

UNIVERSIDADE FEDERAL DE ALFENAS

SADITT ROCIO ROBLES COLONIA

**EXPLORANDO A VARIAÇÃO GENÉTICA DE CARACTERÍSTICAS CONTÍNUAS E
BINÁRIAS DE BOVINOS LEITEIROS SOB ABORDAGEM BAYESIANA**

Alfenas/MG

2020

SADITT ROCIO ROBLES COLONIA

**EXPLORANDO A VARIAÇÃO GENÉTICA DE CARACTERÍSTICAS CONTÍNUAS E
BINÁRIAS DE BOVINOS LEITEIROS SOB ABORDAGEM BAYESIANA**

Dissertação apresentada ao Programa de Pós-Graduação em Estatística Aplicada e Biometria, área de concentração em Estatística Aplicada e Biometria da Universidade Federal de Alfenas, MG, como parte dos requisitos para a obtenção do título de Mestre. Linha de Pesquisa: Modelagem Estatística e Estatística Computacional.

Orientador: Prof. Dra. Juliana Petrini.

Coorientador: Prof. Dr. Denismar Alves Nogueira.

Alfenas/MG

2020

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema de Bibliotecas da Universidade Federal de Alfenas

Robles Colonia, Saditt Rocio.
R666e Explorando a variação genética de características contínuas e binárias de bovinos leiteiros sob abordagem Bayesiana. / Saditt Rocio Robles Colonia. - Alfenas/MG, 2020.
87f.: il. --

Orientadora: Juliana Petrini.
Dissertação (Mestrado em Estatística Aplicada e Biometria) -
Universidade Federal de Alfenas, 2020.
Bibliografia.

Herdabilidade. 2. Incidência de mastite subclínica. 3. Modelo de limiar. 4. Tendência genética. 5. Teor de caseína. I. Petrini, Juliana. II. Título.

CDD-519.5

**EXPLORANDO A VARIAÇÃO GENÉTICA DE CARACTERÍSTICAS CONTÍNUAS E BINÁRIAS
DE BOVINOS LEITEIROS SOB ABORDAGEM BAYESIANA**

A Banca examinadora abaixo-assinada aprova a Dissertação apresentada como parte dos requisitos para a obtenção do título de Mestre em Estatística Aplicada e Biometria pela Universidade Federal de Alfenas. Área de concentração: Estatística Aplicada e Biometria.

Aprovada em: 17 de fevereiro de 2020.

Profa. Dra. Juliana Petrini

Instituição: UNIFAL-MG

Prof. Dr. Gerson Barreto Mourão

Instituição: ESALQ-USP

Prof. Dr. Luiz Alberto Beijo

Instituição: UNIFAL-MG



Documento assinado eletronicamente por **Luiz Alberto Beijo, Coordenador do Programa de Pós-Graduação em Estatística Aplicada e Biometria**, em 17/02/2020, às 11:41, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Gerson Barreto Mourão, Usuário Externo**, em 17/02/2020, às 11:46, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Juliana Petrini, Professor do Magistério Superior**, em 17/02/2020, às 11:51, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://sei.unifal-mg.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0265039** e o código CRC **2B11E1FB**.

*Dedico este trabalho a minha
família.*

AGRADECIMENTOS

Primeiramente a Deus, por sua graça e misericórdia porque ele estive ao meu lado desde o primeiro momento da minha vida ajudando-me a nunca desistir e por me ter dado forças em situações difíceis.

Agradeço ao meu pais, Manuel Luis Robles Martín e Esperanza Neli Colonia Mellisho, a minha irmã Elianne Anabel Robles Colonia e a meus tios que desde o começo eles confiaram em mim e me deram seu apoio e amor incondicional.

A Juliana Petrini e Denismar Alves Nogueira, meus orientadores por sua confiança, dedicação, paciência e entregue no apoio ao desenvolvimento e revisão constante do presente tese.

Agradeço meus professores por sua paciência e dedicação em formar novos mestres e aos funcionários do UNIFAL/MG por toda a atenção e disponibilidade brindada.

Ao Prof. Dr. Gerson Barreto Mourão, Prof. Dr. Paulo Fernando Machado e à Clínica do Leite (ESALQ-USP) pela concessão dos dados utilizados neste estudo, pela parceria e apoio no desenvolvimento da linha de pesquisa em Estatística aplicada ao melhoramento genético de bovinos leiteiros, e pela confiança depositada.

Ao Prof. Dr. Luiz Lehmann Coutinho e ao Laboratório de Biotecnologia (ESALQ-USP) pelo suporte computacional para a realização das análises desta pesquisa.

Agradeço muito à família peruana pelo apoio incondicional na minha estadia em Alfenas: Maurício e Ângela e a seus filhos Ítalo e Lucas. Assim mesmo agradecer a Bruno Novaes, Frank Lázaro e Douglas Ferreira pelo carinho e consideração como parte de sua família estando sempre em todo tempo. Além disso, eu dou graças a Deus pela vida do meus pastores Francisco e Matheus da Igreja Pentecostal “Água da Rocha” da cidade de Alfenas.

Finalmente, agradeço aos meus amigos que participaram direta e indiretamente na minha formação acadêmica.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Os levarei sempre no meu coração, OBRIGADA POR TUDO!

“Render-se não é uma opção”

[\(PRIETO, 2016\)](#)

RESUMO

A eficiência econômica em bovinocultura leiteira está associada à capacidade produtiva e reprodutiva do animal. Dessa forma, para minimizar as perdas no sistema de produção e maximizar o lucro; os critérios de seleção devem incluir características de produção e composição do leite, e adicionalmente características relacionadas à saúde e fertilidade dos animais. No entanto, visto que geralmente fenótipos associados à saúde e reprodução apresentam distribuição não normal ou desconhecida, o uso do modelo de limiar sob abordagem Bayesiana apresenta-se como adequado à avaliação genética dos mesmos. Assim, o objetivo deste trabalho foi estimar os componentes de variância, parâmetros genéticos e prever os valores genéticos para características contínuas e binárias de bovinos leiteiros da raça Holandesa mediante o uso de modelo linear e de limiar binário sob abordagem Bayesiana. Para tanto, foram utilizadas 64.657 mensurações de produção de leite (PROD, em kg/dia), teor de caseína (CAS, g/100g de leite), incidência de mastite subclínica (MSC) e precocidade (PREC) de 5.439 vacas de primeira à sétima lactação e com dias em lactação entre 5 e 365, pertencentes a quatro fazendas. A análise genética foi baseada no modelo linear (para PROD e CAS) e de limiar (para MSC e PREC), os quais incluíram os efeitos de grupo de contemporâneas e a covariável de efeito cúbico de dias em lactação, e os efeitos aditivos genéticos, permanente de ambiente e residual. Os componentes de variância e valores genéticos, bem como os demais parâmetros desconhecidos do modelo, foram obtidos sob *prioris* não informativas e informativas, utilizando o amostrador de Gibbs para a obtenção das marginais das distribuições *a posteriori*. As médias *a posteriori* das herdabilidades para PROD, CAS, MSC e PREC com *priori* não informativa foram iguais a 0,142, 0,399, 0,134 e 0,377, respectivamente, enquanto para *priori* informativa esses valores foram de 0,131, 0,271, 0,134 e 0,436. Ganhos genéticos significativos foram observados apenas para CAS e PREC, com valores iguais a 0,014 g/100g de leite e 0,032 unidade ao ano, respectivamente. Esses resultados destacam que os modelos lineares e limiares são adequados na estimação dos componentes de variância e parâmetros genéticos das características estudadas.

Palavras-chave: Herdabilidade. Incidência de mastite subclínica. Modelo de limiar. Tendência genética. Teor de caseína.

ABSTRACT

The economic efficiency in dairy cattle is associated to the animal's productive and reproductive capacity. This way, to reduce the losses in the production system and to increase the profit; selection criteria should include production and milk composition traits plus traits related to animal health and fertility. However, since health and reproduction traits are generally associated to non-normal or unknown distributions, the use of threshold model under Bayesian approach is adequate to the genetic evaluation of such phenotypes. Therefore, the aim in this study was to estimate variance components, genetic parameters and to predict breeding values for continuous and discrete distributed traits of Holstein cattle by using linear and threshold models under Bayesian approach. A total of 64,657 records for milk yield (MY, kg/day), casein content (CP, g/100g milk), subclinical mastitis incidence (SMI) and precocity (PREC) of 5,439 first to seventh parity cows, with days in milk between 5 and 365, and belonging to four farms were used. Genetic analysis was based on a linear model (for MY and CN continuous traits) and on a threshold model (for SMI and PREC binary traits), which included the effects of contemporary groups, days in milk, lactation number, and the additive genetic, permanent environmental and residual effects. Variance components, breeding values as well as other unknown parameters of the model, were obtained under non-informative and informative priors, by using Gibbs sampler to obtain the marginal posterior distributions. The heritability posterior means for MY, CP, SMI and PREC, with non-informative prior were 0.142, 0.399, 0.134 e 0.377, respectively. For informative prior, these values were 0.131, 0.271, 0.134 e 0.436. Significant genetic gains were only observed for CP and PREC, with values equal to 0.014 g/100g of milk and 0.032 unit by year, respectively. These results highlight that linear and threshold models are adequate to estimate the variance components and genetic parameters of the studied traits.

Key-words: Casein content. Genetic trend. Heritability. Subclinical mastitis incidence. Threshold model.

LISTA DE FIGURAS

<p>Figura1 – Frequência percentual das características binárias: incidência de mastite subclínica (MSC) (a) e Precocidade (PREC) (b). Incidência de mastite subclínica igual a um (1) para contagem de células somáticas (CCS) menor que 200.000 células/ml de leite, e igual a zero (0) caso contrário; PREC igual a um (1) para idade ao primeiro parto menor que 730 dias, e igual a zero (0) caso contrário.</p>	50
<p>Figura2 – Tendências genéticas preditas e valores genéticos para as características contínuas de produção de leite (PROD) (a), teor de caseína (CAS) (b), incidência de mastite subclínica (MSC) (c) e a precocidade (PREC) (d), período de 2001 a 2014.</p>	63
<p>Figura3 – Gráficos de densidade das cadeias <i>a posteriori</i> dos parâmetros estimados $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_P^2$, h^2, c^2 e Rep resultante da característica de produção de leite, utilizando <i>a priori</i> não informativa, período 2012 a 2016.</p>	79
<p>Figura4 – Gráficos de densidade das cadeias <i>a posteriori</i> dos parâmetros estimados $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_P^2$, h^2, c^2 e Rep resultante da característica de produção de leite, utilizando <i>a priori</i> informativa, período 2012 a 2016.</p>	80
<p>Figura5 – Gráficos de densidade das cadeias <i>a posteriori</i> dos parâmetros estimados $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_P^2$, h^2, c^2 e Rep resultante da característica de caseína, utilizando <i>a priori</i> não informativa, período 2012 a 2016.</p>	81
<p>Figura6 – Gráficos de densidade das cadeias <i>a posteriori</i> dos parâmetros estimados $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_P^2$, h^2, c^2 e Rep resultante da característica de caseína, utilizando <i>a priori</i> informativa, período 2012 a 2016.</p>	82
<p>Figura7 – Gráficos de densidade das cadeias <i>a posteriori</i> dos parâmetros estimados $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_P^2$, h^2, c^2 e Rep resultantes da característica de incidência de mastite subclínica (MSC), utilizando <i>a priori</i> não informativa, período 2012 a 2016.</p>	83
<p>Figura8 – Gráficos de densidade das cadeias <i>a posteriori</i> dos parâmetros estimados $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_P^2$, h^2, c^2 e Rep resultante da característica de incidência de mastite subclínica (MSC), utilizando <i>a priori</i> informativa, período 2012 a 2016.</p>	84

Figura9 – Gráficos de densidade das cadeias <i>a posteriori</i> dos parâmetros estimados $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_P^2$ e h^2 resultantes da característica de precocidade (PREC) com <i>priori</i> não informativa, período 2012 a 2016.	85
Figura10 –Gráficos de densidade das cadeias <i>a posteriori</i> dos parâmetros estimados $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_P^2$ e h^2 resultantes da característica de precocidade (PREC) com <i>priori</i> informativa, período 2012 a 2016.	86
Figura11 –Gráficos do traço e densidade das cadeias <i>a posteriori</i> dos parâmetros β_0 , β_1 e τ resultantes do ajuste do modelo linear simples aos dados de valores genéticos para as características de produção de leite (a), teor de caseína (b), incidência de mastite (c) e precocidade (d) período de 2001 a 2014.	87

LISTA DE TABELAS

Tabela1 –	Resumo das <i>prioris</i> informativas utilizadas para as variâncias iniciais aditivas (σ_γ), permanente de ambiente (σ_{pe}) e a variância residual (σ_ε)	45
Tabela2 –	Estatísticas descritivas obtidos para as características contínuas e binárias com seu respectivo número de mensurações (N), número de vacas (NV), média, desvio-padrão (DP), coeficiente de variação (CV %), valores mínimos (MIN) e máximos (MAX).	49
Tabela3 –	Estimativas (média <i>a posteriori</i>) dos componentes de variância e parâmetros (componentes de variância e parâmetros genéticos) obtidos com <i>priori</i> não informativa (NI) e informativa (INF) com seus respectivos intervalos ($HPD_{95\%}$), critérios de convergência e erros de Monte Carlo (EMC) para a característica contínua de produção de leite (PROD), período 2012 a 2016.	52
Tabela4 –	Estimativas (média <i>a posteriori</i>) dos componentes de variância e parâmetros (componentes de variância e parâmetros genéticos) obtidos com <i>priori</i> não informativa (NI) e informativa (INF) com seus respectivos intervalos ($HPD_{95\%}$), critérios de convergência e erros de Monte Carlo (EMC) para a característica contínua de teor de caseína (CAS), período 2012 a 2016.	54
Tabela5 –	Estimativas (média <i>a posteriori</i>) dos parâmetros (componentes de variância e parâmetros genéticos) obtidos com <i>prioris</i> não informativa (NI) e informativa (INF) com seus respectivos intervalos ($HPD_{95\%}$), critérios de convergência e erros de Monte Carlo (EMC) para a característica binária de incidência de mastite subclínica (MSC), período 2012 a 2016.	57
Tabela6 –	Estimativas (média <i>a posteriori</i>) dos parâmetros (componentes de variância e parâmetros genéticos) obtidos com <i>prioris</i> não informativa (NI) e informativa (INF) com seus respectivos intervalos ($HPD_{95\%}$), critérios de convergência e erros de Monte Carlo (EMC) para a característica binária de precocidade (PREC), período 2012 a 2016. . .	59

Tabela7 –	Critério de informação do desvio (DIC) do modelo animal com <i>priori</i> não informativa (NI) e informativa (INF) para das características avaliadas de produção (PROD), teor de caseína (CAS), incidência de mastite subclínica (MSC) e precocidade (PREC), período 2012 a 2016.	60
Tabela8 –	Média <i>a posteriori</i> dos parâmetros da regressão linear dos valores genéticos sobre o ano de nascimento, com seus respectivos intervalos de credibilidade ($HPD_{95\%}$) e critérios de convergência.	62

LISTA DE ABREVIATURAS E SIGLAS

- EMM** - Equação dos modelos mistos
- IBGE** - Instituto Brasileiro de Geografia e Estatística
- PROD** - Produção de leite
- CAS** - Teor de caseína
- PREC** - Precocidade
- MSC** - Incidência de mastite subclínica
- CCS** - Contagem de células somáticas
- MCMC** - Monte Carlo via Cadeias de Markov
- DIC** - Critério de desvio de informação
- IPP** - Idade ao primeiro parto
- BLUE** - Melhor Estimador Linear não-viesada
- BLUP** - Melhor Predição Linear não-viesada
- GC** - Grupo contemporâneo

SUMÁRIO

1	INTRODUÇÃO	15
1.1	OBJETIVOS	16
1.1.1	Objetivo Geral	16
1.1.2	Objetivos Específicos	16
2	REFERENCIAL TEÓRICO	18
2.1	MELHORAMENTO GENÉTICO EM BOVINOS LEITEIROS	18
2.1.1	Objetivos e critérios de seleção	19
2.1.2	Componentes de variância genética e parâmetros genéticos	21
2.1.3	Ganho genético e tendência genética	23
2.2	MODELO LINEAR MISTO	24
2.3	MODELO MISTO NO MELHORAMENTO GENÉTICO ANIMAL	27
2.3.1	Modelo misto com medidas repetidas	28
2.4	ABORDAGEM BAYESIANA NO CONTEXTO DO MELHORAMENTO ANIMAL	29
2.4.1	Abordagem Bayesiana na estimação de componentes de variância e efeitos genéticos	31
2.4.2	Modelo de limiar binário	35
2.4.3	Algoritmo MCMC: Amostrador de Gibbs	37
2.4.4	Diagnóstico de convergência	39
2.4.5	Critério de Informação do Desvio (DIC)	40
3	MATERIAIS E MÉTODOS	41
3.1	ANIMAIS E DADOS	41
3.2	ANÁLISE GENÉTICA	42
3.2.1	Análise genética para características contínuas	43
3.2.2	Análise genética para características binárias	43
3.2.3	Estimação dos componentes de variância e valores genéticos	44
3.2.3.1	Especificação da <i>priori</i> não informativa	44
3.2.3.2	Especificação da <i>priori</i> informativa	44
3.2.4	Avaliação da tendência genética	47
4	RESULTADOS E DISCUSSÃO	49

4.1	ANÁLISE EXPLORATÓRIA DOS DADOS	49
4.2	COMPONENTES DE VARIÂNCIA E PARÂMETROS GENÉTICOS	51
4.3	TENDÊNCIA GENÉTICA	61
5	CONCLUSÕES	65
	REFERÊNCIAS	66
	APÊNDICES	79

1 INTRODUÇÃO

Em 2017, a produção mundial de leite bovino foi igual a 675,6 bilhões de litros. Neste período, o principal produtor de leite foi a União Europeia com 30,8% da produção, seguida pela Ásia (30,2%), América (27,3%) e com contribuições menores da África, Oceania e outras regiões da Europa (RESENDE *et al.*, 2019). Trata-se de um valioso produto agrícola, sendo produzido e consumido em grande parte dos países, principalmente por conter uma variedade de nutrientes, como as proteínas (caseína), que são de grande valor para a indústria (BRASIL *et al.*, 2015).

