,31 ; 0,32)	32.469,40
NB	0,29 (0,22 ; 0,36)	0,15 (0,10 ; 0,19)	0,55 (0,45 ; 0,65)	0,06 (0,04 ; 0,08)	35.012,74
ZINB	0,28 (0,21 ; 0,35)	0,15 (0,10 ; 0,20)	0,55 (0,45 ; 0,64)	0,06 (0,04 ; 0,08)	34.999,56
<i>Eimeria</i> spp.					
N	2,82 (0,59 ; 5,17)	11,72 (8,64 ; 14,71)	23,82 (19,41 ; 28,38)	55,26 (46,88 ; 63,64)	54.808,43
Nlog	0,01 (0,01 ; 0,01)	0,01 (0,01 ; 0,01)	0,10 (0,08 ; 0,12)	0,08 (0,08 ; 0,09)	-6.198,44
P	0,23 (0,15 ; 0,30)	0,16 (0,09 ; 0,22)	4,92 (4,00 ; 5,80)	1,53 (1,44 ; 1,61)	19.087,85
ZIP	0,12 (0,09 ; 0,14)	0,14 (0,11 ; 0,16)	4,04 (3,33 ; 4,80)	0,90 (0,83 ; 0,97)	20.374,99
NB	0,21 (0,13 ; 0,28)	0,17 (0,10 ; 0,24)	4,84 (3,98 ; 5,72)	0,27 (0,19 ; 0,36)	23.168,03
ZINB	0,21 (0,12 ; 0,29)	0,17 (0,10 ; 0,24)	4,84 (3,99 ; 5,73)	0,27 (0,19 ; 0,35)	23.158,80

das amostras) da probabilidade de zeros adicionais para as distribuições Poisson inflacionada de zero (ZIP) e binomial negativa inflacionada de zero (ZINB) para contagem de carrapatos,

nematoides gastrointestinais (NGI) e *Eimeria* spp. em bovinos Nelore obtidos por meio de análises unicaracterísticas

Parasito	ZIP	ZINB
Carrapatos	0.1767 (0.1758 ; 0.1775)	0.0053 (0.0052 ; 0.0054)
NGI	0.2184 (0.2173 ; 0.2194)	0.0034 (0.0016 ; 0.0054)
<i>Eimeria</i> spp.	0.3321 (0.3159 ; 0.3451)	0.0093 (0.0031 ; 0.0175)

Selecionando-se 10% dos tourinhos com melhores valores genéticos para as três características de contagem nas seis distribuições em análise, foram observados 100% de animais selecionados em comum entre a distribuição N e a ZIP (Tabela 4). Comparando-se P e ZIP, essa proporção variou entre as características sendo 56% de tourinhos selecionados em comum para carrapatos, 74% para NGI e 46% para *Eimeria* spp. (Tabela 4). A proporção de tourinhos selecionados em comum foi elevada para as três características chegando a 98% para carrapatos, 100% para NGI e 99% para *Eimeria* spp. entre NB e ZINB (Tabela 4). Resultados mais detalhados podem ser observados no material suplementar onde são apresentadas a proporção de tourinhos selecionados em comum entre as distribuições nas seis diferentes estações de nascimento, para cada característica (Anexo C, Anexo D, Anexo E).

Tabela 4. Médias das frequências relativas (acima da diagonal) e absolutas (abaixo da diagonal) de tourinhos selecionados em comum, em análises unicaracterísticas sob diferentes distribuições, quando foram selecionados 10% daqueles com melhores valores genéticos para

as contagens de carrapatos, nematódeos gastrointestinais (NGI) e *Eimeria* spp. (em seis estações de nascimento)

Distribuição⁽¹⁾	N	Nlog	P	ZIP	NB	ZINB
Carrapatos (n⁽²⁾ = 169)						
N	-	59%	56%	100%	56%	56%
Nlog	100	-	87%	59%	86%	86%
P	94	147	-	56%	99%	98%
ZIP	169	100	94	-	56%	56%
NB	94	146	167	94	-	98%
ZINB	94	145	165	94	166	-
NGI (n⁽²⁾ = 161)						
N	-	79%	74%	100%	74%	74%
Nlog	128	-	88%	79%	86%	86%
P	120	142	-	74%	96%	96%
ZIP	161	128	120	-	74%	74%
NB	120	138	155	120	-	100%
ZINB	120	138	155	120	161	-
<i>Eimeria</i> spp. (n⁽²⁾ = 161)						
N	-	48%	46%	100%	51%	52%
Nlog	77	-	76%	48%	68%	69%
P	75	123	-	46%	85%	85%
ZIP	161	77	75	-	51%	52%
NB	83	110	137	83	-	99%
ZINB	84	111	138	84	160	-

⁽¹⁾N, normal; Nlog, normal com transformação logarítmica dos dados; P, Poisson; ZIP, Poisson inflacionada de zero; NB, binomial negativa; ZINB, binomial negativa inflacionada de zero. ⁽²⁾n = número de animais TOP10 no parasito correspondente

A média de tourinhos selecionados em comum também foi elevada entre as distribuições Nlog, P, NB e ZINB, superando 86% para a contagem de carrapatos. No entanto, menos animais em comum foram selecionados quando foram comparados os resultados obtidos por N ou por ZIP com os resultados de Nlog, P, NB ou ZINB sendo que as similaridades apresentadas nessas comparações chegaram a ser inferiores a 60% (Tabela 4).

Também foi observado esse comportamento para a contagem de NGI. Entre as distribuições Nlog, P, NB e ZINB, a média de tourinhos selecionados em comum também foi superior a 86% e, quando foram comparados os resultados obtidos por N ou por ZIP com os resultados de Nlog, P, NB ou ZINB, as similaridades sempre foram inferiores a 79% (Tabela 4).

Já para a contagem de *Eimeria* spp., a diferença entre os tourinhos selecionados foi mais acentuada em relação às contagens dos outros parasitos, estando sempre inferior a 52% quando comparamos os resultados das distribuições N ou ZIP com as demais distribuições, (Tabela 4). Além disso, para esse parasito a similaridade de tourinhos selecionados em comum entre as distribuições P e NB ou ZINB foi elevada, superando 85% (Tabela 4).

Os touros que possuíam maior número de filhos no banco de dados foram ordenados e ranqueados de acordo com os valores genéticos obtidos, sendo que, o touro que ocupa a primeira posição é o que apresenta menor valor genético para a característica (Figura 2). A numeração de classificação inicial dos touros foi definida conforme a ordem dos VG obtidos para a distribuição N. Os touros que ocuparam as posições 1 a 5, para contagem de carrapatos, apresentaram alterações em suas classificações entre as distribuições N, Nlog e P, enquanto os touros 6 e 7 sofreram mudanças de classificação na distribuição ZIP. Para os que ocupavam os níveis 8 a 12, as mudanças foram mais acentuadas entre as distribuições N, Nlog, P e ZIP. O touro classificado na 13ª posição, manteve sua classificação ao longo de todas as 6 distribuições (Figura 2). Apesar das mudanças descritas, todos os touros mantiveram a mesma classificação entre NB e ZINB (Figura 2).

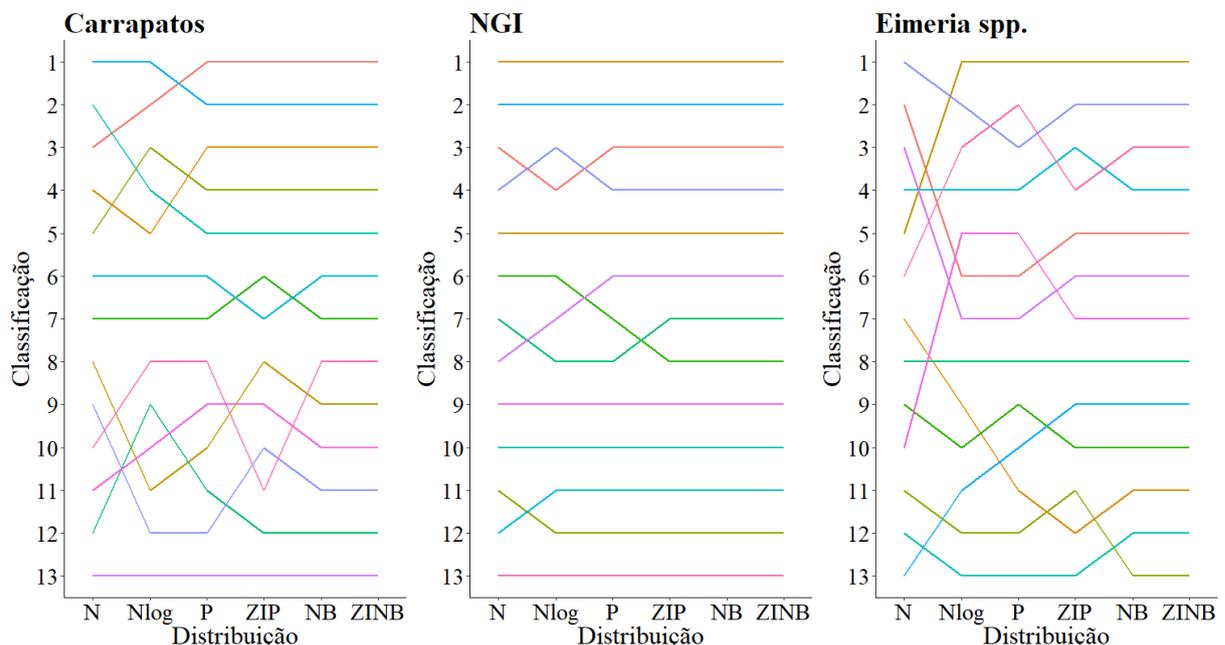


Figura 2. Classificações dos 13 touros com maior número de filhos no banco de dados de acordo com os valores genéticos para contagem de carrapatos, nematódeos gastrointestinais (NGI) e *Eimeria* spp. em bovinos Nelore segundo diferentes distribuições (N, normal; Nlog, normal com transformação logarítmica; P, Poisson; ZIP, Poisson inflacionada de zero; NB, binomial negativa; ZINB, binomial negativa inflacionada de zero)

A ordenação dos touros para o VG de contagem de NGI sofreu alterações mais discretas se compararmos com as mudanças observadas nas outras duas características (Figura 2). Para NGI, os touros classificados como 1, 2, 5, 9, 10 e 13 não passaram por reclassificação entre as distribuições. Os touros 3 e 4 alternaram suas classificações apenas na distribuição Nlog e os animais 11 e 12 alternaram em N. Já os touros 6, 7 e 8 sofreram alteração de ranqueamento mais acentuadas entre N, Nlog, P e ZIP. Para essa característica, apenas dois touros não mantiveram suas posições entre P e as distribuições ZIP, NB e ZINB e todos mantiveram a ordenação entre NB e ZINB (Figura 2).