No melhoramento genético de gado leiteiro, a eficiência econômica está associada à capacidade produtiva e reprodutiva do animal, com a produção de leite atuando como importante critério de seleção na avaliação genética (PÉREZ-CABAL *et al.*, 2009). No entanto, atualmente, características associadas à fertilidade e saúde do gado leiteiro também estão sendo consideradas como critérios de seleção para minimizar as perdas no sistema de produção. O propósito da inclusão dessas características na seleção genética tem sido melhorar a conformação da vaca e aumentar a sua longevidade, reduzindo os custos referentes ao uso de medicamentos e reposição de animais (PETRINI, 2016).

O ganho genético nestas características depende da predição acurada do mérito genético dos indivíduos sob seleção e, para tanto, da estimação dos componentes de variância, parâmetros genéticos e efeitos ambientais associados a tais características. Nesse contexto, a abordagem Bayesiana tem-se apresentado como metodologia adequada à estimação na avaliação genética, visto que fornece algumas vantagens práticas. A análise Bayesiana proporciona estimativas confiáveis e possibilita a construção de intervalos de credibilidade mais precisos sobre a incerteza de todos os parâmetros do modelo, baseando-se nos dados disponíveis e na distribuição *a priori* que permitem a incorporação de informações passadas, caso existam, que enriquecem o processo de inferência (GIANOLA; FERNANDO, 1986; RESENDE; SILVA; AZEVEDO, 2014; VILLEMEREUIL, 2019).

Além disso, essa abordagem possibilita utilizar pequenos e grandes conjuntos de dados, envolvendo modelos simples e complexos na análise de dados normalmente distribuídos, com distribuição não-normal, ou com distribuição desconhecida. Especificamente para dados discretos, destaca-se o modelo de limiar (*threshold model*) (HADFIELD; NAKAGAWA, 2010; MRODE, 2014) o qual baseia-se nos modelos generalizados e que relaciona a variável

categórica observada à uma variável subjacente em escala contínua (HADFIELD; NAKAGAWA, 2010). Esta metodologia é adequada para a modelagem de características de natureza binária (sucesso ou fracasso) e torna-se importante no estudo da variação genética de características de interesse econômico, como as associadas à saúde e reprodução, que geralmente não possuem distribuição normal. Exemplos da utilização desse modelo na avaliação genética de bovinos leiteiros foram reportados em Faraji-Arough *et al.* (2015), Irano *et al.* (2014), Ødegård *et al.* (2010) e Oliveira *et al.* (2015).

No entanto, ainda não foram desenvolvidos estudos utilizando o modelo de limiar sob abordagem Bayesiana para características com distribuição desconhecida e com base em *prioris* informativas, na análise genética de populações de bovinos leiteiros manejados sob clima tropical.

1.1 OBJETIVOS

Nesta seção são apresentados os objetivos deste estudo.

1.1.1 Objetivo Geral

Estimar os componentes de variância, parâmetros genéticos e prever os valores genéticos para características de produção, qualidade do leite, saúde e fertilidade de bovinos leiteiros mediante o uso de modelos mistos lineares e de limiar sob abordagem Bayesiana.

1.1.2 Objetivos Específicos

- a) Estimar os componentes de variância e prever os efeitos aleatórios aditivos, para produção de leite e teor de caseína, utilizando modelo linear de repetibilidade sob abordagem Bayesiana, considerando distribuições *a priori* informativas e não-informativas;
- b) Estimar os componentes de variância e prever os efeitos aleatórios aditivos para

incidência de mastite e precocidade utilizado modelo de limiar sob abordagem Bayesiana, considerando distribuições *a priori* informativas e não-informativas;

- c) Obter as estimativas de herdabilidade e repetibilidade para as características em estudo;
- d) Estimar o ganho genético e tendência genética para as características em estudo;
- e) Gerar informações que auxiliem na definição de estratégias de seleção genética visando ao progresso genético nessas características.

2 REFERENCIAL TEÓRICO

Nesta seção apresenta-se o referencial teórico sobre melhoramento genético, modelo linear misto e de limiar binário sob abordagem Bayesiana na estimação dos componentes de variância, parâmetros genéticos e predição dos valores genéticos.

2.1 MELHORAMENTO GENÉTICO EM BOVINOS LEITEIROS

A bovinocultura leiteira é considerada uma das atividades mais importantes do setor agropecuário brasileiro (ARAÚJO *et al.*, 2019). Em 2018, a produção leiteira no Brasil foi de 33,8 bilhões de litros com produção média de 1.963 litros/vaca/ano, sendo o sexto maior produtor de leite no mundo (IBGE, 2018). Nos últimos anos, a produção de leite bovino está aumentando consideravelmente, mesmo com uma redução no número de animais em lactação, indicando um aumento na especialização do setor (ZOCCAL, 2018). Esse aumento se deve a melhorias no manejo, sanidade, nutrição e genética dos rebanhos que afetam a produtividade e qualidade do leite. Contudo, o sistema de produção no Brasil ainda é considerado de baixa rentabilidade para o produtor rural, sendo iminente a adoção de ferramentas que visem o aumento da eficiência do setor, como o melhoramento genético.

O melhoramento genético é um processo avaliativo cujo objetivo é a seleção de animais com maior mérito genético para atuarem como pais da geração seguinte, aproveitando a variação genética existente entre os animais de uma mesma espécie (PEÑA; RUIZ, 2018). A avaliação genética é o processo pelo qual o valor genético dos indivíduos é predito para uma ou mais características de interesse econômico (por exemplo, produção de leite, porcentagem de proteína, gordura, sólidos totais em bovinos, etc.), tornando-se a principal ferramenta para o progresso genético das próximas gerações (ARANGO; ECHEVERRI, 2014; MUÑOZ; GONZÁLES, 2016; MUÑOZ *et al.*, 2012).

2.1.1 Objetivos e critérios de seleção

Objetivos de seleção são características de importância econômica de acordo com um determinado sistema de produção (CARDOSO *et al.*, 2004) enquanto critérios de seleção representam o conjunto de informações ou características nas quais baseia-se a seleção genética, visando-se avaliar e classificar os candidatos à seleção para o objetivo do melhoramento (VAYEGO, 2007). Estes objetivos geralmente estão associados à produção de leite e derivados, à reprodução e saúde da vaca, bem como eficiência alimentar e longevidade. Nisto, os critérios envolvem características como produção de leite, teor de sólidos (proteína, gordura, lactose), incidência de mastite e demais doenças, consumo alimentar, ganho de peso diário, vida produtiva, entre outros.

A redução de custos em uma propriedade leiteira está fortemente ligada à redução na incidência de doenças, como a mastite, e ao desempenho reprodutivo das vacas, de forma que a relevância destes objetivos na seleção genética tem aumentado. A mastite corresponde a uma inflamação na glândula mamária acompanhada de mudanças físicas, químicas e fisiológicas no leite e no tecido mamário (SAMAD, 2008). Essa doença pode ser classificada como subclínica, clínica e crônica. Na mastite subclínica não são observadas alterações macroscópicas. No entanto, há mudanças químicas e microbiológicas no leite, além de um aumento na contagem de células somáticas. Por sua vez, a mastite clínica é caracterizada por respostas inflamatórias severas, resultando em mudanças na secreção de leite, mudanças visíveis no tecido mamário e, em alguns casos, hipertermia, prostração e tremores musculares (PRESTES; FILAPPI; CECIM, 2002). Trata-se da doença de maior prevalência e custo dentre os rebanhos leiteiros. Os microrganismos comumente associados a essa patologia são dos gêneros *Staphylococcus*, *Streptococcus* e coliformes. Os patógenos do úbere variam entre regiões e também entre os tipos de mastite - clínica e subclínica.

A mastite em sua forma subclínica é importante porque é mais prevalente do que a forma clínica, usualmente precede a forma clínica, tem longa duração, é de difícil detecção, e reduz a produção de leite com efeitos adversos sobre a sua qualidade (SEEGERS; FOURICHON; BEAUDEAU, 2003). Estima-se que cerca de 70% das perdas econômicas associadas à incidência de mastite sejam resultado da mastite subclínica (VARSHNEY; NARESH, 2004). Segundo Abdel-Rady e Sayed (2009), citando a Harmon (1994), tais perdas são devido à redução de produção de leite, ao descarte de leite anormal ou retido de vacas

tratadas com antibióticos, queda na qualidade do leite e no preço devido à alta contagem bacteriana e de células somáticas, custos referentes ao tratamento, aumento do risco de mastite subsequente e problemas associados à resíduos de antibióticos no leite e derivados. No Brasil, a prevalência de mastite subclínica varia entre 14 e 65% de acordo com a região estudada (OLIVEIRA *et al.*, 2011).

O desempenho reprodutivo também é um fator de alto impacto nos custos de produção e determina grande parte do lucro de um sistema produtivo (M-ROCHA *et al.*, 2012). Especificamente a idade ao primeiro parto e, conseqüentemente, a precocidade da vaca, é um dos muitos fatores que afetam o lucro de uma propriedade, visto a sua relação direta com o custo de criação das novilhas e efeito sobre o desempenho futuro do animal. Estima-se que o custo de criação de uma novilha da raça Holandesa no Brasil seja de aproximadamente R\$ 3.000,00 (ano de 2009) e tal custo está predominantemente associado à alimentação, mão-de-obra e despesas diversas (SANTOS; LOPES, 2014). Dessa forma, um primeiro parto precoce pode ser benéfico ao reduzir os custos de criação e ao retornar o investimento mais rapidamente (TOZER; HEINRICHS, 2001).

A relação entre idade ao primeiro parto e produção de leite foi reportada por Berry e Cromie (2009), Nilforooshan e Edriss (2004) e Pirlo, Miglior e Speroni (2000). Berry e Cromie (2009), a partir de dados de 60.802 vacas da raça Holandesa Friesiana na Irlanda, no qual estimaram uma redução linear na produção de leite, gordura e proteína na primeira lactação de 55,5 kg, 0,6 kg e 2,3 kg, respectivamente, a cada adiantamento em um mês na idade ao primeiro parto. Por sua vez, Nilforooshan e Edriss (2004) observaram uma relação não-linear entre IPP e produção de leite na primeira lactação de 12.082 vacas Holandesas no Irã, com máxima produção associada à IPP igual a 24 meses e efeito negativo sobre produção para IPP menor do que 21 meses. Similarmente, Pirlo, Miglior e Speroni (2000) documentaram um aumento na produção de leite na lactação com IPP entre 20 e 36 meses de idade, utilizando 1.048.942 registros de vacas da raça Holandesa de rebanhos italianos. Recentemente, Hutchison *et al.* (2017) verificaram uma redução de 2,4 meses na IPP de vacas da raça Holandesa nos EUA desde 2006 associada a um aumento na produção e fertilidade, exceto para IPP entre 18 e 20 meses. A influência da IPP sobre a produção de leite pode ser explicada pelo tamanho e peso corporal da vaca à puberdade (ETTEMA; SANTOS, 2004; SEJRSEN; PURUP, 1997) e pelo manejo nutricional das novilhas no pós-parto (BERRY; CROMIE, 2009). A idade ao primeiro parto pode ainda influenciar à vida produtiva, ocorrência de distocia no parto e viabilidade dos

bezerros (PIRLO; MIGLIOR; SPERONI, 2000).

A redução na idade ao primeiro parto tem um efeito positivo sobre o progresso genético do rebanho, dado que reduz-se o intervalo de gerações e os touros podem ser avaliados mais precocemente (PIRLO; MIGLIOR; SPERONI, 2000).

2.1.2 Componentes de variância genética e parâmetros genéticos

Os componentes de variância genética associados às características de interesse compõem as estimativas dos parâmetros genéticos (como coeficiente de herdabilidade e correlações genéticas) e, conseqüentemente, têm papel importante sobre o delineamento dos programas de melhoramento genético animal e para a predição da resposta à seleção. Adicionalmente, a estimação dos componentes de variância é necessária à predição do valor genético para o sucesso do melhoramento animal. Os componentes de variância e parâmetros genéticos podem ser estimados por diferentes métodos, entre os quais citam-se a máxima verossimilhança restrita (REML) e a abordagem Bayesiana (PAIVA *et al.*, 2019).

No melhoramento genético, a variabilidade entre os indivíduos é própria de cada característica de interesse e estimada pela variância de seus valores fenotípicos (P), os quais são obtidos a partir da mensuração direta no indivíduo e que geralmente se aproximam razoavelmente a uma curva normal (ELER, 2017; GASQUE, 2008). Tais fenótipos são utilizados para a predição dos valores genéticos (não observável), por meio do modelo genético $P = G + E + E \times G$, em que P é o valor fenotípico estimado a partir de dados mensurados dos indivíduos, e que é resultante da ação do genótipo G , do ambiente E e da interação entre eles $G \times E$. Analogamente, a variância fenotípica (VP) é composta pela variância genotípica (VG) ou (σ_G^2) e pela variância atribuída aos desvios no fenótipo originados pela influência do ambiente (VE) (FALCONER, 1987). Esse modelo tem sido muito utilizado em estudos acerca da estrutura genética (FALCONER, 1987; QUIJANO; ECHEVERRI, 2015), possibilitando o desdobramento da variância genética (VG) em outros três componentes:

$$\sigma_G^2 = \sigma_A^2 + \sigma_D^2 + \sigma_I^2, \quad (2.1)$$

sendo σ_A^2 a variância genética aditiva correspondente à soma dos efeitos de todos os alelos; σ_D^2 é a variância genética de dominância, que surge da propriedade do domínio entre os alelos

de um mesmo locu, ou seja, da interação entre eles; e σ_I^2 a variância genética epistática que é originada pela interação de alelos de locus diferentes. A variação aditiva apresenta-se como a mais importante na seleção de animais visando o melhoramento genético, pois representa a fração herdável da variância genética (MRODE, 2014).

A partir dos componentes de variância são estimados os parâmetros genéticos (herdabilidade e repetibilidade), os quais são indicadores populacionais das características avaliadas, utilizados na estruturação dos programas de melhoramento genético e dos procedimentos de avaliação genética. Dentre os parâmetros genéticos destaca-se a herdabilidade (no sentido restrito, $h^2 = \sigma_A^2/\sigma_P^2$), que representa a proporção da variância fenotípica devida ao efeito genético aditivo, e indica a validade do fenótipo como preditor do valor genético (LOPES *et al.*, 2005). Os valores da herdabilidade podem variar de 0 a 1, de modo que valores baixos indicam uma predominância do efeito do ambiente sobre a variação fenotípica observada, enquanto que valores próximos a 1 sugerem que grande parte da variação é de natureza genética (ELER, 2017; GUTIERREZ, 2010).