Para a contagem de *Eimeria* spp., grandes mudanças foram observadas na classificação dos touros entre as distribuições N, Nlog, P e ZIP. Para essa característica, todos os animais também mantiveram suas colocações entre NB e ZINB e apenas o touro 8 manteve sua posição independentemente da distribuição testada (Figura 2).

5. DISCUSSÃO

A definição da distribuição de Poisson como mais adequada, segundo o critério de ajuste DIC, não confirma os resultados apresentados na literatura que definiram modelos distintos para a análise desse tipo de dados. Chipeta et al. (2014) avaliaram contagens de ovos de helmintos em amostras humanas e concluíram que os modelos inflacionados de zero são melhores para modelar esse tipo de dados, definindo o ZINB como mais adequada. Sae-Lim et al. (2017), por outro lado, concluíram que a distribuição ZIP é a mais adequada ao analisarem contagens de carrapatos em cordeiros. Essa diferença ocorre porque os dados de contagem atendem pressuposições tanto das distribuições Poisson e binomial negativa quanto de seus modelos inflacionados de zero. Sendo assim, a definição de qual distribuição melhor

modela esse tipo de dados fica a cargo da estrutura do banco de dados em questão e, também, dos fatores de variação incluídos no modelo estatístico.

Os valores de variância maiores que as médias obtidas para cada característica aqui analisada (Tabela 1) induzem que a distribuição do Poisson não se ajustaria aos dados, já que uma das pressuposições dessa distribuição é que média e variância sejam iguais. Entretanto, Sae-Lim et al. (2017) utilizaram um banco de dados em que a variância é três vezes maior que a média dos dados e a distribuição que melhor se ajustou foi a ZIP, indicando que a diferença existente entre média e variância pode não ser suficientemente grande a ponto de prejudicar a análise dos dados pela distribuição de Poisson ou seu modelo inflacionado de zero. Uma diferença entre esse trabalho e o de Sae-Lim et al. (2017), entretanto, é que a probabilidade de zeros para os dados dos autores na distribuição ZIP em relação à Poisson, é menor (0.042) que a probabilidade de zeros encontrada no presente trabalho (0.17 para carrapatos, 0.21 para NGI e 0.33 para *Eimeria* spp.) para a distribuição ZIP. Sendo assim, os dados aqui apresentados se ajustam bem à distribuição de Poisson, apesar de possuírem média e variância diferentes e percentual de inflacionamento de zeros superiores aqueles relatados por Sae-Lim et al. (2017). Também foi observado que a probabilidade de zeros obtida para *Eimeria* spp. é maior, tanto em ZIP quanto em ZINB, se compararmos com os outros dois parasitas analisados. Isso ocorre porque no banco de dados, o número de contagens iguais a zero é muito maior quando consideramos os registros de *Eimeria* spp. do que em carrapatos e NGI (Tabela 1) já que a eimeriose acomete principalmente animais jovens, de três semanas a seis meses de vida (Chartier & Paraud, 2012), e os dados analisados são de animais com mais de sete meses de vida.

Dados de contagem não se encaixam aos pressupostos da distribuição Normal, portanto, ela não é tradicionalmente utilizada na análise desses dados e, se utilizada, pode levar a previsões errôneas das estimativas dos componentes de variância (Ayres et al., 2013). Além disso, os valores de DIC e as estimativas de variância obtidas com a distribuição normal não são comparáveis com os valores obtidos pelas demais distribuições pois, nesses casos, as estimativas são obtidas para os valores reais dos dados enquanto que nas distribuições P, ZIP, NB e ZINB as estimativas são feitas para os parâmetros específicos de cada distribuição.

Dados de contagem que passam por transformação logarítmica não podem ter os valores de DIC e dos componentes de variância diretamente comparados com os resultados obtidos para as outras distribuições. A transformação, feita para aproximar os dados da distribuição normal, pode aumentar a distorção e a variabilidade dos dados, comprometendo os resultados das estimativas dos parâmetros genéticos já que os dados transformados são

muito diferentes dos originais (Feng et al., 2014). Apesar de não serem comparáveis com as demais distribuições aqui apresentadas, as estimativas de variância obtidas para a distribuição Nlog foram semelhantes às apresentadas na literatura (Cardoso et al., 2015; Kornás et al., 2015).

As médias a posteriori das variâncias obtidas para as três características analisadas considerando as distribuições P, ZIP, NB e ZINB são dadas para os parâmetros das distribuições e, portanto, são comparáveis. Em geral, esses resultados foram próximos entre as distribuições, dentro de cada característica (Tabela 2), o que também foi observado por Sae-Lim et al. (2017) que encontrou valores semelhantes para as estimativas de variâncias entre essas distribuições. Outros estudos com dados de contagem analisados pelas distribuições P e ZIP apresentaram resultados semelhantes entre as distribuições e não foram capazes de definir qual modelo melhor se adequaria à estrutura de dados (Rodrigues-Motta et al., 2007; Naya et al., 2008). A maior semelhança entre as estimativas de variâncias das distribuições NB e ZINB em relação à semelhança entre P e ZIP ocorre pois a distribuição ZINB possui uma probabilidade de zeros em relação à NB menor do que a distribuição ZIP em relação à distribuição P, tornando as estimativas mais próximas entre as mesmas em todas as características.

Comparando-se os dois modelos inflacionados de zero, observa-se que os intervalos de alta densidade são maiores, na maioria das vezes, na distribuição ZINB indicado menor qualidade de ajuste nessa distribuição (Tabela 2). Em contrapartida, as estimativas de probabilidade de zeros para a ZINB são bem menores que as estimativas obtidas para a distribuição ZIP para as três características avaliadas (Tabela 3), indicando que a distribuição ZINB pode modelar melhor dados com superdispersão devido à menor flexibilidade da Poisson que considera média e variância iguais (Wang et al., 2018). Dessa forma, comparando-se os modelos pelo DIC, a distribuição ZIP se ajusta melhor à modelagem dos dados utilizados nesse trabalho. Esse resultado pode indicar que a inflação de zeros na distribuição ZINB é menor que a inflação de zeros na ZIP porque o número esperado de zeros na distribuição NB é maior que o número esperado de zeros em P, portanto, a distribuição ZINB precisa se ajustar menos ao excesso de zeros do banco de dados.

A análise dos dados de contagem de carrapatos pelas distribuições P, NB e ZINB não alteraram, significativamente, quais tourinhos estão entre os TOP10 para a característica (Tabela 4). Entretanto, a utilização da distribuição ZIP pode ocasionar a seleção de tourinhos diferentes para a característica e descarte de animais superiores, já que ela apresentou menor similaridade de tourinhos selecionados em comum com as outras três distribuições (P, NB e

ZINB). As elevadas proporções de tourinhos selecionados em comum nas três características entre P, NB e ZINB (Tabelas 4) indicam quem o uso de qualquer uma dessas distribuições não interfere ou interfere pouco na seleção dos tourinhos considerados superiores segundo seus valores genéticos.

Por fim, a classificação dos touros com maior número de filhos no banco de dados (Figura 2) mostra que a utilização das distribuições NB e ZINB são as que menos induzem à seleção de touros diferentes, para as três características, por não provocarem alterações na ordenação com base nos valores genéticos. Essa maior similaridade entre NB e ZINB, também pode ser justificada pela menor inflação de zeros entre NB e ZINB em relação à P e ZIP. Além disso, de forma geral, as maiores alterações na classificação dos touros para todas as características ocorreram entre as distribuições N e Nlog, tendendo a reduzir as alterações de classificação, e até mesmo manter a ordem dos touros, quando comparamos as distribuições P, ZIP, NB e ZINB.

6. CONCLUSÃO

A distribuição que melhor se ajusta aos dados de contagens de carrapatos, nematóides gastrointestinais e *Eimeira* spp. em bovinos é a Poisson.

Considerando-se a seleção de tourinhos e o ordenamento dos touros com base nos valores genéticos, a distribuição de Poisson apresentou resultados semelhantes aos apresentados pelas distribuições binomial negativa e binomial negativa inflacionada de zero. Dessa forma, sugere-se que ambas as distribuições também são adequadas para avaliação de características de contagem de parasitos uma vez que não há comprometimento no ranqueamento e seleção dos animais.

REFERÊNCIAS BIBLIOGRÁFICAS

ADACHI, Y.; MAKITA, K. Time series analysis based on two-part models for excessive zero count data to detect farm-level outbreaks of swine echinococcosis during meat inspections. **Preventive Veterinary Medicine**, v.148, p.49-57, 2017.

AGUERRE, S.; JACQUIET, P.; BRODIER, H. et al. Resistance to gastrointestinal nematodes in dairy sheep: Genetic variability and relevance of artificial infection of nucleus rams to select for resistant ewes on farms. **Veterinary Parasitology**, v.256, p.16-23, 2018.

AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, v.19, p.716-723, 1974.

ALEGRÍA-LÓPEZ, M.A.; RODRÍGUEZ-VIVAS, R.I.; TORRES-ACOSTA, J.F.J. et al. Use of Ivermectin as endoparasiticide in tropical cattle herds generates resistance in gastrointestinal nematodes and the tick *Rhipicephalus microplus* (Acari:Ixodidae). **Journal of Medical Entomology**, v.52, p.215-221, 2015.

AYRES, D.R.; PEREIRA, R.J.; BOLIGON, A.A. et al. Linear and Poisson models for genetic evaluation of tick resistance in cross-bred Hereford x Nellore cattle. **Journal of Animal Breeding and Genetics**, v.130, p.417-424, 2013.

AYRES, D.R.; PEREIRA, R.J.; BOLIGON, A.A. et al. Genetic parameters and investigation of genotype x environment interactions in Nellore x Hereford crossbred for resistance to cattle ticks in different regions of Brazil. **Journal of Applied Genetics**, v.56, p.107-113, 2015.

BARLETT, M.S. The use of transformations. **Biometrics**, v.3, p.39-52, 1947.

BIEGELMEYER, P.; GULIAS-GOMES, C.C.; ROSO, V.M. et al. Tick resistance genetic parameters and its correlations with production traits in Hereford and Braford cattle. **Livestock Science**, v.202, p.96-100, 2017.

BOBBO, T.; PENASA, M.; FINOCCHIARO, R. et al. Alternative somatic cell count traits exploitable in genetic selection for mastitis resistance in Italian Holsteins. **Journal of Dairy Science**, v.101, p.10001-10010, 2018.

BURNHAM, K.P.; ANDERSON, D.R. Multimodel inference: understanding AIC and BIC in model selection. **Sociological Methods and Research**, v.33, p.261-304, 2004.

CARDOSO, F.F.; YOKOO, M.J.; GOMES, C.C.G. et al. Avaliação Genômica de Touros Hereford e Braford. Bagé: Embrapa Pecuária Sul, 2012.