A repetibilidade ($Rep = \frac{\sigma_A^2 + \sigma_E^2}{\sigma_P^2}$) é a proporção da variância total em múltiplas medidas de uma característica devida às diferenças entre indivíduos; a qual possibilita quantificar em que extensão o desempenho ou comportamento de um indivíduo permanece constante ao longo do tempo (DOHM, 2002). Pode ainda ser definida como a correlação entre medidas repetidas de uma característica nos mesmos indivíduos (FALCONER; MACKAY, 1996). Segundo Falconer e Mackay (1996), as medidas repetidas devem ter variâncias homogêneas e estarem sob o mesma ação genética (isto é, a correlação genética entre as medidas deve ser igual a um).

A repetibilidade é dita como o limite superior da herdabilidade, dado que inclui as fontes de variação genética e ambiental enquanto a herdabilidade inclui apenas as diferenças genéticas entre os indivíduos (FALCONER; MACKAY, 1996). Pode também ser importante em determinar o quão efetiva a seleção natural será em mudar uma característica ao longo do tempo (BOAKE, 1989; HUEY; DUNHAM, 1987). Adicionalmente, valores alto de repetibilidade indicam que os indivíduos tendem a apresentar um desempenho consistente e, portanto, não há razão prática na obtenção de múltiplas medidas (FALCONER; MACKAY, 1996). Nesse sentido, a repetibilidade possibilita verificar se as múltiplas medidas se referem, geneticamente, à mesma ou diferentes características (DOHM, 2002). Valores baixos de h^2 e correspondente variação ambiental alta podem reduzir a repetibilidade entre medidas sucessivas; contudo, baixa repetibilidade pode ter diversas origens genéticas ou ambientais

(DOHM, 2002). Similarmente à herdabilidade, os valores de repetibilidade oscilam entre 0 e 1.

Quando medidas são tomadas repetidamente sobre um mesmo indivíduo é possível ainda estimar a variância devida ao efeito permanente de ambiente. Este é um efeito ambiental que contribui permanentemente ao fenótipo de um indivíduo, sendo constante entre as medidas repetidas (KRUUK; HADFIELD, 2007) e cumulativo ao longo do tempo (SCHAEFFER, 2011). A variação do efeito permanente de ambiente pode ser devida a diferenças constantes quanto às condições experimentadas pelos indivíduos ao longo da vida (ex. qualidade do território) ou a efeitos de longo prazo ocasionados por condições experimentadas em um estágio crítico da vida (ex. desenvolvimento inicial) (KRUUK; HADFIELD, 2007). Efeitos permanentes de ambiente podem também estar associados a efeitos epigenéticos, que são mudanças na expressão gênica causadas por modificações no DNA e histonas (responsáveis pela modulação de genes regulatórios), sem alterar a sequência do DNA (BIRD, 2002; WOLFFE; GUSCHIN, 2000). Um exemplo citado por Kelleher *et al.* (2016) do trabalho de Berry *et al.* (2008) são estímulos ao ambiente intrauterino, que podem impactar o desenvolvimento fetal e, posteriormente, o desempenho fenotípico. Adicionalmente, a variação devido ao efeito permanente de ambiente pode ser inflacionada pela incorporação de efeitos não aditivos (dominância e epistasia) (KRUUK; HADFIELD, 2007).

Animais com apenas uma observação de determinada característica também possuem efeito ambiental sobre o seu registro; contudo, não é possível separar esse efeito dos efeitos aditivos e residuais. Neste caso, o efeito ambiental torna-se parte do termo residual no modelo de análise (SCHAEFFER, 2011).

2.1.3 Ganho genético e tendência genética

O sucesso de um programa de melhoramento genético está condicionado à definição correta dos objetivos e critérios de seleção bem como da predição acurada e adequada utilização dos valores genéticos na seleção genética (FERRAZ; ELER, 2010). Adicionalmente, o ganho genético está também condicionado à variabilidade da característica na população, à intensidade de seleção e ao intervalo de gerações (BOURDON, 2000). O ganho genético pode ser acompanhado por meio da tendência genética das características de importância econômica, a qual permite identificar a taxa (aumento ou redução) do progresso

anual alcançado nessas características por meio da seleção genética (CARRARA *et al.*, 2019; LAUREANO *et al.*, 2011). A tendência genética é obtida por meio do modelo de regressão linear simples denotado por:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.2)$$

em que \mathbf{X} é uma matriz de variáveis de dimensão $n \times p$, associada a um vector $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ de parâmetros de dimensão $n \times p$, e $\boldsymbol{\varepsilon}$ é um vector dos erros aleatórios de dimensão $n \times 1$ e são homoscedásticos, não correlacionados e distribuídos normalmente $N(0, \sigma^2 \mathbf{I})$ (HOFFMANN, 2016).

2.2 MODELO LINEAR MISTO

Modelos lineares mistos podem ser definidos como modelos que incluem efeitos fixos e aleatórios adicionalmente ao efeito do resíduo e a média, e é a principal ferramenta estatística para avaliação e seleção de animais domésticos, tanto para características contínuas quanto discretas (MARTINEZ; PERDOMO; ELZO, 2012; VAZQUEZ *et al.*, 2010).

Mrode (2014) menciona que os modelos mistos são usados quando os dados têm algum tipo de estrutura hierárquica ou de agrupamento, como delineamentos de medidas repetidas, que permitem ter coeficientes fixos, aleatórios e o erro. Nesse contexto, Faraway (2016) apresenta matricialmente o modelo linear misto univariado desenvolvido por Henderson em 1949:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \quad (2.3)$$

em que, \mathbf{y} é o vector de observações, de dimensão $n \times 1$, com n igual ao número de observações; \mathbf{X} é a matriz de incidência dos efeitos fixos (conhecida), de dimensão $n \times p$, com p igual ao número de parâmetros fixos; $\boldsymbol{\beta}$ é o vector $p \times 1$ de efeitos fixos desconhecidos; \mathbf{Z} é a matriz de incidência dos efeitos aleatórios (conhecida), de dimensões $n \times q$, com q parâmetros aleatórios; $\boldsymbol{\gamma}$ é o vector de efeitos aleatórios desconhecidos, de dimensão $q \times 1$; $\boldsymbol{\varepsilon}$ é o vector de erros aleatórios, de dimensão $n \times 1$. Assume-se que \mathbf{y} , os efeitos aleatórios e os erros têm distribuição normal, pelo qual a, $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$, $E(\mathbf{y}|\boldsymbol{\gamma}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma}$, $E(\boldsymbol{\gamma}) = E(\boldsymbol{\varepsilon}) = 0$, $Var(\boldsymbol{\gamma}) = \mathbf{G}$, $Var(\boldsymbol{\varepsilon}) = \mathbf{R}$, $cov(\boldsymbol{\gamma}, \boldsymbol{\varepsilon}') = 0$, de modo que, $Var(\mathbf{y}) = Var(\mathbf{X}\boldsymbol{\beta}) + Var(\mathbf{Z}\boldsymbol{\gamma}) + Var(\boldsymbol{\varepsilon}) = \mathbf{Z}Var(\boldsymbol{\gamma})\mathbf{Z}' + Var(\boldsymbol{\varepsilon}) = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R} = \mathbf{V}$.

Matricialmente tem-se,

$$\begin{bmatrix} \mathbf{y} \\ \gamma \\ \varepsilon \end{bmatrix} \sim N \left(\begin{bmatrix} \mathbf{X}\beta \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{ZGZ}' + \mathbf{R} & \mathbf{ZG} & \mathbf{R} \\ \mathbf{GZ}' & \mathbf{G} & \phi \\ \mathbf{R} & \phi & \mathbf{R} \end{bmatrix} \right)$$

sendo, ϕ a matriz nula. Considerando-se as variâncias dos resíduos como homogêneas tem-se $\mathbf{R} = \mathbf{I}\sigma_\varepsilon^2$, em que σ_ε^2 é a variância residual e \mathbf{I} a matriz identidade. Similarmente, \mathbf{G} é a matriz de covariância dos efeitos aleatórios. Assim,

$$\mathbf{y}|\gamma \sim N(\mathbf{X}\beta + \mathbf{Z}\gamma, \sigma^2\mathbf{I}) \quad (2.4)$$

assumindo-se normalidade, a função densidade de probabilidade de \mathbf{y} pode ser escrita da seguinte maneira,

$$f(\mathbf{y}) = \frac{1}{(2\pi)^{n/2}[\mathbf{ZGZ}' + \mathbf{R}]^{1/2}} \exp \left\{ -\frac{1}{2} [(\mathbf{y} - \mathbf{X}\beta)'(\mathbf{ZGZ}' + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{X}\beta)] \right\} \quad (2.5)$$

de forma que a função de probabilidade conjunta de \mathbf{y} e γ pode ser obtida como o produto da função de probabilidade condicional de \mathbf{y} dado γ , pela função de probabilidade γ ,

$$f(\mathbf{y}, \gamma) = f(\mathbf{y}|\gamma) \times f(\gamma) \quad (2.6)$$

ou seja,

$$f(\mathbf{y}, \gamma) = \frac{1}{(2\pi)^{n/2}|\mathbf{R}|^{1/2}} \exp \left\{ -\frac{1}{2} [(\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}\gamma)'\mathbf{R}^{-1}(\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}\gamma)] \right\} \times \frac{1}{(2\pi)^{j/2}|\mathbf{G}|^{1/2}} \exp \left\{ -\frac{1}{2} [\gamma'\mathbf{G}^{-1}\gamma] \right\} \quad (2.7)$$

em que, n e j são os postos das matrizes \mathbf{R} e \mathbf{G} , respectivamente.

As soluções dos efeitos fixos e aleatórios são obtidas por meio da maximização da função de densidade de probabilidade conjunta da [equação \(2.7\)](#) ([HENDERSON, 1973](#); [HENDERSON, 1975](#)). Para efeito de simplificação, utiliza-se o logaritmo neperiano da função

de verossimilhança. Assim, fazendo-se $l(\beta, \gamma) = \ln[f(\mathbf{y}, \gamma)]$, tem-se:

$$l(\beta, \gamma) = -\frac{1}{2}(n+j)\ln(2\pi) - \frac{1}{2}(\ln|\mathbf{R}| + \ln|\mathbf{G}|) - \frac{1}{2}(\mathbf{y}'\mathbf{R}^{-1}\mathbf{y} - 2\mathbf{y}'\mathbf{R}^{-1}\mathbf{X}\beta - 2\mathbf{y}'\mathbf{R}^{-1}\mathbf{Z}\gamma + 2\beta'\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\gamma + \beta'\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\beta + \gamma'\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\gamma + \gamma'\mathbf{G}^{-1}\gamma) \quad (2.8)$$

Derivando-se $l(\beta, \gamma)$ em relação a β e γ , obtêm-se as seguintes equações:

$$\begin{aligned} \frac{l(\beta, \gamma)}{\partial \beta} &= -\frac{1}{2} \left[\frac{\partial(\mathbf{y}'\mathbf{R}^{-1}\mathbf{X}\beta)}{\partial \beta} \right] - \frac{1}{2} \left[\frac{\partial(\beta'\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\gamma)}{\partial \beta} \right] - \frac{1}{2} \left[\frac{\partial(\beta'\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\beta)}{\partial \beta} \right] \\ &= -\frac{1}{2} (-\mathbf{X}'\mathbf{R}^{-1}\mathbf{y} + \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\gamma + \mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\beta) \end{aligned} \quad (2.9)$$

e

$$\begin{aligned} \frac{l(\beta, \gamma)}{\partial \gamma} &= \left[\frac{\partial(\mathbf{y}'\mathbf{R}^{-1}\mathbf{Z}\gamma)}{\partial \gamma} \right] - \left[\frac{\partial(\beta'\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\gamma)}{\partial \gamma} \right] - \frac{1}{2} \left[\frac{\partial(\gamma'\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\gamma)}{\partial \gamma} \right] - \frac{1}{2} \left[\frac{\partial(\gamma'\mathbf{G}^{-1}\gamma)}{\partial \gamma} \right] \\ &= (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} - \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}\beta - \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\gamma - \mathbf{G}^{-1}\gamma) \end{aligned} \quad (2.10)$$

tomando $\hat{\beta}$ e $\hat{\gamma}$ de tal maneira que as equações (2.9) e (2.10) se tornem nulas, chega-se às equações de modelos mistos (EMM):

$$\begin{aligned} \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} - \mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\hat{\beta} - \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\hat{\gamma} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} - \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}\hat{\beta} - \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\hat{\gamma} - \mathbf{G}^{-1}\hat{\gamma} \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\hat{\beta} + \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\hat{\gamma} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}\hat{\beta} + \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\hat{\gamma} + \mathbf{G}^{-1}\hat{\gamma} \end{bmatrix} &= \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \\ \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\hat{\gamma} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} &= \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \end{aligned} \quad (2.11)$$

a resolução da equação (2.11) fornece soluções para efeitos fixos $\hat{\beta}$ e efeitos aleatórios $\hat{\gamma}$:

$$\hat{\beta} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \quad (2.12)$$

$$\hat{\gamma} = \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}) \quad (2.13)$$

com,

$$\mathbf{V}^{-1} = \mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}.$$

As soluções das estimativas dos efeitos fixos (β) são denominadas de Melhor Estimador Linear não-viesado (BLUE) enquanto que a predição dos efeitos aleatórios é conhecida como a Melhor predição Linear não-viesada (BLUP) sob um modelo misto (HENDERSON, 1975).

2.3 MODELO MISTO NO MELHORAMENTO GENÉTICO ANIMAL

No modelo misto utilizado para a avaliação genética em programas de melhoramento animal, os efeitos genéticos são assumidos como efeitos aleatórios enquanto os efeitos de origem ambiental são incluídos como fixos. Este modelo apresenta como propriedades a maximização da correlação entre os valores genéticos verdadeiros e preditos; a estimação e predição não viesados em um método único; e a predição do mérito genético de indivíduos observados e não observados (GARCIA, 2004; HENDERSON *et al.*, 1959; WHITE; HODGE, 1989).

Os indivíduos avaliados geneticamente são relacionados por parentesco, de forma que os efeitos genéticos são correlacionados. Essas correlações genéticas aditivas podem ser estimadas a partir do registro genealógico dos animais (pedigree) e organizados em uma matriz chamada de matriz de relacionamentos aditivos.

Quando a característica de interesse é mensurada apenas uma vez na vida do animal (MUÑOZ; GONZÁLES, 2016), o modelo misto incluirá dois componentes aleatórios: γ , que corresponde ao vetor de efeitos genéticos aditivos aleatórios, proporcionando soluções diretas para os indivíduos sob avaliação em termos de valor genético aditivo ou mérito genético; e ε que é o vetor de efeitos aleatórios residuais (MARTINEZ; PERDOMO; ELZO, 2012; MRODE, 2014). A estrutura de covariâncias entre os efeitos aleatórios é dada por $\mathbf{G} = \mathbf{A}\sigma_{\gamma}^2$, sendo que \mathbf{A} é a matriz de relacionamentos construída a partir da informação genealógica dos indivíduos; enquanto a estrutura de covariâncias dos erros é assumida como $Var(\varepsilon) = \mathbf{R} = \mathbf{I}\sigma_{\varepsilon}^2$, sendo \mathbf{I} uma matriz de identidade quando se assume que os erros não são correlacionados (MARTINEZ; PERDOMO; ELZO, 2012; MRODE, 2014).

Segundo Vanderick (2017) ao fatorar-se \mathbf{R}^{-1} em ambos lados a matriz da

equação (2.11), tem-se:

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\lambda_1 \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\boldsymbol{\gamma}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix} \quad (2.14)$$

em que o valor de λ_1 é devida a herdabilidade em sentido restrito sendo calculada através da expressão:

$$\lambda_1 = \frac{1 - h^2}{h^2} = \frac{\sigma_\varepsilon^2}{\sigma_\gamma^2} \quad \text{e} \quad h^2 = \frac{\sigma_\gamma^2}{\sigma_P^2}.$$

Esse modelo em sua forma simplificada pode ser facilmente estendido para a avaliação de mais de uma característica, e para a inclusão do efeito aleatório genético materno, efeitos aleatórios genéticos não aditivos (tais como dominância e epistasia), bem como do efeito permanente de ambiente (MARTINEZ; PERDOMO; ELZO, 2012).