CARDOSO, F.F.; GOMES, C.C.G.; SOLLERO, B.P. et al. Genomic prediction for tick resistance in Braford and Hereford cattle. **Journal of Animal Science**, v.93, p.2693-2705, 2015.

CASTILHO, V.L.P.; GUIZELINI, E.; TURRI, E.S. et al. Exame parasitológico quantitativo das fezes: estudo comparativo entre os métodos de McMaster, Stoll-Hausheer e Kato-Katz. **Revista da Sociedade Brasileira de Medicina Tropical**, v.17, p.209-212, 1984.

CAMPOS, A.M. *Uma abordagem bayesiana para alguns modelos de crescimento na presença de assimetria e heteroscedasticidade*. 2011. 74f. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional). Universidade de São Paulo, São Carlos.

CHARTIER, C.; PARAUD, C. Coccidiosis due to Eimeria in sheep and goats, a review. **Small Ruminant Research**, v.103, p.84-92, 2012.

CHIPETA, M.G.; NGWIRA, B.M.; SIMOONGA, C. et al. Zero adjusted models with applications to analysing helminths count data. **BMC Research Notes**, v.7, p.1-11, 2014.

COSTA, J.V.; SILVEIRA, L.V.A.; DONALÍSIO, M.R. Análise espacial dos dados de contagem com excesso de zeros aplicado ao estudo da incidência de dengue em Campinas, São Paulo, Brasil. **Cadernos de Saúde Pública**, v.32, p.1-14, 2016.

COPPIETERS, W.; MES, T.H.M.; DRUET, T. et al. Mapping QTL influencing gastrointestinal nematode burden in Dutch Holstein-Friesian dairy cattle. **BMC Genomics**, v.10 p.1-13, 2009.

CORDEIRO, G.M.; DEMÉTRIO, C.G.B. Modelo linear generalizado: Modelos especiais. In: CORDEIRO, G.M. & DEMÉTRIO, C.G.B. *Modelos lineares generalizados e extensões*. Piracicaba/SP: Editora ESALQ USP, 2013, p.23-68.

EMILIANO, P.C. *Fundamentos e aplicações dos critérios de informação: Akaike e Bayesiano*. 2009. 105f. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras.

EMILIANO, P.C. *Crítérios de informação: como eles se comportam em diferentes modelos?* 2013. 195f. Tese (Doutorado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras.

FANG, R.; WAGNER, B.D.; HARRIS, J.K.; FILLON, S.A. Application of zero-inflated negative binomial mixed model to human microbiota sequence data. **PeerJ PrePrints**, v.2, p.1-12, 2014.

FENG, C.; WANG, H.; LU, N. et al. Log-transformation and its implications for data analysis. **Shangai Archives of Psychiatry**, v.26, p.105-109, 2014.

GADDIS, K.L.P.; DIKMEN, S.; NULL, D.J. et al. Evaluation of genetic components in traits related to superovulation, in vitro fertilization, and embryo transfer in Holstein cattle. **Journal of Dairy Science**, v.100, p.2877-2891, 2017.

GELMAN, A.; CARLIN, J.B.; STERN, H.S. et al. *Bayesian Data Analysis*. Boca Raton/FL: CRC Press – Taylor & Francis Group, 3 ed., 2014. 640p.

GRISI, L.; LEITE, R.C.; MARTINS, J.R.S. et al. Reassessment of the potential economic impact of cattle parasites in Brazil. **Brazilian Journal of Veterinary Parasitology**, v.23, p.150-156, 2014.

HANSEN, J.; PERRY, B. 1994. *The epidemiology, diagnosis and the control of helminth parasites in ruminants*. International Laboratory for Research on Animal Diseases, Nairobi, Kenya, 2th ed., 1994.

HARRISON, X.A. Using observation-level random effects to model overdispersion in count data in ecology and evolution. **PeerJ PrePrints**, p.1-19, 2014.

HAWKINS, J.A. Economic benefits of parasite control in cattle. **Veterinary Parasitology**, v.46, p.159-173, 1993.

HERTL, J.A.; CHUKKEN, Y.H.; TAUER, L.W. et al. Does clinical mastitis in the first 100 days of lactation predict increased mastitis occurrence and shorter herd life in dairy cows? **Journal of Dairy Science**, v.11, p.2309-2323, 2018.

HOLAND, A.M.; STEINSLAND, I.; MARTINO, S. et al. Animal models and integrated nested Laplace approximations. **G3**, v.3, p.1241-1251, 2013.

JELINSKI, M.; GILLEARD, J.; ROCHELEAU, L. et al. Epidemiology of gastrointestinal nematode infections in grazing yearling beef cattle in Saskatchewan. **The Canadian Veterinary Journal**, v.58, p.1044-1050, 2017.

KORNÁS, S.; SALLÉ, G.; SKALSKA, M. et al. Estimation of genetic parameters for resistance to gastrointestinal nematodes in pure blood Arabian horses. **International Journal of Parasitology**, v.45, p.237-242, 2015.

LAMBERT, D. Zero-inflated Poisson regression, with an application to defects in manufacturing. **Technomet**, v.34, p.1-14, 1992.

LE GALL, V.L.; KLAFKE, G.M.; TORRES, T.T. Detoxification mechanisms involved in ivermectin resistance in the cattle tick, *Rhipicephalus (Boophilus) microplus*. **Scientific Report**, v.8, p.1-10, 2018.

LOPES, L.B.; NICOLINO, R.; CAPANEMA, R.O. et al. Economics impacts of parasitic diseases in cattle. **CAB Reviews**, v.10, p.1-13, 2015.

MAPHOLI, N.O.; MAIWASHE, A.; MATIKA, O. et al. Genome-wide association study of tick resistance in South African Nguni cattle. **Ticks and Tick-borne Diseases**, v.7, p.487-497, 2016.

MAPHOLI, N.O.; MAIWASHE, A.; MATIKA, O. et al. Genetic parameters for tick counts across months for different tick species and anatomical locations in South African Nguni cattle. **Tropical Animal Health and Production**, v.45, p.1201-1210, 2017.

MARTINA, R. KAY, R. VAN MAANEN, R.; RIDDER, A. The analysis of incontinence episodes and other count data in patients with overactive bladder by Poisson and negative binomial regression. **Pharmaceut Statist**, v.14, p.151-160, 2015.

MATHEW, B.; HOLAND, A.M.; KOISTINEN, P. et al. Reparametrization-based estimation of genetic parameters in multi-trait animal model using Integrated Nested Laplace Approximation. **Theoretical and Applied Genetics**, v.129, p.215-225, 2015.

MAY, K.; BRÜGEMANN, K.; YIN, T. et al. Genetic line comparisons and genetic parameters for endoparasite infections and test-day milk production traits. **Journal of Dairy Science**, v.100, p.7330-7344, 2017.

MOREL, N.; SIGNORINI, M.L.; MANGOLD, A.J. et al. Strategic control of *Rhipicephalus (Boophilus) microplus* infestation on beef cattle grazed in *Panicum maximum* grasses in a subtropical semi-arid region of Argentina. **Preventive Veterinary Medicine**, v.144, p.179-183, 2017.

MORRIS, C.A.; AMYES, N.C. Heritability and repeatability of resistance to nematode parasites in commercial beef cattle. **Proceedings of the New Zealand Society of Animal Production**, v.72, p.236-239, 2012.

NAYA, H.; URIOSTE, J.I.; CHANG, Y. et al. A comparison between Poisson and zero-inflated Poisson regression models with an application to number of black spots in Corriedale sheep. **Genetics Selection Evolution**, v.40, p.379-394, 2008.

OLIVEIRA, R.C.; BARROS, S.T.D.; ROSSI, R.M. Aplicação da metodologia bayesiana para o estudo reológico da polpa de uva. **Revista Brasileira de Produtos Agroindustriais**, v.11, p.73-80, 2009.

PASSAFARO, T.L.; CARRERA, J.P.B.; DOS SANTOS, L.L. et al. Genetic analysis of resistance ticks, gastrointestinal nematodes and *Eimeria* spp. In Nellore cattle. **Veterinary Parasitology**, v.210, p.224-234, 2015.

PEREZ, B.C.; SILVA, F.F.; VENTURA, R.V. et al. Count bayesian models for genetic analysis of in vitro embryo production traits in Guzará cattle. **Animal**, v.11, p.1440-1448, 2017.

PINO, F.A. A questão da não normalidade: uma revisão. **Revista de Economia Agrícola**, v. 61, p.17-33, 2014.

R DEVELOPMENT CORE TEAM, 2018. R: a language and environment for statistical computing: R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <<https://www.r-project.org/>>. Acesso em: 2 de fevereiro de 2018.

RESENDE, M.D.V.; AZEVEDO, C.F.; SILVA, F.F. Análise estatística de dados longitudinais: Ajuste de modelos curvilíneos. In: RESENDE, M.D.V.; SILVA, F.F.; AZEVEDO, C.F. *Estatística matemática, biométrica e computacional*. Viçosa/MG: Editora UFV, 1ed., 2014. p.327-368.

RIDOUT, M.; DEMÉTRIO, C.G.B.; HINDE, J. Models for count data with many zeros. In: International Biometric Conference, 1998, Cape Town. 1998.

RODRIGUES, J. Bayesian analysis of zero-inflated distributions. **Communications in Statistics – Theory and Methods**, v.32, p.281-289, 2003.

RODRIGUES-MOTTA, M.; GIANOLA, D.; HERINGSTAD, B. et al. A zero-inflated Poisson model for genetic analysis of the number of mastitis cases in Norwegian Red cows. **Journal of Dairy Science**, v.90, p.5306-5315, 2007.

RODRÍGUEZ-VIVAS, R.I.; GRISI, L.; LÉON, A.A.P. et al. Potential economic impact assessment for cattle parasites in Mexico. Review. **Revista Mexicana de Ciencias Pecuarias**, v.8, p.61-74, 2017.

RUE, H.; MARTINO, S. Approximate Bayesian inference for hierarchical Gaussian Markov random field models. **Journal of the Royal Statistical Society**, v.137, p.3711-3192, 2007.

RUE, H.; MARTINO, S. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations. **Journal of the Royal Statistical Society**, v.71, p.319-392, 2009.

SAE-LIM, P.; GRÓVA, L.; OLESEN, I. et al. A comparison of nonlinear mixed models and response to selection of tick-infestation on lambs. **PLoS One**, v.12, p.1-16, 2017.

SCHWARZ, G. Estimating the dimensional of a model. **Annals of Statistics**, v.6, p.461-464, 1978.

SPEEDING, C.R.W. Effect of a sub-clinical worm burden on the digestive efficiency of sheep. **Journal of Comparative Pathology**, v.64, p.5-14, 1954.