2.3.1 Modelo misto com medidas repetidas

O modelo com medidas repetidas (também denominado modelo de repetibilidade) é aplicado quando a característica de interesse é mensurada repetidamente em um mesmo indivíduo durante um determinado período ou trajetória. Em bovinocultura leiteira, características de produção e composição do leite geralmente apresentam registros repetidos dentro de uma mesma lactação ou em várias lactações de uma mesma vaca (BLASCO, 2017; MUÑOZ; GONZÁLES, 2016).

Segundo Mrode (2014) e Resende, Silva e Azevedo (2014) o modelo com medidas repetidas segue a mesma lógica do modelo misto geral e é dado por:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{W}\mathbf{pe} + \boldsymbol{\varepsilon} \quad (2.15)$$

sendo que \mathbf{W} é a matriz de incidência dos efeitos aleatórios de ambiente permanente e, \mathbf{pe} , o vetor de efeitos aleatórios permanentes de ambiente. O modelo apresenta a seguinte estrutura de variâncias:

$$\text{Var}(\boldsymbol{\gamma}) = \mathbf{A}\sigma_\gamma^2 = \mathbf{G}$$

$$\text{Var}(\mathbf{pe}) = \mathbf{I}\sigma_{pe}^2$$

$$\text{Var}(\boldsymbol{\varepsilon}) = \mathbf{I}\sigma_\varepsilon^2 = \mathbf{R}$$

$$\text{Var}(\mathbf{y}) = \mathbf{ZAZ}'\sigma_\gamma^2 + \mathbf{WI}\sigma_{pe}^2\mathbf{W}' + \mathbf{R}$$

As equações de modelos mistos são dadas por:

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} & \mathbf{X}'\mathbf{W} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\lambda_1 & \mathbf{Z}'\mathbf{W} \\ \mathbf{W}'\mathbf{X} & \mathbf{W}'\mathbf{Z} & \mathbf{W}'\mathbf{W} + \mathbf{I}\lambda_2 \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\gamma} \\ \widehat{\mathbf{pe}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \\ \mathbf{W}'\mathbf{y} \end{bmatrix}$$

em que,

$$\lambda_1 = \frac{\sigma_\varepsilon^2}{\sigma_\gamma^2} \quad ; \quad \lambda_2 = \frac{\sigma_\varepsilon^2}{\sigma_{pe}^2}$$

2.4 ABORDAGEM BAYESIANA NO CONTEXTO DO MELHORAMENTO ANIMAL

A abordagem Bayesiana tem se mostrado como uma metodologia válida para estudos em melhoramento genético, visto que incorpora a incerteza e o conhecimento prévio para a obtenção dos componentes de variância e os valores genéticos com intervalos exatos para as suas estimativas (GIANOLA; FERNANDO, 1986). Nesse contexto, a abordagem Bayesiana baseia-se na distribuição conjunta *a posteriori* de todos os parâmetros $\boldsymbol{\theta}$ (componentes de variância, efeitos ambientais e valores genéticos) tratados como variáveis aleatórias dado o vetor \mathbf{y} ou de observações fenotípicas, sendo esta resultante da combinação da distribuição *a priori* e da função de verossimilhança (dados amostrais), que permitem modelar e atualizar as estimativas *a posteriori* dos parâmetros desconhecidos $\boldsymbol{\theta}$ (GELMAN *et al.*, 2013; ROSSI, 2011).

Gelman *et al.* (2013) e Paulino, Turkman e Murteira (2003) apresentam o teorema de Bayes da seguinte maneira:

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\mathbf{y}, \boldsymbol{\theta})}{p(\mathbf{y})} = \frac{p(\mathbf{y}, \boldsymbol{\theta})}{\int_{\boldsymbol{\theta} \in \Theta} p(\mathbf{y}, \boldsymbol{\theta}) d\boldsymbol{\theta}} = \frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int_{\boldsymbol{\theta} \in \Theta} p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) d\boldsymbol{\theta}} \quad (2.16)$$

em que $\boldsymbol{\theta}$ é o vetor dos parâmetros; \mathbf{y} é o vetor das observações; $p(\boldsymbol{\theta}|\mathbf{y})$ é a distribuição conjunta *a posteriori* dos parâmetros dadas as observações e que contém toda a informação do parâmetro $\boldsymbol{\theta}$; $p(\mathbf{y}|\boldsymbol{\theta})$ é a distribuição condicional de \mathbf{y} dados os parâmetros, denominada de função de verossimilhança; $p(\boldsymbol{\theta})$ é a distribuição *a priori* ou também chamada de densidade marginal de $\boldsymbol{\theta}$. Por fim, a $\int_{\boldsymbol{\theta} \in \Theta} p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) d\boldsymbol{\theta}$ é a distribuição marginal ou preditiva de \mathbf{y} com respeito

a θ , sendo Θ a amplitude da distribuição de θ . Desse modo, o teorema de Bayes pode ser apresentado como:

$$p(\theta|\mathbf{y}) \propto p(\mathbf{y}|\theta) \times p(\theta)$$

em que \propto significa proporcional, ou seja, a distribuição conjunta *a posteriori* equivale ao produto entre a função de verossimilhança e a distribuição *a priori*.

A especificação da distribuição *a priori* é importante na inferência Bayesiana, pois influencia a distribuição *a posteriori* podendo ser especificada através das experiências e/ou crenças do indivíduo. A sua escolha não é trivial, dado que esta pode levar a diferentes distribuições *a posteriori* e, possivelmente, a diferentes conclusões (SMITH, 1991). A *a priori* é definida como informativa quando reflete algum conhecimento prévio sobre o parâmetro desconhecido; e como não informativa, difusa ou vaga quando não existem informações disponíveis sobre o parâmetro ou estas são superficiais. Neste último caso, é especificada uma distribuição *a priori* que não influencia a distribuição conjunta *a posteriori*, deixando que os dados mostrem todas as informações sobre θ (GAMERMAN; LOPES, 2006; JEFFREYS, 1961; STONE, 2013).

A distribuição conjunta *a posteriori* é o resultado final da análise Bayesiana e toda a informação sobre os parâmetros é obtida a partir dela. Na maioria dos problemas há mais de um parâmetro de interesse e é necessário diferenciar entre a distribuição conjunta *a posteriori* e a distribuição marginal *a posteriori* de cada parâmetro (NTZOUFRAS, 2011). O primeiro objetivo é encontrar a distribuição conjunta *a posteriori* de todos os parâmetros e, em seguida, obter a distribuição marginal de cada um dos parâmetros. Para isso, se integra a distribuição conjunta *a posteriori* em relação aos outros parâmetros $(\theta_1, \theta_2, \dots, \theta_p)$ da mesma (NTZOUFRAS, 2011). Portanto, se o interesse é em um determinado parâmetro de θ , é necessário obter a distribuição marginal $p(\theta_1|\mathbf{y})$, a qual é dada por:

$$p(\theta_1|\mathbf{y}) = \int_{\theta \neq \theta_1} p(\theta|\mathbf{y}) d\theta_{\theta \neq \theta_1}$$

o cálculo segue análogo para os demais parâmetros. Geralmente, a distribuição marginal *a posteriori* não é analiticamente tratável, sendo difícil de analisar porque envolve integrais de alta dimensão (NTZOUFRAS, 2011). Por isso, são utilizados métodos de Monte Carlo via cadeias de Markov (MCMC), para a obtenção das marginais condicionais completas via métodos numéricos de reamostragem e simulação. A obtenção destas permite calcular as

estimativas pontuais dos parâmetros e intervalos de credibilidade (GAMERMAN; LOPES, 2006; LINK; BARKER, 2010).

As estimativas dos parâmetros baseia-se nas distribuições marginais *a posteriori* de cada parâmetro de interesse e a estimativa pontual do parâmetro desconhecido θ_1 pode ser obtida a partir da esperança matemática de sua distribuição marginal *a posteriori*, $E(\theta_1|\mathbf{y})$. Também se pode obter o intervalo (a,b) de probabilidade do parâmetro (intervalo de credibilidade) ou seja, $\int_a^b p(\boldsymbol{\theta}|\mathbf{y})d\boldsymbol{\theta} = 1 - \alpha$ ($0 \leq \alpha \leq 1$), de forma que para α fixo quanto menor a amplitude do intervalo, mais concentrada está a distribuição do parâmetro (SMITH, 1991).

Em geral, a metodologia Bayesiana fornece uma solução para o problema de tamanho finito da amostra, porque existe uma distribuição exata *a posteriori* para cada conjunto de dados grande ou pequeno a partir do qual inferências podem ser extraídas. Quando um grande conjunto de dados é analisado, as informações *a priori* tendem a ser subjugadas pela função de verossimilhança no estabelecimento da distribuição *a posteriori* (RESENDE; SILVA; AZEVEDO, 2014). Nesse caso, as estimativas de parâmetros são próximas daquelas obtidas por métodos clássicos baseados em funções verossimilhança. No entanto, isso pode não ser verdade quando o tamanho da amostra é limitado, porque o procedimento de máxima verossimilhança apresenta apenas propriedades bem definidas quando o tamanho da amostra é suficiente (GIANOLA; FERNANDO, 1986; RESENDE; SILVA; AZEVEDO, 2014).

2.4.1 Abordagem Bayesiana na estimação de componentes de variância e efeitos genéticos

A utilização da inferência Bayesiana considera o modelo linear hierárquico em vários estágios que contempla a estimação simultânea dos parâmetros devido às atualizações realizadas no processo iterativo (GELMAN *et al.*, 2013). Neste contexto, todos os parâmetros do modelo, sejam fixos ou aleatórios (quando analisados de maneira clássica), são considerados como variáveis aleatórias (RESENDE; SILVA; AZEVEDO, 2014).

Nesta seção, a estimação dos parâmetros baseia-se no modelo uni-caraterística da equação (2.15) adotado do desenvolvimento de Gelman *et al.* (2013) e Misztal (2016) que

assumiram a distribuição condicional dos dados como uma distribuição normal:

$$\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{p}e, \sigma_\gamma^2, \sigma_{pe}^2, \sigma_\varepsilon^2 \sim N(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{W}\mathbf{p}e, \mathbf{I}\sigma_\varepsilon^2) \quad (2.17)$$

que a partir desta distribuição normal, tem-se a seguinte densidade:

$$p(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{p}e, \sigma_\gamma^2, \sigma_{pe}^2, \sigma_\varepsilon^2) \propto \left(\frac{1}{\sigma_\varepsilon^2}\right)^{\binom{n}{2}} \exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\right\} \quad (2.18)$$

sendo,

$$\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\gamma} - \mathbf{W}\mathbf{p}e).$$

e uma distribuição *a priori* não-informativa ou *flat* é estabelecida para o parâmetro $\boldsymbol{\beta}$ enquanto uma distribuição normal é adotada para os demais parâmetros, como:

$$p(\boldsymbol{\beta}) \propto \text{constante}; \quad -\infty < \boldsymbol{\beta} < \infty$$

$$\boldsymbol{\gamma}|\mathbf{A}, \sigma_\gamma^2 \sim N(0, \mathbf{A}\sigma_\gamma^2)$$

$$\mathbf{p}e|\mathbf{I}_{pe}, \sigma_{pe}^2 \sim N(0, \mathbf{I}_{pe}\sigma_{pe}^2)$$

apresentando a seguinte densidade,

$$p(\boldsymbol{\gamma}|\mathbf{A}\sigma_\gamma^2) \propto \left(\frac{1}{\sigma_\gamma^2}\right)^{\binom{n_\gamma}{2}} \exp\left\{-\frac{1}{2\sigma_\gamma^2}(\boldsymbol{\gamma}'\mathbf{A}^{-1}\boldsymbol{\gamma})\right\} \quad (2.19)$$

$$p(\mathbf{p}e|\sigma_{pe}^2) \propto \left(\frac{1}{\sigma_{pe}^2}\right)^{\binom{n_{pe}}{2}} \exp\left\{-\frac{1}{2\sigma_{pe}^2}(\mathbf{p}e'\mathbf{p}e)\right\} \quad (2.20)$$

em que \mathbf{A} é a matriz de relacionamentos.

Nesse contexto, os componentes de variância vem da distribuição qui-quadrado dada por uma razão de variâncias $\chi_v^2 = \frac{S^2v}{\sigma^2}$, em que v é um hiperparâmetro referente ao grau de confiança no componente de variância *a priori* S^2 (parâmetro de escala). Dessa expressão tem-se que $\sigma^2 = \frac{S^2v}{\chi_v^2}$ é a distribuição qui-quadrado invertida escalonada ou densidade *a priori* para o componente de variância σ^2 (RESENDE; SILVA; AZEVEDO, 2014). Especificamente, a distribuição qui-quadrado inversa escalonada pode ser usada como uma *priori* conjugada para o parâmetro de variância de uma distribuição normal (GELMAN *et al.*, 2013). Então tem-se as

seguintes densidades:

$$p(\sigma_\gamma^2 | v_\gamma, S_\gamma^2) \propto \left(\frac{1}{\sigma_\gamma^2} \right)^{\left(\frac{v_\gamma}{2} + 1 \right)} \exp \left\{ -\frac{v_\gamma S_\gamma^2}{2\sigma_\gamma^2} \right\} \quad (2.21)$$

$$p(\sigma_{pe}^2 | v_{pe}, S_{pe}^2) \propto \left(\frac{1}{\sigma_{pe}^2} \right)^{\left(\frac{v_{pe}}{2} + 1 \right)} \exp \left\{ -\frac{v_{pe} S_{pe}^2}{2\sigma_{pe}^2} \right\} \quad (2.22)$$

$$p(\sigma_\varepsilon^2 | v_\varepsilon, S_\varepsilon^2) \propto \left(\frac{1}{\sigma_\varepsilon^2} \right)^{\left(\frac{v_\varepsilon}{2} + 1 \right)} \exp \left\{ -\frac{v_\varepsilon S_\varepsilon^2}{2\sigma_\varepsilon^2} \right\} \quad (2.23)$$

em que v_γ , v_{pe} , v_ε representam os parâmetros da distribuição que correspondem ao graus de liberdade e S_γ^2 , S_{pe}^2 , S_ε^2 são os parâmetros de escala. Assumindo valores de $v_\gamma = v_{pe} = v_\varepsilon = -2$ e, $S_\gamma^2 = S_{pe}^2 = S_\varepsilon^2 = 0$, a distribuição qui quadrada inversa escalonada se reduz a uma distribuição constante, caracterizando *prioris* não informativas ou de conhecimento vago com intervalo das variâncias maiores do que zero (RESENDE; SILVA; AZEVEDO, 2014). Não obstante, ao se especificar valores aos graus de liberdade e aos valores iniciais das variâncias, essa distribuição torna-se informativa (SILVA, 2012).