SPIEGELHALTER, D.J.; BEST, N.G.; CARLIN, B.P. et al. Bayesian measures of model complexity and fit. **Journal of Royal Statistical Society B (Statistical Methodology)**, v.64, p.583-639, 2002.

UENO, H.; GONÇALVES, P.C. **Manual para diagnóstico das helmintoses de ruminantes**. 4.ed. Tóquio: Japan International Cooperation Agency, 1998. 143p.

WHARTON, R.H.; UTECH, K.B.W. The relation between engorgement and dropping of *B. microplus* (Canestrini) (Ixodidae) to the assessment of tick numbers on cattle. **Journal Australian Entomology Society**, v.9, p.171-182, 1970.

WANG, X.; YUE, Y.R.; FARAWAY, J.J. *Bayesian regression modeling with INLA*. Boca Raton/FL: CRC Press – Taylor & Francis Group, 2018. 312p.

ZINSSTAG, J.; ANKERS, P.; NIJE, M. et al. Heritability of gastrointestinal nematode faecal egg counts in West African village N'Dama cattle and its relation to age. **Veterinary Parasitology**, v.89, p.71-78, 2000.

MATERIAL COMPLEMENTAR

Anexo A. Estatísticas descritivas⁽¹⁾ para contagem de carrapatos, nematódeos gastrointestinais (NGI) e *Eimeria* spp. diferentes de zero

	<i>n</i>	RGN	Média	sd	Mediana	Mín.	Máx.
Carrapato	6.210	1.665	8,28	9,47	5	1	131
NGI	5.476	1.581	6,54	7,02	4	1	80
<i>Eimeria</i> spp.	3.065	1.164	8,95	18,94	4	1	328

⁽¹⁾*n*, número de dados; RGN, número de animais com registros de nascimento distintos; sd, desvio padrão; Mín., mínimo; Máx., máximo.

Anexo B. Critérios de informação da deviance (DIC) para contagem de carrapatos, nematoides gastrointestinais (NGI) e *Eimeria* spp. em bovinos Nelore obtidos no modelo completo e no modelo simples (sem efeito genético aditivo) sob diferentes distribuições

Distribuições ⁽¹⁾	Modelo Completo	Modelo Simples
Carrapatos		
N	49.698,59	48.799,06
Nlog	-9.859,46	-45.599,53
P	35.198,44	35.217,59
ZIP	36.592,52	36.559,70
NB	38.847,57	38.886,03
ZINB	39.018,86	39.107,87
NGI		
N	43.625,58	43.627,36
Nlog	-7.671,79	-41.335,99
P	30.879,84	30.898,63
ZIP	32.469,40	32.733,58
NB	35.012,74	35.064,05
ZINB	34.999,56	35.054,43
<i>Eimeria</i> spp.		
N	54.808,43	55.506,83
Nlog	-6.198,44	-40.946,68
P	19.087,85	19.103,52
ZIP	20.374,99	20.677,10
NB	23.168,03	23.198,73
ZINB	23.158,80	23.191,99

⁽¹⁾N, normal; Nlog, normal com transformação logarítmica dos dados; P, Poisson; ZIP, Poisson inflacionada de zero; NB, binomial negativa; ZINB, binomial negativa inflacionada de zero.

ANEXO C. Frequências relativas (acima da diagonal) e absolutas (abaixo da diagonal) de tourinhos selecionados em comum, em análises unicaracterísticas sob diferentes distribuições, quando foram selecionados 10% daqueles com melhores valores genéticos para a contagem de carrapatos em cada estação de nascimento

Distribuição ⁽¹⁾	N	Nlog	P	ZIP	NB	ZINB
2010/2011 (n⁽²⁾=34)						
N	-	56%	59%	100%	59%	59%
Nlog	19	-	94%	56%	91%	91%
P	20	32	-	59%	97%	97%
ZIP	34	19	20	-	59%	59%
NB	20	31	33	20	-	100%
ZINB	20	31	33	20	34	-
2011/2012 (n=32)						
N	-	62%	66%	100%	66%	66%
Nlog	20	-	94%	62%	94%	94%
P	21	30	-	66%	100%	100%
ZIP	32	20	21	-	66%	66%
NB	21	30	32	21	-	100%
ZINB	21	30	32	21	32	-
2012/2013 (n=28)						
N	-	68%	71%	100%	71%	71%
Nlog	19	-	93%	68%	89%	89%
P	20	26	-	71%	96%	96%
ZIP	28	19	20	-	71%	71%
NB	20	25	27	20	-	100%
ZINB	20	25	27	20	28	-
2013/2014 (n=26)						
N	-	65%	69%	100%	69%	73%
Nlog	17	-	88%	65%	88%	85%
P	18	23	-	69%	100%	96%
ZIP	26	17	18	-	69%	73%
NB	18	23	26	18	-	96%
ZINB	19	22	25	19	25	-
2014/2015 (n=24)						
N	-	46%	50%	100%	46%	46%
Nlog	11	-	87%	46%	87%	83%
P	12	21	-	50%	96%	92%
ZIP	24	11	12	-	46%	46%
NB	11	21	23	11	-	92%
ZINB	11	20	22	11	22	-
2015/2016 (n=27)						
N	-	67%	70%	100%	70%	70%
Nlog	18	-	92%	67%	89%	89%
P	19	25	-	70%	96%	96%
ZIP	27	18	19	-	70%	70%
NB	19	24	26	19	-	100%
ZINB	19	24	26	19	27	-

⁽¹⁾N, normal; Nlog, normal com transformação logarítmica dos dados; P, Poisson; ZIP, Poisson inflacionada de zero; NB, binomial negativa; ZINB, binomial negativa inflacionada de zero. ⁽²⁾n = número de animais TOP10 na estação de nascimento correspondente

ANEXO D. Frequências relativas (acima da diagonal) e absolutas (abaixo da diagonal) de tourinhos selecionados em comum, em análises unicaracterísticas sob diferentes distribuições, quando foram selecionados 10% daqueles com melhores valores genéticos para a contagem de nematódeos gastrointestinais em cada estação de nascimento

Distribuição ⁽¹⁾	N	Nlog	P	ZIP	NB	ZINB
2010/2011 (n⁽²⁾=27)						
N	-	78%	74%	100%	74%	78%
Nlog	21	-	89%	78%	85%	85%
P	20	24	-	78%	96%	96%
ZIP	20	21	21	-	74%	74%
NB	20	23	26	20	-	100%
ZINB	21	23	26	20	27	-
2011/2012 (n=31)						
N	-	84%	77%	100%	77%	77%
Nlog	26	-	84%	84%	84%	84%
P	24	26	-	77%	100%	100%
ZIP	31	26	24	-	77%	77%
NB	24	26	31	24	-	100%
ZINB	24	26	31	24	31	-
2012/2013 (n=27)						
N	-	85%	85%	100%	81%	81%
Nlog	23	-	92%	85%	92%	92%
P	23	25	-	85%	92%	92%
ZIP	27	23	23	-	81%	81%
NB	22	25	25	22	-	100%
ZINB	22	25	25	22	27	-
2013/2014 (n=26)						
N	-	81%	85%	100%	81%	81%
Nlog	21	-	88%	81%	88%	88%
P	22	23	-	85%	96%	96%
ZIP	26	21	22	-	81%	81%
NB	21	23	25	21	-	100%
ZINB	21	23	25	21	26	-
2014/2015 (n=24)						
N	-	71%	67%	100%	67%	67%
Nlog	17	-	92%	71%	87%	87%
P	16	22	-	67%	96%	96%
ZIP	24	17	16	-	67%	67%
NB	16	21	23	16	-	100%
ZINB	16	21	23	16	24	-
2015/2016 (n=27)						
N	-	85%	85%	100%	85%	85%
Nlog	23	-	96%	85%	92%	92%
P	23	26	-	85%	92%	92%
ZIP	27	23	23	-	85%	85%
NB	23	25	25	23	-	100%
ZINB	23	25	25	23	27	-

⁽¹⁾N, normal; Nlog, normal com transformação logarítmica dos dados; P, Poisson; ZIP, Poisson inflacionada de zero; NB, binomial negativa; ZINB, binomial negativa inflacionada de zero. ⁽²⁾n = número de animais TOP10 na estação de nascimento correspondente

ANEXO E. Frequências relativas (acima da diagonal) e absolutas (abaixo da diagonal) de tourinhos selecionados em comum, em análises unicaracterísticas sob diferentes distribuições, quando foram selecionados 10% daqueles com melhores valores genéticos para a contagem de *Eimeria* spp. em cada estação de nascimento

Distribuição ⁽¹⁾	N	Nlog	P	ZIP	NB	ZINB
2010/2011 (n⁽²⁾=27)						
N	-	41%	37%	100%	48%	48%
Nlog	11	-	81%	41%	85%	85%
P	10	22	-	37%	89%	89%
ZIP	27	11	10	-	48%	48%
NB	13	23	24	13	-	100%
ZINB	13	23	24	13	27	-
2011/2012 (n=31)						
N	-	39%	29%	100%	29%	29%
Nlog	12	-	81%	39%	73%	74%
P	9	25	-	29%	84%	87%
ZIP	31	12	9	-	29%	29%
NB	9	23	27	9	-	97%
ZINB	9	23	26	9	30	-
2012/2013 (n=27)						
N	-	44%	44%	100%	48%	48%
Nlog	12	-	78%	44%	70%	70%
P	12	21	-	44%	89%	89%
ZIP	27	12	12	-	48%	48%
NB	13	19	24	13	-	100%
ZINB	13	19	24	13	27	-
2013/2014 (n=26)						
N	-	54%	50%	100%	61%	61%
Nlog	14	-	77%	54%	73%	73%
P	13	20	-	50%	85%	85%
ZIP	26	14	13	-	61%	61%
NB	16	19	22	16	-	100%
ZINB	16	19	22	16	26	-
2014/2015 (n=24)						
N	-	46%	46%	100%	50%	50%
Nlog	11	-	92%	46%	87%	87%
P	11	22	-	46%	87%	87%
ZIP	24	11	11	-	50%	50%
NB	12	21	21	12	-	96%
ZINB	12	21	21	12	23	-
2015/2016 (n=27)						
N	-	48%	44%	100%	55%	55%
Nlog	13	-	78%	48%	74%	74%
P	12	21	-	44%	81%	81%
ZIP	27	13	12	-	55%	55%
NB	15	20	22	15	-	100%
ZINB	15	20	22	15	27	-

⁽¹⁾N, normal; Nlog, normal com transformação logarítmica dos dados; P, Poisson; ZIP, Poisson inflacionada de zero; NB, binomial negativa ZINB, binomial negativa inflacionada de zero. ⁽²⁾n = número de animais TOP10 na estação de nascimento correspondente.