De acordo com o teorema de Bayes, como descrito por Silva *et al.* (2013), obtém-se a distribuição conjunta *a posteriori* a partir de *prioris* uniformes:

$$p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{pe}, \sigma_\gamma^2, \sigma_{pe}^2, \sigma_\varepsilon^2 | \boldsymbol{y}) \propto p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{pe}, \sigma_\gamma^2, \sigma_{pe}^2, \sigma_\varepsilon^2) \times p(\boldsymbol{y} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{pe}, \sigma_\gamma^2, \sigma_{pe}^2, \sigma_\varepsilon^2)$$

$$p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{pe}, \sigma_\gamma^2, \sigma_{pe}^2, \sigma_\varepsilon^2 | \boldsymbol{y}) \propto p(\boldsymbol{y} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{pe}, \sigma_\gamma^2, \sigma_{pe}^2, \sigma_\varepsilon^2) p(\boldsymbol{\beta}) p(\boldsymbol{\gamma} | \boldsymbol{A} \sigma_\gamma^2) \times$$

$$p(\boldsymbol{pe} | \boldsymbol{I}_{pe}) p(\sigma_\gamma^2 | S_\gamma^2, v_\gamma) p(\sigma_{pe}^2 | S_{pe}^2, v_{pe}) p(\sigma_\varepsilon^2 | S_\varepsilon^2, v_\varepsilon)$$

ou seja,

$$\begin{aligned}
p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{pe}, \sigma_\gamma^2, \sigma_{pe}^2, \sigma_\varepsilon^2 | \mathbf{y}) &\propto \left(\frac{1}{\sigma_\varepsilon^2}\right)^{\left(\frac{n_\varepsilon + v_\varepsilon}{2} + 1\right)} \exp\left\{-\frac{1}{2\sigma_\varepsilon^2}[\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} + v_\varepsilon S_\varepsilon^2]\right\} \times \\
&\left(\frac{1}{\sigma_\gamma^2}\right)^{\left(\frac{n_\gamma + v_\gamma}{2} + 1\right)} \exp\left\{-\frac{1}{2\sigma_\gamma^2}[\boldsymbol{\gamma}'\mathbf{A}^{-1}\boldsymbol{\gamma} + v_\gamma S_\gamma^2]\right\} \times \\
&\left(\frac{1}{\sigma_{pe}^2}\right)^{\left(\frac{n_{pe} + v_{pe}}{2} + 1\right)} \exp\left\{-\frac{1}{2\sigma_{pe}^2}[\mathbf{pe}'\mathbf{pe} + v_{pe} S_{pe}^2]\right\}
\end{aligned} \tag{2.24}$$

As inferências sobre cada um dos parâmetros desconhecidos $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ e \mathbf{pe} são baseadas em suas respectivas densidades marginais, que são obtidas por uma integração sucessiva da densidade conjunta *a posteriori* da equação (2.24) em relação a cada um dos parâmetros de interesse. Para implementar o amostrador de Gibbs é necessário determinar todas as densidades condicionais completas dos parâmetros desconhecidos que são obtidos por meio da distribuição *a posteriori* conjunta tratando todos os outros parâmetros como conhecidos da equação (2.24). Portanto tem-se:

$$\boldsymbol{\beta} | \boldsymbol{\gamma}, \mathbf{pe}, \sigma_\gamma^2, \sigma_{pe}^2, \sigma_\varepsilon^2, \mathbf{y} \sim N(\hat{\boldsymbol{\beta}}, (\mathbf{X}'\mathbf{X})^{-1}\sigma_\varepsilon^2) \tag{2.25}$$

$$\boldsymbol{\gamma} | \boldsymbol{\beta}, \mathbf{pe}, \sigma_\gamma^2, \sigma_{pe}^2, \sigma_\varepsilon^2, \mathbf{y} \sim N(\hat{\boldsymbol{\gamma}}, (\mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\lambda_1)^{-1}\sigma_\varepsilon^2) \tag{2.26}$$

$$\mathbf{pe} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma_\gamma^2, \sigma_{pe}^2, \sigma_\varepsilon^2, \mathbf{y} \sim N(\widehat{\mathbf{pe}}, (\mathbf{W}'\mathbf{W} + \lambda_3)^{-1}\sigma_\varepsilon^2) \tag{2.27}$$

em que:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{y} - \mathbf{Z}\boldsymbol{\gamma} - \mathbf{W}\mathbf{pe})$$

$$\hat{\boldsymbol{\gamma}} = (\mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\lambda_1)^{-1}\mathbf{Z}'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{W}\mathbf{pe})$$

$$\widehat{\mathbf{pe}} = (\mathbf{W}'\mathbf{W} + \mathbf{I}\lambda_3)^{-1}\mathbf{W}'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\gamma})$$

obtendo-se as distribuições condicionais *a posteriori* completas dos componentes de variância são descritas a seguir:

$$\sigma_\gamma^2 | \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{pe}, \sigma_{pe}^2, \sigma_\varepsilon^2, \mathbf{y} \sim \tilde{v}_\gamma \tilde{S}_\gamma^2 \chi_{\tilde{v}_\gamma}^{-2} \tag{2.28}$$

$$\sigma_{pe}^2 | \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{p}\mathbf{e}, \sigma_\gamma^2, \sigma_\varepsilon^2, \mathbf{y} \sim \tilde{v}_{pe} \tilde{S}_{pe}^2 \chi_{\tilde{v}_{pe}}^{-2} \quad (2.29)$$

$$\sigma_\varepsilon^2 | \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{p}\mathbf{e}, \sigma_\gamma^2, \sigma_{pe}^2, \mathbf{y} \sim \tilde{v}_\varepsilon \tilde{S}_\varepsilon^2 \chi_{\tilde{v}_\varepsilon}^{-2} \quad (2.30)$$

em que,

$$\begin{aligned} \tilde{v}_\gamma &= n_\gamma + v_\gamma & \mathbf{e} & & \tilde{S}_\gamma^2 &= \boldsymbol{\gamma}' \mathbf{A}^{-1} \boldsymbol{\gamma} + v_\gamma S_\gamma^2 \\ \tilde{v}_{pe} &= n_{pe} + v_{pe} & \mathbf{e} & & \tilde{S}_{pe}^2 &= \mathbf{p}\mathbf{e}' \mathbf{p}\mathbf{e} + v_{pe} S_{pe}^2 \\ \tilde{v}_\varepsilon &= n_\varepsilon + v_\varepsilon & \mathbf{e} & & \tilde{S}_\varepsilon^2 &= \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} + v_\varepsilon S_\varepsilon^2 \end{aligned}$$

com \tilde{v} graus de liberdade (*degrees of belief*) e parâmetros de escala \tilde{S}^2 .

2.4.2 Modelo de limiar binário

Além do modelo linear existem modelos alternativos que podem ser utilizados para a obtenção dos componentes de variância, parâmetros e valores genéticos de certas características com distribuição não normal. Autores como [Ødegård et al. \(2010\)](#), [Perez et al. \(2017\)](#), [Silva et al. \(2019\)](#), [Sorensen e Gianola \(2002\)](#) e [Ventura et al. \(2015\)](#) sugerem utilizar modelos como Poisson, Binomial e Limiar (*threshold*) para as características que são de natureza discreta, binária ou categórica. Neste contexto, [Mrode \(2014\)](#), recomenda a utilização do modelo de limiar (*Threshold*) para modelar dados binários, o qual determina a probabilidade genética de um indivíduo assumir sucesso ou fracasso. O modelo de limiar é baseado na metodologia dos modelos mistos generalizados, que assumem uma variável contínua subjacente normalmente distribuída, introduzida como uma variável latente (*liability*) denotada por l ([HADFIELD; NAKAGAWA, 2010](#); [SORENSEN; GIANOLA, 2002](#)).

Conforme [Hadfield e Nakagawa \(2010\)](#), [Heringstad et al. \(2001\)](#), [Ødegård et al. \(2010\)](#) e [Sorensen e Gianola \(2002\)](#), o vetor \mathbf{y} das observações binárias é associado a uma variável latente (variável contínua não observada), denotada por l , de tal modo que:

$$\mathbf{y} = \begin{cases} 0, & \text{para } l_i \leq z \\ 1, & \text{para } l_i > z \end{cases} \quad (2.31)$$

assim, o valor do limiar z pode ser definido como zero em observações binárias e, l_i é o valor subjacente para a mencionada observação. Desta forma, utiliza-se o modelo de repetibilidade com variável latente dado por:

$$\mathbf{l} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{W}\mathbf{pe} + \boldsymbol{\varepsilon} \quad (2.32)$$

tal que \mathbf{l} é o vetor latente de todos os l_i ; \mathbf{X} , \mathbf{Z} e \mathbf{W} são as matrizes de incidência que relacionam os vetores $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ e \mathbf{pe} ao vetor \mathbf{l} , respectivamente; $\boldsymbol{\beta}$ é o vetor dos efeitos fixos; $\boldsymbol{\gamma}$ é o vetor de efeitos genéticos aditivos aleatórios de todos os indivíduos (animal), com $\boldsymbol{\gamma} \sim N(0, \mathbf{A}\sigma_\gamma^2)$, com \mathbf{A} sendo a matriz de relacionamentos genética aditiva, σ_γ^2 é a variância genética aditiva; $\boldsymbol{\varepsilon}$ é o vetor de resíduos aleatórios com $\boldsymbol{\varepsilon} \sim N(0, \mathbf{I}\sigma^2)$, sendo \mathbf{I} uma matriz de identidade com dimensão igual ao número de observações, e σ^2 a variação residual. No modelo de limiar (*threshold*) binário a variância residual não é estimável, sendo fixada em 1 (SORENSEN; GIANOLA, 2002).

Porém, os parâmetros do modelo seguem as mesmas distribuições e pressuposições utilizadas na [subseção \(2.4.1\)](#), com a diferença que \mathbf{y} é agora a variável latente \mathbf{l} e $\sigma_\varepsilon^2 = 1$:

$$\mathbf{l}|\boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{pe} \sim N(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{W}\mathbf{pe}, 1) \quad (2.33)$$

assim, a resolução é dada por meio da mesma estratégia vista anteriormente na construção do modelo linear hierárquico apresentada na [subseção \(2.4.1\)](#). As densidades condicionais completas dos efeitos $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ e \mathbf{pe} são:

$$\boldsymbol{\beta}|\boldsymbol{\gamma}, \mathbf{pe}, \sigma_\gamma^2, \sigma_{pe}^2, \mathbf{l} \sim N(\hat{\boldsymbol{\beta}}, (\mathbf{X}'\mathbf{X})^{-1}) \quad (2.34)$$

$$\boldsymbol{\gamma}|\boldsymbol{\beta}, \mathbf{pe}, \sigma_\gamma^2, \sigma_{pe}^2, \mathbf{l} \sim N(\hat{\boldsymbol{\gamma}}, (\mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\lambda_1)^{-1}) \quad (2.35)$$

$$\mathbf{pe}|\boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma_\gamma^2, \sigma_{pe}^2, \mathbf{l} \sim N(\widehat{\mathbf{pe}}, (\mathbf{W}'\mathbf{W} + \lambda_2)^{-1}) \quad (2.36)$$

em que:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{l} - \mathbf{Z}\boldsymbol{\gamma} - \mathbf{W}\mathbf{pe})$$

$$\hat{\boldsymbol{\gamma}} = (\mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\lambda_1)^{-1}\mathbf{Z}'(\mathbf{l} - \mathbf{X}\boldsymbol{\beta} - \mathbf{W}\mathbf{pe}), \quad \text{e} \quad \lambda_1 = (\sigma_\gamma^2)^{-1}$$

$$\widehat{\mathbf{p}}_e = (\mathbf{W}'\mathbf{W} + \mathbf{I}\lambda_3)^{-1}\mathbf{W}'(\mathbf{l} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\gamma}), \quad \text{e} \quad \lambda_2 = (\sigma_{pe}^2)^{-1}$$

a obtenção das distribuições condicionais *a posteriori* completas dos componentes de variância são descritas a seguir:

$$\sigma_\gamma^2 | \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{p}_e, \sigma_{pe}^2, \sigma_\varepsilon^2, \mathbf{l} \sim \tilde{v}_\gamma \tilde{S}_\gamma^2 \chi_{\tilde{v}_\gamma}^{-2} \quad (2.37)$$

$$\sigma_{pe}^2 | \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{p}_e, \sigma_\gamma^2, \sigma_\varepsilon^2, \mathbf{l} \sim \tilde{v}_{pe} \tilde{S}_{pe}^2 \chi_{\tilde{v}_{pe}}^{-2} \quad (2.38)$$

$$\sigma_\varepsilon^2 | \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{p}_e, \sigma_\gamma^2, \sigma_{pe}^2, \mathbf{l} \sim \tilde{v}_\varepsilon \tilde{S}_\varepsilon^2 \chi_{\tilde{v}_\varepsilon}^{-2} \quad (2.39)$$

em que,

$$\begin{aligned} \tilde{v}_\gamma &= n_\gamma + v_\gamma & \text{e} & \quad \tilde{S}_\gamma^2 = \boldsymbol{\gamma}'\mathbf{A}^{-1}\boldsymbol{\gamma} + v_\gamma S_\gamma^2 \\ \tilde{v}_{pe} &= n_{pe} + v_{pe} & \text{e} & \quad \tilde{S}_{pe}^2 = \mathbf{p}_e'\mathbf{p}_e + v_{pe} S_{pe}^2 \\ \tilde{v}_\varepsilon &= n_\varepsilon + v_\varepsilon & \text{e} & \quad \tilde{S}_\varepsilon^2 = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} + v_\varepsilon S_\varepsilon^2 \end{aligned}$$

e no modelo de limiar, a probabilidade do vetor \mathbf{y} dado a variável latente \mathbf{l} é dada por:

$$p(\mathbf{y}) = f(y_i | 1 - \Phi(l_i))$$

em que, f é a função de probabilidade das observações, y_i são as observações que pertencem à primeira ou à segunda categoria, respetivamente e, Φ é a função da distribuição acumulada de uma variável normal que indica que a função tem valor 1 se a expressão é avaliada como verdadeira e 0 caso contrário (HERINGSTAD *et al.*, 2001; ØDEGÅRD *et al.*, 2010; SORENSEN; GIANOLA, 2002).

2.4.3 Algoritmo MCMC: Amostrador de Gibbs

O método MCMC mais popular é o amostrador de Gibbs que foi introduzido por Geman e Geman em 1984, sendo um caso particular do algoritmo Metropolis-Hastings. Esses algoritmos são baseados na integração de Monte Carlo, que utiliza as cadeias de Markov para gerar amostras da distribuição de interesse sem a necessidade de conhecer a função exata da

densidade *a posteriori* (LINK; BARKER, 2010; GELMAN *et al.*, 2013; NTZOUFRAS, 2011). As cadeias devem convergir para uma distribuição estacionária *a posteriori* (equilibrada), por meio do método iterativo no qual em cada etapa é gerado um valor condicionado ao valor da etapa anterior (GELMAN *et al.*, 2013; NTZOUFRAS, 2011).

O Amostrador de Gibbs consiste no uso de distribuições condicionais completas dos parâmetros, sendo utilizado quando a distribuição condicional *a posteriori* está na forma de uma densidade de probabilidade com um padrão conhecido (GELMAN *et al.*, 2013). Este algoritmo consiste em amostrar sucessivamente as equações (2.25) a (2.30) (modelo linear) e as equações (2.34) a (2.39) (modelo de limiar). O algoritmo, explicado por Mrode (2014), é dado por:

1. Inicia-se o contador com interações em $t = 0$, definindo-se o número de iterações (k);
2. Especifica-se os valores iniciais para $\theta^{(0)}$. Neste passo pode-se atribuir *prioris* a β , γ , \mathbf{pe} , σ_γ^2 , σ_{pe}^2 , σ_ε^2 .
3. Avança-se t fazendo $t = t + 1$ e se obtém $\theta^{(t)}$ a partir de $\theta^{(t-1)}$ por geração sucessiva de valores como:

$$\begin{aligned}\theta_1^{(t)} &\sim p(\theta_1 | \theta_2^{(t-1)}, \theta_3^{(t-1)}, \theta_4^{(t-1)}, \dots, \theta_m^{(t-1)}) \\ \theta_2^{(t)} &\sim p(\theta_2 | \theta_1^{(t-1)}, \theta_3^{(t-1)}, \theta_4^{(t-1)}, \dots, \theta_m^{(t-1)}) \\ \theta_3^{(t)} &\sim p(\theta_3 | \theta_1^{(t-1)}, \theta_2^{(t-1)}, \theta_4^{(t-1)}, \dots, \theta_m^{(t-1)}) \\ &\vdots \\ \theta_m^{(t)} &\sim \pi(\theta_m | \theta_1^{(t)}, \theta_2^{(t)}, \theta_3^{(t)}, \dots, \theta_{m-1}^{(t)})\end{aligned}$$

nesta etapa calcula-se sequencialmente as amostras das distribuições condicionais:

$$\begin{aligned}\beta | \gamma, \mathbf{pe}, \sigma_\gamma^2, \sigma_{pe}^2, \sigma_\varepsilon^2 \\ \gamma | \beta, \mathbf{pe}, \sigma_\gamma^2, \sigma_{pe}^2, \sigma_\varepsilon^2 \\ \vdots \\ \sigma_\varepsilon^2 | \beta, \mathbf{pe}, \sigma_\gamma^2, \sigma_{pe}^2\end{aligned}$$

4. Se $t = k$, encerra-se;

5. Retorna-se ao passo 3.

2.4.4 Diagnóstico de convergência

Na prática, as distribuições estacionárias dos parâmetros de interesse são geradas por meio do processo estocástico *a posteriori*, sendo difícil estabelecer o número de iterações necessárias para que as cadeias geradas atinjam a convergência (próxima da distribuição de equilíbrio, ou seja, da distribuição marginal desejada) (NOGUEIRA; SÁFADI; FERREIRA, 2004). Nestas análises é comum descartar as primeiras observações amostrais para que o amostrador se distancie dos valores iniciais e atinja uma distribuição estacionária do parâmetro (ROSSI, 2011). Geralmente as primeiras 1.000 a 5.000 iterações são descartadas (10% a 50% do total) em um processo conhecido como período de queima, *burn-in* ou aquecimento (NOGUEIRA; SÁFADI; FERREIRA, 2004). Além disso, precisa-se de um espaçamento ou saltos (*thin*) de tamanho m para evitar a dependência no processo iterativo. O diagnóstico de convergência das cadeias detecta problemas fora do período de aquecimento (*burn-in*) podendo ser verificado graficamente e pelos critérios, dentre os quais destacam-se os diagnósticos de Geweke (1992), Heidelberger e Welch (1983) e Raftery e Lewis (1992) (NOGUEIRA; SÁFADI; FERREIRA, 2004).

O critério de Geweke (1992) realiza a comparação dos valores iniciais e finais das cadeias para verificar falhas de convergência, de modo que, ao testar a hipótese de nulidade (convergência da cadeia), o valor seja menor do que 1,96 (estatística Z da distribuição Normal padrão) e o nível de significância seja maior que 0,05 (5%) afirmando-se assim que não existe ausência de convergência (NOGUEIRA; SÁFADI; FERREIRA, 2004; ROSSI, 2011).

O critério de Heidelberger e Welch (1983) testa se a cadeia resultante é estacionária por meio da estatística Cramer-Von-Misses quando o *valor - p* é maior a 0,05 (ROSSI, 2011).

O critério de Raftery e Lewis (1992) sugere o valor do comprimento total da cadeia, o número de iterações para o período de aquecimento (*burn-in*) e o valor do espaçamento entre as observações (*thin*) (GAMERMAN; LOPES, 2006; PAULINO; TURKMAN; MURTEIRA, 2003). Ademais, outra informação importante é o fator de dependência (FD), que mede o aumento de iterações necessárias para alcançar a convergência. Quando essas iterações apresentam valores de FD muito maiores que 1 indicam um alto nível de dependência e

correlações altas dentro da cadeia provocando falhas de convergência. Ainda, FD maiores do que 5 indicam problemas na convergência com presença de dependência entre as iterações (RAFTERY; LEWIS, 1992b).

2.4.5 Critério de Informação do Desvio (DIC)

O critério DIC é uma medida muito utilizada na comparação e seleção de modelos com base na distribuição *a posteriori* obtida pela simulação de Monte Carlo via cadeias de Markov (MCMC) (SPIEGELHALTER *et al.*, 2002). Esse critério é facilmente calculado e é considerado como uma generalização hierárquica do critério de Informação de Akaike (AIC) e do critério de Informação Bayesiano de Schwarz (BIC) (NTZOUFRAS, 2011). O DIC foi desenvolvido por Spiegelhalter *et al.* (2002) e é apresentado da seguinte forma:

$$DIC = D(\bar{\theta}) + 2pD \quad (2.40)$$

em que $D(\bar{\theta}) = -2\ln L(\bar{\theta}) + C$ e, o $D(\bar{\theta})$ é o desvio avaliado na média *a posteriori* e θ é o vetor dos parâmetros desconhecidos do modelo; $L(\theta)$ a função de verossimilhança; C é uma constante, a qual não precisa ser conhecida na comparação dos modelos visto que é cancelada; e pD é o número efetivo de parâmetros, o qual é dado por $pD = \bar{D} - D(\bar{\theta})$, sendo $\bar{D} = E[D(\bar{\theta})]$ o desvio esperado na média *a posteriori* e que mede a qualidade do ajuste do modelo aos dados. Modelos com menores valores de DIC são considerados mais adequados, pois apresentam ajuste ponderado pelo grau de complexidade. Os mesmos autores também sugerem utilizar o módulo da diferença entre os valores de DIC de dois modelos A e B analisados, tal como é citado por Rossi (2011), dado da seguinte forma:

$$\Delta = |DIC_A - DIC_B| \quad (2.41)$$

com as seguintes decisões:

- se $\Delta < 5$ é não substancial;
- se $5 \leq \Delta \leq 10$, é significativo ou a diferença dos modelos são substanciais;
- se $\Delta > 10$, conclui-se que o modelo com maior valor DIC é excluído e é significativo.

3 MATERIAIS E MÉTODOS

Nesta seção detalha-se a obtenção dos dados e explicam-se os procedimentos utilizados nas análises estatísticas.

3.1 ANIMAIS E DADOS

Foram utilizadas 127.434 mensurações referentes às características de produção, composição do leite e saúde de 7.664 vacas da raça Holandesa, coletadas mensalmente entre maio de 2012 e dezembro de 2016 em quatro fazendas, estando três fazendas localizadas no estado de São Paulo e uma no estado do Paraná. Esses dados foram obtidos com a colaboração da Clínica do Leite (Departamento de Zootecnia, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, ESALQ/USP).

Foram disponibilizadas informações individuais acerca da genealogia, data de parto, dias em lactação, ordem de lactação e idade ao primeiro parto (IPP) adicionalmente aos dados de desempenho individual, como produção de leite (PROD, em kg por dia), contagem de células somáticas (CCS, em células por ml de leite), e teor de caseína no leite (CAS, em g por 100g de leite).

Anteriormente às análises, foram eliminados registros:

- a) sem informação de data de nascimento, data de parto, data de coleta, data de análise, ordem de lactação e idade à mensuração;
- b) com medidas de PROD e CAS fora da amplitude aceitável (mais ou menos três desvios-padrão) ([PETRINI, 2016](#));
- c) associados a dias de lactação inferior a cinco ou superior a 365;
- d) associados à idade maior que 4.213 dias (aproximadamente onze anos).

Adicionalmente, foram mantidos apenas animais com mais de 3 mensurações na lactação e também com número de lactações igual ou menor a sete com a finalidade de obter uma estimativa mais precisa do efeito ambiental.

Os grupos contemporâneos (GC) foram formados de acordo com o ano e mês de parto, ano e mês da coleta, ano e mês de análise, estação do parto (estação seca - abril a setembro; estação das águas - outubro a março) e rebanho (fazenda), conforme descrito em (PETRINI *et al.*, 2016). Grupos de contemporâneos associados a menos de 5 indivíduos foram descartados.

Ao final restaram 64.657 mensurações de 5.439 vacas, pertencentes a 471 GC e com dias de lactação entre 5 e 365 (média igual a 170,6) e o número de vacas de primeira à sétima lactação foi igual a 3.941, 2.860, 1.671, 767, 349, 128 e 42, respectivamente.

Além disso, foram obtidos dois fenótipos binários:

- a) Incidência de mastite subclínica (MSC): assumida como igual a zero (0) quando CCS \leq 200.000 células por ml de leite e igual a um (1) quando CCS $>$ 200.000 células por mililitro de leite. De acordo com Schukken *et al.* (2003) a incidência de mastite subclínica pode ser identificada pelo aumento do CCS, ou seja, CCS superior a 200.000 células por mililitro de leite é geralmente considerado evidência de MSC;
- b) Precocidade (PREC): assumida como igual a zero (0) quando a idade da vaca ao primeiro parto foi superior a 730 dias (24 meses) e igual a um (1) quando a idade da vaca ao primeiro parto foi menor ou igual a 730 dias (24 meses). A idade ideal para a primeira concepção de vacas da raça Holandesa é em torno de 24 meses (NILFOROOSHAN; EDRISS, 2004).

Similarmente ao realizado para as demais características, para a obtenção de PREC foram somente consideradas informações de IPP dentro da amplitude aceitável (mais ou menos três-desvios padrão). Os GC para PREC consideraram as informações de rebanho, ano e estação de parto (igualmente ao adotado em Petrini (2016) para idade ao primeiro parto), sendo formados 48 GC.

3.2 ANÁLISE GENÉTICA

As estimativas dos componentes de variância e parâmetros genéticos para as características contínuas e binárias foram obtidos a partir de diferentes modelos em relação à sua estrutura genética em relação à sua respectiva distribuição.

3.2.1 Análise genética para características contínuas

Para as características de produção de leite (PROD) e teor de caseína (CAS) foi ajustado o seguinte modelo de repetibilidade:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{W}\mathbf{pe} + \boldsymbol{\varepsilon} \quad (3.1)$$

em que \mathbf{y} é o vetor dos fenótipos das características em estudo (PROD e CAS); $\boldsymbol{\beta}$ é o vetor que incluiu os efeitos de ordem de lactação, grupos contemporâneos, e como covariável os efeitos linear, quadrático e cúbico de dias em lactação; $\boldsymbol{\gamma}$ é o vetor dos efeitos genéticos aditivos; \mathbf{pe} é o vetor dos efeitos permanentes de ambiente (não correlacionados); e $\boldsymbol{\varepsilon}$ é o vetor dos efeitos residuais; \mathbf{X} , \mathbf{Z} e \mathbf{W} são matrizes de incidência que relacionam \mathbf{y} a $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, e \mathbf{pe} , respectivamente.

3.2.2 Análise genética para características binárias

Para as características binárias de incidência de mastite (MSC) e precocidade (PREC), foi adotado o modelo de limiar da [equação \(3.2\)](#) incluindo os mesmos efeitos apresentados na [equação \(3.1\)](#) (exceto o efeito permanente de ambiente para PREC)

$$\mathbf{l} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{W}\mathbf{pe} + \boldsymbol{\varepsilon} \quad (3.2)$$

no entanto, neste modelo, o vetor \mathbf{l} é referente a uma variável contínua não observada (*liability*) do vetor dos fenótipos das características em estudo (MSC e PREC). Para a característica MSC, o vetor $\boldsymbol{\beta}$ incluiu os efeitos de ordem de lactação, grupos de contemporâneos, e como covariável o efeito linear, quadrático e cúbico de dias em lactação. Para a PREC, o vetor $\boldsymbol{\beta}$ incluiu somente o efeito de grupos de contemporâneos.

3.2.3 Estimação dos componentes de variância e valores genéticos

Os componentes de variâncias e os valores genéticos foram obtidos por meio de abordagem Bayesiana sob modelo linear [equação \(3.1\)](#) e de limiar [equação \(3.2\)](#), utilizando distribuições *a priori* não informativas e informativas para as características contínuas (PROD e CAS) e binárias (MSC e PREC).

3.2.3.1 Especificação da *priori* não informativa

Considerou-se o $\beta \propto$ constante, com $-\infty < \beta < \infty$, indicando vago conhecimento sobre esse vetor; $\gamma | \sigma_\gamma^2 \sim N(0, \mathbf{A}\sigma_\gamma^2)$; $pe | \sigma_{pe}^2 \sim N(0, \mathbf{I}\sigma_{pe}^2)$ (exceto para PREC) e $\varepsilon | \sigma_\varepsilon^2 \sim N(0, \mathbf{I}\sigma_\varepsilon^2)$; com σ_γ^2 sendo a variância genética aditiva, considerada desconhecida, σ_{pe}^2 a variância do efeito permanente de ambiente desconhecida, e σ_ε^2 a variância residual, também desconhecida. Adicionalmente, para os componentes de variância foi adotada como *priori* a distribuição qui-quadrado invertida escalonada: $\sigma_\gamma^2 \sim \chi_{v_\gamma}^{-2} v_\gamma S_\gamma^2$; $\sigma_{pe}^2 \sim \chi_{v_{pe}}^{-2} v_{pe} S_{pe}^2$ e $\sigma_\varepsilon^2 \sim \chi_{v_\varepsilon}^{-2} v_\varepsilon S_\varepsilon^2$, tais que v_γ , v_{pe} e v_ε correspondem aos graus de liberdade ("*degrees of belief*") e S_γ^2 , S_{pe}^2 e S_ε^2 são os parâmetros de escala ou valores *a priori* da distribuição qui-quadrado invertida para a variância aditiva, permanente de ambiente e residual, respectivamente ([CARRARA et al., 2019](#)).

Considerou-se como zero os valores iniciais da média para os efeitos γ e pe , visto que eles vêm aleatoriamente de uma distribuição normal. Para os componentes de variância das características contínuas e binárias, valores iniciais arbitrários de $\sigma_\gamma^2 = 1$, $\sigma_{pe}^2 = 1$ e $\sigma_\varepsilon^2 = 1$ (fixada para as binárias) foram adotados para inicializar o amostrador de Gibbs, de maneira que $\lambda_1 = \sigma_\varepsilon^2 / \sigma_\gamma^2 = 1$ e $\lambda_2 = \sigma_\varepsilon^2 / \sigma_{pe}^2 = 1$ (exceto para PREC) ([MISZTAL, 2016](#); [NOGUEIRA et al., 2003](#); [RESENDE; SILVA; AZEVEDO, 2014](#)).

3.2.3.2 Especificação da *priori* informativa

Da mesma forma vista na [seção \(3.2.3.1\)](#) considerou-se o $\beta \propto$ constante e os efeitos com uma distribuição normal $\gamma | \sigma_\gamma^2 \sim N(0, \mathbf{A}\sigma_\gamma^2)$; $pe | \sigma_{pe}^2 \sim N(0, \mathbf{I}\sigma_{pe}^2)$ (exceto para PREC) e $\varepsilon | \sigma_\varepsilon^2 \sim N(0, \mathbf{I}\sigma_\varepsilon^2)$. Como os componentes de variância assumem uma distribuição qui-quadrado invertida seus graus de liberdade e parâmetros de escala foram obtidos com base no trabalho de [Carneiro Júnior et al. \(2005\)](#), que tinha como objetivo verificar a influência da informação *a priori* na avaliação genética animal com valores de graus de liberdade para a

variância genética aditiva (v_γ) e permanente de ambiente v_{pe} iguais a 60 (exceto PREC), e para a variância residual v_ε igual a 100; e S_γ^2 , S_{pe}^2 e S_ε^2 iguais aos valores verdadeiros dos componentes de variância a serem estimados para as características em estudo (PROD, CAS, MSC e PREC). Assim, quanto maior o valor dos graus de liberdade menor é a variância da distribuição *a priori* do valor indicado por S_γ^2 , S_{pe}^2 e S_ε^2 (CARNEIRO JÚNIOR *et al.*, 2005).

Logo, os componentes de variância (σ_γ^2 , σ_{pe}^2 e σ_ε^2) para as características PROD e CAS foram calculados a partir dos resultados apresentados em Tiezzi *et al.* (2013). Neste, foram reportadas as estimativas de herdabilidade (h^2), repetibilidade (Rep), proporção da variância fenotípica devido ao efeito de rebanho (h_{Reb}^2), e desvio-padrão fenotípico (Dp) para PROD e CAS em uma população de 16.089 vacas da raça Holandesa, pertencentes a 345 rebanhos e com 63.470 registros de desempenho. No caso de MSC, foram utilizados como valores iniciais para estas variâncias as estimativas obtidas em Oliveira *et al.* (2015) que utilizaram uma população de 5.084 vacas da raça Holandesa com registros de 11.738 lactações tomados no período de 1995 a 2010. Este trabalho foi desenvolvido com a finalidade de estimar os parâmetros genéticos para mastite clínica, considerando um modelo de limiar, em que 1 e 0 representavam a ocorrência ou não da doença, respectivamente. Finalmente, para PREC foi utilizado o trabalho de Ayalew, Aliy e Negussie (2017), no qual foi estimada a herdabilidade para idade ao primeiro parto em 1.125 vacas da raça Holandesa.

Na Tabela 1, encontram-se resumidamente os valores iniciais adotados para as variâncias, utilizados para inicializar o amostrador de Gibbs.

Tabela 1 – Resumo das *prioris* informativas utilizadas para as variâncias iniciais aditivas (σ_γ), permanente de ambiente (σ_{pe}) e a variância residual (σ_ε)

Característica	h^2	Rep	h_{Reb}^2	Dp	σ_P^2	σ_{Reb}^2	σ_γ^2	σ_{pe}^2	σ_ε^2
PROD ¹	0,13	0,46	0,21	8,79	77,26 *	16,15 *	10,04 *	25,57 *	25,50 *
CAS ¹	0,11	0,53	0,14	0,30	0,09 *	0,01 *	0,01 *	0,04 *	0,03 *
MSC ²	0,11	-	-	-	-	-	0,14	0,122	1
PREC ³	0,47	-	-	-	-	-	0,89	-	1

Fonte: Da autora.

Nota: * Variâncias obtidas com as seguintes expressões: $\sigma_P^2 = (Dp)^2$; $\sigma_\gamma^2 = h^2 \times \sigma_P^2$; $\sigma_{Reb}^2 = h_{Reb}^2 \times \sigma_P^2$; $\sigma_{pe}^2 = Rep \times \sigma_P^2 - \sigma_\gamma^2$ e $\sigma_\varepsilon^2 = \sigma_P^2 - \sigma_{Reb}^2 - \sigma_{pe}^2 - \sigma_\gamma^2$.

¹ Informações *a priori* obtidas do trabalho de Tiezzi *et al.* (2013).

² Informações *a priori* obtidas do trabalho de Oliveira *et al.* (2015b).

³ Informações *a priori* obtidas do trabalho de Ayalew, Aliy e Negussie (2017).

Logo, as variâncias (σ_γ^2 , σ_{pe}^2 , σ_ε^2) da Tabela 1, foram utilizados para dar início ao

amostrador de Gibbs dos modelos lineares (PROD e CAS) e de limiar binários (MSC, PREC). Assim, $\lambda_1 = \frac{25,50}{10,0} = 2,540$ e $\lambda_2 = \frac{25,50}{25,57} = 0,997$ para PROD; $\lambda_1 = \frac{0,03}{0,01} = 3,00$ e $\lambda_2 = \frac{0,03}{0,04} = 0,75$ para CAS; $\lambda_1 = \frac{1}{0,14} = 7,143$ e $\lambda_2 = \frac{1}{0,12} = 8,333$ para MSC e $\lambda_1 = \frac{1}{0,89} = 1,124$ para PREC.

As análises Bayesianas com *priori* não informativa e informativa, baseadas no modelo linear e de limiar (*threshold*), foram realizadas por meio do software GIBBS3F90 (modelo linear) e THRGIBBS3F90 (modelo de limiar), utilizando o algoritmo de Gibbs (MISZTAL *et al.*, 2019). Foram geradas cadeias de 5.000.000 iterações, com queima inicial (*burn-in*) de 1.000.000 para minimizar os efeitos dos valores iniciais, e saltos (*thin*) de 800 iterações para garantir a independência da amostra.

As estimativas das variâncias fenotípicas e herdabilidades foram calculadas a partir dos componentes de variância de cada distribuição *a posteriori* da iteração. A variância fenotípica (σ_P^2) foi calculada como a soma das variâncias genética aditiva (σ_γ^2), permanente de ambiente (σ_{pe}^2) e residual (σ_ε^2), enquanto a herdabilidade (h^2) foi calculada como a razão entre a variância genética aditiva e a variância fenotípica para cada característica estudada. Da mesma forma, a repetibilidade (*Rep*) foi calculada como a razão entre a soma da variância genética aditiva e a variância do efeito permanente de ambiente, e a variância fenotípica, isto é, $Rep = \frac{(\sigma_\gamma^2 + \sigma_{pe}^2)}{(\sigma_P^2)}$. Por sua vez, a proporção da variância fenotípica associada ao efeito permanente de ambiente (c^2) foi calculada como a razão entre a variância do efeito permanente de ambiente (σ_{pe}^2) e a variância fenotípica (σ_P^2) com a finalidade de isolar o efeito ambiental.

Intervalos de máxima densidade *a posteriori* (HPD) foram construídos para todos os componentes de variância e parâmetros genéticos estimados com um nível de credibilidade de 95%. A convergência das cadeias foram verificadas por meio do pacote *Bayesian Output Analysis* (BOA) de Smith *et al.* (2007) no software R versão 3.6.0 R Core Team (2019), graficamente e utilizando os diagnósticos de Geweke (1992), Heidelberger e Welch (1983) e Raftery e Lewis (1992). O erro de Montecarlo (EMC) também foi estimado para cada parâmetro de interesse, a fim de indicar quanto erro há na estimativa devido à simulação MCMC (NTZOUFRAS, 2011). Valores do EMC baixos indicam uma boa qualidade das estimativas dos parâmetros (BEIJO; VIVANCO; MUNIZ, 2009).

A seleção do melhor modelo em relação ao uso de *priori* não informativa e informativa para cada característica foi baseada no critério DIC, calculado pela equação (2.41), descrita na seção (2.4.5).

3.2.4 Avaliação da tendência genética

As estimativas das tendências genéticas para as características avaliadas foram obtidas pela regressão linear simples dos valores genéticos médio preditos sobre o ano de nascimento dos animais:

$$\hat{\gamma} = \mathbf{X}\hat{\beta} + \varepsilon \quad (3.3)$$

em que $\hat{\gamma}$ representa valor genético predito em escala normal (valores genéticos preditos para PROD e CAS) e latente (valores genéticos preditos para MSC e PREC), β é o vetor que contém o intercepto β_0 e o coeficiente angular β_1 associado ao ano de nascimento do vetor \mathbf{X} dos animais, e ε é o vetor de resíduos. O coeficiente angular estimado corresponde ao ganho genético anual na referida característica. Desta forma, consegue-se quantificar o ganho genético ocorrido em uma determinada característica (LACERDA *et al.*, 2014).

Para γ assumiu-se como *priori* a distribuição distribuição normal $N(\mu, \tau)$ com média $\mu = \beta_0 + \beta_1 x$ e precisão $\tau = \frac{1}{\sigma^2}$, definido como o inverso da variância com uma distribuição gama e hiperparâmetros ($\alpha_0 = 0,01; \alpha_1 = 0,01$). Os parâmetros $\beta_0 \sim N(0, 10^{-6})$ e $\beta_1 \sim N(0, 10^{-6})$ seguem uma distribuição normal com *prioris* não informativas para os hiperparâmetros. Para todas as características estudadas, foram estabelecidos como valores iniciais para os parâmetros $\beta_0 = 1$, $\beta_1 = 1$ e $\tau = 1$ (ROSSI, 2011).

Foi gerada uma cadeia com 5.000.000 de iterações, com descarte de 1.000.000 iterações e salto de 800 iterações, obtendo-se uma cadeia final de tamanho igual a 5.000 amostras. O ajuste do modelo e obtenção das distribuições *a posteriori* dos parâmetros β_0 , β_1 e τ foram realizados por meio do software OpenBUGS de Thomas e O'Hara (2004) pela interface do pacote R2OpenBUGS (LIGGES *et al.*, 2017). Verificou-se a convergência das cadeias por meio do pacote *Output Analysis and Diagnostics for MCMC* (CODA) de Plummer *et al.* (2006), implementado no software estatístico R 3.6.0 (R CORE TEAM, 2019). A convergência das cadeias foi verificada graficamente e pelos diagnósticos de Geweke (1992), Heidelberger e Welch (1983) e Raftery e Lewis (1992).

Para a verificação da significância dos parâmetros estimados de β_0 e β_1 utilizou-se o $HPD_{95\%}$ e estes foram testados pelas seguintes hipóteses:

$$\begin{cases} H_0 : \beta_0 = 0 \\ H_1 : \beta_1 \neq 0 \end{cases}$$

sendo rejeitada H_0 se zero não pertence ao $HPD_{95\%}$, ou seja, o parâmetro é significativo e, caso contrário, se $0 \in HPD_{95\%}$ (RIBEIRO *et al.*, 2018; ROBERT, 2007). A qualidade do ajuste foi verificada pelo coeficiente de determinação (R^2), definido como a proporção dos dados observados que podem ser explicados pelo modelo, sendo expresso pela seguinte equação: $R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$ (HOFFMANN, 2016).

4 RESULTADOS E DISCUSSÃO

Nesta seção apresentam-se as estatísticas descritivas, as estimativas dos componentes de variância, parâmetros genéticos, e a predição do ganho genético por meio da regressão dos valores genéticos sobre o ano de nascimento.

4.1 ANÁLISE EXPLORATÓRIA DOS DADOS

As estatísticas descritivas para as características contínuas PROD, CAS, CCS e IPP, e binárias MSC e PREC (obtidas a partir de CCS e IPP, respectivamente) estão inseridas na [Tabela 2](#).

Tabela 2 – Estatísticas descritivas obtidos para as características contínuas e binárias com seu respectivo número de mensurações (N), número de vacas (NV), média, desvio-padrão (DP), coeficiente de variação (CV %), valores mínimos (MIN) e máximos (MAX).

Caraterística*	N	NV	Média	DP	CV %	MIN	MAX
PROD	43.501	5.218	33,66	10,672	31,7	5,00	66,20
CAS	59.453	5.327	2,37	0,286	12,0	1,48	3,28
CCS	64.372	5.439	428,48	1081,941	252,5	1,00	15.974,00
IPP	4.501	4.501	816,90	132,102	16,2	476,00	1.945,00
MSC	64.372	5.439	0,30	0,456	154,3	0,00	1,00
PREC	4.501	4.501	0,25	0,434	172,2	0,00	1,00

Fonte: Da autora.

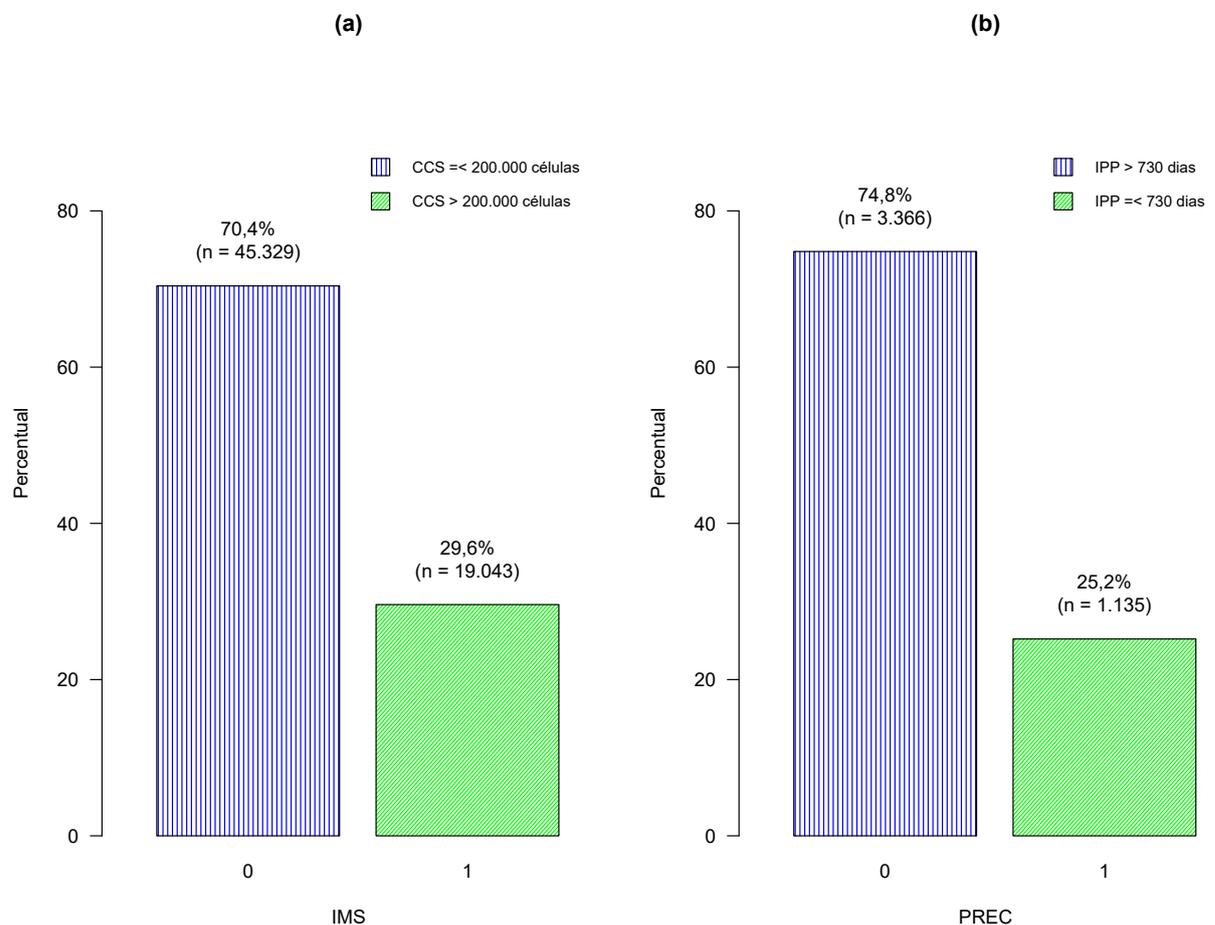
Nota: * PROD (produção de leite medida em kg), CAS (teor de caseína no leite medida em g/100g de leite), CCS (contagem de células somáticas no leite medida em 1.000 células/ml de leite) e IPP (idade ao primeiro parto em dias).

Observa-se que a produtividade de leite no dia do controle (PROD) apresenta um valor médio de 33,6 kg/vaca/dia com coeficiente de variação (CV) de 31,7%. Esses valores foram próximos aos relatados por [Costa et al. \(2018\)](#), [Tiezzi et al. \(2013\)](#) e [Toffanin et al. \(2015\)](#) em estudos anteriores para vacas da raça Holandesa. Eles obtiveram médias de 27,58, 32,00 e 30,19 kg/dia, respectivamente. O valor médio de teor de caseína foi de 2,37 g/100 g de leite e foram também próximo aos relatados nos estudos de [Gaya et al. \(2013\)](#) e [Tiezzi et al. \(2013\)](#) provenientes de 2.382 e 25.590 vacas da raça da raça Holandesa, respectivamente. Esses autores

em média obtiveram caseína no leite de 2,31% e 2,64%, respectivamente.

Para uma melhor visualização apresenta-se a frequência percentual das características binárias (MSC e PREC) na [Figura 1](#).

Figura 1 – Frequência percentual das características binárias: incidência de mastite subclínica (MSC) (a) e Precocidade (PREC) (b). Incidência de mastite subclínica igual a um (1) para contagem de células somáticas (CCS) maior que 200.000 células/ml de leite, e igual a zero (0) caso contrário; PREC igual a um (1) para idade ao primeiro parto menor que 730 dias, e igual a zero (0) caso contrário.



Fonte: Da autora.

Das 64.372 mensurações de contagem de células somáticas de 5.439 vacas, 19.043 mensurações ou 29,6% indicam incidência de mastite subclínica nos animais avaliados.

Em relação às vacas com registro de idade ao primeiro parto, verificou-se que 1.135 vacas, ou seja, 25,2% pariram antes de 24 meses de idade, sendo consideradas precoces.

4.2 COMPONENTES DE VARIÂNCIA E PARÂMETROS GENÉTICOS

A convergência foi atingida para todos os parâmetros, indicando que o período de queima (*burn-in*), salto (*thin*) e número de iterações utilizados foram suficientes para a estimação dos parâmetros de interesse. O módulo do valor de Geweke $|ZG|$ variou entre 0,235 e 1,833; portanto, menor do que 1,96, não apresentando indícios de ausência de convergência. Os valores próximos a 1 obtidos para o fator de dependência do critério de Raftery e Lewis confirmam a independência entre as iterações enquanto que pelo critério de Heidelberger e Welch foi constatado que a série é estacionária ($p > 0,05$) (Tabelas 3, 4,5 e 6). As densidades *a posteriori* dos parâmetros apresentaram maior proximidade da normalidade (Apêndice A e Apêndice B).

Na Tabela 3 são apresentadas as estimativas dos componentes de variância, coeficiente de herdabilidade (h^2), repetibilidade (*Rep*) e proporção da variância fenotípica devido ao efeito permanente de ambiente (c^2) para a característica de produção de leite (PROD).

Tabela 3 – Estimativas (média *a posteriori*) dos componentes de variância e parâmetros (componentes de variância e parâmetros genéticos) obtidos com *priori* não informativa (NI) e informativa (INF) com seus respectivos intervalos ($HPD_{95\%}$), critérios de convergência e erros de Monte Carlo (EMC) para a característica contínua de produção de leite (PROD), período 2012 a 2016.

Prioris	Parâmetros ¹	Estimativas	$HPD_{95\%}^2$		EMC	Critérios de convergência ³		
			LI	LS		$ ZG $	RL	HW
NI	$\hat{\sigma}_{\gamma}^2$	11,089	8,175	14,020	0,023021	0,635	1,015	0,062
	$\hat{\sigma}_{pe}^2$	18,599	16,120	20,900	0,019687	0,235	1,022	0,172
	$\hat{\sigma}_{\varepsilon}^2$	48,549	47,910	49,300	0,004279	1,101	1,043	0,064
	$\hat{\sigma}_P^2$	78,238	76,447	79,903	0,013346	0,312	0,989	0,132
	h^2	0,142	0,105	0,177	0,000284	0,641	1,015	0,071
	c^2	0,238	0,205	0,268	0,000255	0,284	1,015	0,148
	Rep	0,379	0,366	0,393	0,000092	1,164	1,015	0,199
INF	$\hat{\sigma}_{\gamma}^2$	10,222	8,041	12,270	0,018749	0,714	1,015	0,081
	$\hat{\sigma}_{pe}^2$	19,409	17,570	21,360	0,015080	0,585	1,022	0,180
	$\hat{\sigma}_{\varepsilon}^2$	48,465	47,780	49,140	0,004974	0,117	1,043	0,069
	$\hat{\sigma}_P^2$	78,096	76,530	79,720	0,013096	0,286	0,989	0,101
	h^2	0,131	0,104	0,156	0,000228	0,739	1,015	0,074
	c^2	0,249	0,224	0,272	0,000196	0,659	1,015	0,153
	Rep	0,379	0,368	0,393	0,000090	0,376	1,015	0,202

Fonte: Da autora.

¹ $\hat{\sigma}_{\gamma}^2$ -variância genética aditiva; $\hat{\sigma}_{pe}^2$ -variância atribuída aos efeitos permanente de ambiente; $\hat{\sigma}_{\varepsilon}^2$ -variância residual; $\hat{\sigma}_P^2$ -variância fenotípica; h^2 -herdabilidade; c^2 -proporção da variância fenotípica associada ao efeito permanente de ambiente; Rep -repetibilidade.

² LI - limite inferior; LS - limite superior de $HPD_{95\%}$.

³ $|ZG|$ - módulo da estatística de Geweke; RL -Fator de Raftery e Lewis; HW -valor-p do teste de Heidelberger e Welch.

As médias *a posteriori* das herdabilidades para PROD utilizando *priori* não informativa e informativa foram baixas, com valores iguais a 0,142 e 0,131, respectivamente. Considerando como *priori* não informativa, a amplitude do intervalo de credibilidade foi menor, indicando maior precisão. Estes valores são menores do que o relatado por [Penasa et al. \(2010\)](#) para produção de leite no dia do controle em uma população de 666 vacas da raça Burlina de primeira à oitava lactação pertencentes a dez rebanhos; e cujo valor da mediana da herdabilidade foi igual a 0,18 com $HPD_{95\%}$ entre 0,09 e 0,28.

Grande parte dos trabalhos com abordagem Bayesiana foram desenvolvidos considerando distribuições *a priori* não informativas e reportaram estimativas de herdabilidade para produção de leite acumulada na lactação, em análises multi-característica. Por exemplo, [Alijani et al. \(2012\)](#) obtiveram médias de herdabilidade para produção de leite na lactação iguais a 0,29, 0,21 e 0,19 considerando primeira, segunda e terceira lactação, respectivamente;

em um conjunto de 254.833 dados de vacas da raça Holandesa no Irã. [Zaabza et al. \(2016\)](#), em uma população da raça Holandesa (113.492 registros de 54.105 vacas) da Tunísia e utilizando modelo linear tri-característica, vestimaram herdabilidade igual a 0,273. Em trabalho mais recente de [Zaabza et al. \(2017\)](#), com 14.069 dados de lactações completas de 7.193 vacas de 11 rebanhos, as herdabilidades variaram entre 0,13 e 0,21.

Em rebanhos brasileiros da raça Holandesa, [Irano et al. \(2014\)](#) estimaram média de herdabilidade igual 0,19 ($HPD_{95\%}$ de 0,14 a 0,22) a partir de dados de 10.335 lactações de 5.090 vacas enquanto [Oliveira et al. \(2015\)](#) obtiveram média de herdabilidade igual a 0,16 ($HPD_{95\%}$ de 0,12 a 0,20) utilizando dados de 11.738 lactações de 5.084 vacas.

Apesar dos trabalhos citados anteriormente apresentarem resultados referentes à produção de leite acumulada aos 305 dias de lactação enquanto o presente estudo estimou os componentes de variância para produção de leite mensurada no dia do controle; as estimativas dos coeficientes de herdabilidade foram similares. A herdabilidade baixa estimada para produção de leite nesses trabalhos pode ser devida ao estudo de populações múltiparas, para as quais têm se estimado menores valores de herdabilidade menor comparativamente a populações formadas somente de vacas de primeira lactação, como observado em [Alijani et al. \(2012\)](#) e discutido em [Petrini et al. \(2016\)](#). No entanto, a produção de leite ainda é um dos critérios de seleção predominantes no melhoramento genético de bovinos leiteiros, já que representa a principal fonte de receita do sistema produtivo.

As estimativas dos componentes de variância, coeficiente de herdabilidade (h^2), repetibilidade (Rep) e proporção da variância fenotípica devido ao efeito permanente de ambiente (c^2) para a característica contínua de teor de caseína (CAS) estão resumidos na [Tabela 4](#).

Tabela 4 – Estimativas (média *a posteriori*) dos componentes de variância e parâmetros (componentes de variância e parâmetros genéticos) obtidos com *priori* não informativa (NI) e informativa (INF) com seus respectivos intervalos ($HPD_{95\%}$), critérios de convergência e erros de Monte Carlo (EMC) para a característica contínua de teor de caseína (CAS), período 2012 a 2016.

<i>Prioris</i>	Parâmetros ¹	Estimativas	$HPD_{95\%}$ ²		EMC	Critérios de convergência ³		
			LI	LS		$ ZG $	<i>RL</i>	<i>HW</i>
NI	$\hat{\sigma}_{\gamma}^2$	0,023	0,020	0,025	0,000018	0,786	0,987	0,125
	$\hat{\sigma}_{pe}^2$	0,004	0,003	0,006	0,000010	0,840	1,015	0,114
	$\hat{\sigma}_{\varepsilon}^2$	0,030	0,029	0,031	0,000003	0,399	1,107	0,060
	$\hat{\sigma}_P^2$	0,057	0,056	0,059	0,000012	0,446	1,032	0,122
	h^2	0,399	0,363	0,436	0,000248	0,846	0,966	0,148
	c^2	0,075	0,046	0,102	0,000185	0,834	0,982	0,123
	<i>Rep</i>	0,474	0,459	0,488	0,000174	0,576	1,015	0,116
INF	$\hat{\sigma}_{\gamma}^2$	0,015	0,014	0,017	0,000014	0,622	0,982	0,286
	$\hat{\sigma}_{pe}^2$	0,010	0,009	0,012	0,000008	1,224	0,982	0,624
	$\hat{\sigma}_{\varepsilon}^2$	0,030	0,029	0,031	0,000002	0,385	1,003	0,228
	$\hat{\sigma}_P^2$	0,056	0,055	0,057	0,000011	0,771	0,966	0,115
	h^2	0,271	0,245	0,298	0,000202	0,899	0,999	0,367
	c^2	0,187	0,166	0,208	0,000152	1,033	0,999	0,614
	<i>Rep</i>	0,458	0,445	0,470	0,000095	0,172	0,999	0,101

Fonte: Da autora.

¹ $\hat{\sigma}_{\gamma}^2$ -variância genética aditiva; $\hat{\sigma}_{pe}^2$ -variância atribuída aos efeitos permanente de ambiente; $\hat{\sigma}_{\varepsilon}^2$ -variância residual; $\hat{\sigma}_P^2$ -variância fenotípica; h^2 -herdabilidade; c^2 -proporção da variância fenotípica associada ao efeito permanente de ambiente; *Rep*-repetibilidade

² LI - limite inferior; LS - limite superior de $HPD_{95\%}$.

³ $|ZG|$ - módulo da estatística de Geweke; *RL*-Fator de Raftery e Lewis; *HW*-valor-p do teste de Heidelberger e Welch.

A média *a posteriori* da herdabilidade ($h^2 = 0,399$) para CAS com *priori* não informativa foi maior do que a média estimada da herdabilidade ($h^2 = 0,271$) com *priori* informativa. Nota-se que, a amplitude do $HPD_{95\%}$ foi ligeiramente menor na análise com *priori* informativa bem como o EMC, e os $HPD_{95\%}$ não encontram-se sobrepostos. Isto indica que as inferências foram afetadas pela inclusão de informações externas aos dados na análise.

O valor de h^2 na análise com *priori* informativa também foi mais próximo aos encontrados por Bittante, Cipolat-Gotet e Cecchinato (2013), Cecchinato *et al.* (2011). O primeiro utilizou o modelo linear animal multi-característica sob abordagem Bayesiana, relatando que a mediana da herdabilidade para o teor de caseína foi igual a 0,364 com ($HPD_{95\%}$ de 0,22 a 0,58) em uma população de 1.025 vacas da raça Holandesa, enquanto para 1.234 vacas da raça Pardo Suíço, a mediana da herdabilidade foi de 0,275 com ($HPD_{95\%}$ de 0,16 a 0,41). No segundo estudo, foi adotado um modelo linear bi-característica, também sob

abordagem Bayesiana, com dados de uma população de 1.167 vacas Pardo Suíço. A mediana da herdabilidade foi igual a 0,216 com $HPD_{95\%}$ de 0,10 a 0,38. Por fim, Bonfatti *et al.* (2011) obtiveram média da herdabilidade igual a 0,23 com intervalo de credibilidade ($HPD_{95\%}$) de 0,14 a 0,34 a partir de dados de 2.167 vacas da raça Simental e utilizando modelo bi-característica em abordagem Bayesiana.

Segundo Kawasaki, Lafont e Sire (2011), as caseínas são as proteínas mais abundantes do leite. Encontram-se na maioria dos mamíferos, possuindo um alto valor biológico. Especificamente na raça Holandesa, a composição do leite é, em média, de 3,2% de proteínas, das quais 80% são caseínas e 20% proteínas do soro de leite (CHANDAN, 1997). As caseínas estão associadas à coagulação do leite, sendo a base para a produção de queijos enquanto as demais proteínas do leite permanecem no soro. No entanto, o teor de caseína não tem sido incluído em índices de seleção de bovinos leiteiros. Segundo Bittante, Cipolat-Gotet e Cecchinato (2013), citando Ikonen *et al.* (2004) e Samoré *et al.* (2012); isto se deve à baixa variação da razão entre caseína e proteína total e à alta correlação entre teor de caseína e teor de proteína do leite.

As estimativas dos componentes de variância, coeficiente de herdabilidade (h^2), repetibilidade (Rep) e proporção da variância fenotípica devido ao efeito permanente de ambiente (c^2) para as característica binária de incidência de mastite subclínica (MSC) estão na Tabela 5.

As estimativas da média *a posteriori* da herdabilidade para a incidência de mastite subclínica (MSC) foram exatamente iguais para *priori* informativa e não informativa, com valores de 0,134 ($HPD_{95\%}$ entre 0,091 e 0,177) em ambos casos, significando que as informações *a priori* empregadas na *priori* informativa não influenciaram na inferência. A proporção da variância fenotípica associada ao efeito permanente de ambiente foram altas, com valores iguais a 0,358 ($HPD_{95\%}$ que variam entre 0,319 e 0,398) e 0,358 ($HPD_{95\%}$ que variam entre 0,320 e 0,399) sob *priori* não informativa e informativa, respectivamente (Tabela 5).

Estimativas de herdabilidade baixas também foram observadas para outras características associadas à saúde da vaca. Nos Estados Unidos, Vallimont *et al.* (2009), utilizaram modelo animal de limiar de repetibilidade uni-caraterística e estimaram herdabilidade para incidência de mastite clínica (alterações visíveis no leite e na glândula mamária) igual a 0,07. Na Espanha, Pérez-Cabal *et al.* (2009) obtiveram herdabilidade para

incidência de mastite clínica de 0,10 usando o modelo touro de repetibilidade e limiar incluindo sete lactações. Além disso, em [Hossein-Zadeh e Ardalán \(2011\)](#) as estimativas de herdabilidade para mastite clínica (classificada como 0 como ausência e, 1 com presença da doença), considerando vacas de primeira, segunda e terceira lactação, foram iguais a 0,06 com $HPD_{95\%}$ de 0,01 e 0,11, 0,08 $HPD_{95\%}$ entre 0,03 e 0,13, e 0,09 com $HPD_{95\%}$ entre 0,03 e 0,14, respectivamente. Nesse trabalho, foram utilizados registros de 57.301 vacas da raça Holandesa em modelos de limiar uni e bi-característica sob abordagem Bayesiana. Isto indica que fatores ambientais têm influência alta sobre tais características, de forma que o ganho genético nestas será mais lento e a predição acurada dependerá de um maior montante de informações fenotípicas e de pedigree comparativamente a características com herdabilidade alta.

Geralmente, nas avaliações genéticas a mastite clínica é uma característica binária, classificada como 1 se a vaca apresentou doença dentro de um determinado período, e 0, caso contrário ([IRANO *et al.*, 2014](#)). No entanto, de acordo com [Schukken *et al.* \(2003\)](#) essa doença pode ser identificada pelo aumento somente de CCS no leite.

Tabela 5 – Estimativas (média *a posteriori*) dos parâmetros (componentes de variância e parâmetros genéticos) obtidos com *prioris* não informativa (NI) e informativa (INF) com seus respectivos intervalos ($HPD_{95\%}$), critérios de convergência e erros de Monte Carlo (EMC) para a característica binária de incidência de mastite subclínica (MSC), período 2012 a 2016.

<i>Prioris</i>	Parâmetros ¹	Estimativas	$HPD_{95\%}$ ²		EMC	Critérios de convergência ³		
			LI	LS		$ ZG $	<i>RL</i>	<i>HW</i>
NI	$\hat{\sigma}_{\gamma}^2$	0,265	0,179	0,354	0,000668	1,833	0,973	0,216
	$\hat{\sigma}_{pe}^2$	0,707	0,621	0,781	0,000561	1,433	1,049	0,244
	$\hat{\sigma}_{\varepsilon}^2$	1,003	0,979	1,027	0,000145	0,139	0,982	0,280
	$\hat{\sigma}_P^2$	1,975	1,907	2,041	0,000556	0,763	0,966	0,051
	h^2	0,134	0,091	0,177	0,000321	1,821	0,966	0,218
	c^2	0,358	0,319	0,398	0,000282	1,624	1,032	0,292
	<i>Rep</i>	0,492	0,476	0,512	0,000119	1,032	1,032	0,136
INF	$\hat{\sigma}_{\gamma}^2$	0,265	0,179	0,354	0,000668	1,833	0,973	0,216
	$\hat{\sigma}_{pe}^2$	0,707	0,621	0,781	0,000561	1,433	1,049	0,244
	$\hat{\sigma}_{\varepsilon}^2$	1,003	0,979	1,027	0,000145	0,139	0,982	0,280
	$\hat{\sigma}_P^2$	1,975	1,907	2,041	0,000556	0,763	0,966	0,051
	h^2	0,134	0,091	0,177	0,000321	1,821	0,966	0,218
	c^2	0,358	0,320	0,399	0,000282	1,624	1,032	0,292
	<i>Rep</i>	0,492	0,476	0,510	0,000119	1,032	1,032	0,136

Fonte: Da autora.

¹ $\hat{\sigma}_{\gamma}^2$ -variância genética aditiva; $\hat{\sigma}_{pe}^2$ -variância atribuída aos efeitos permanente de ambiente; $\hat{\sigma}_{\varepsilon}^2$ -variância residual; $\hat{\sigma}_P^2$ -variância fenotípica, h^2 -herdabilidade; c^2 -proporção da variância fenotípica associada ao efeito permanente de ambiente; *Rep*-repetibilidade.

² LI - limite inferior; LS - limite superior de $HPD_{95\%}$

³ $|ZG|$ - módulo da estatística de Geweke; *RL*-Fator de Raftery e Lewis; *HW*-valor-p do teste de Heidelberger e Welch.

Valores altos de proporção da variância fenotípica associada ao efeito permanente de ambiente (c^2) foram estimados para PROD, CAS (quando *priori* informativa) e MSC, com valores entre 0,187 (CAS com *priori* informativa) e 0,358 (MSC) (Tabelas 3, 4 e 5). Para PROD e MSC, inclusive, os valores de c^2 foram maiores do que os coeficientes de herdabilidade estimados para as respectivas características. Este comportamento foi reportado também por [Penasa et al. \(2015\)](#) e [Zaabza, Gara e Rekik \(2018\)](#) para PROD, e [Petrini et al. \(2016\)](#) para PROD e escore de células somáticas (ECS, obtida como transformação da contagem de células somáticas). Em [Penasa et al. \(2015\)](#), 72.848 dados de 17.873 vacas da raça Holandesa foram analisados considerando um modelo de repetibilidade sob abordagem Bayesiana, sendo estimada média *a posteriori* de h^2 igual a 0,104 e c^2 igual a 0,31. Sob um modelo de regressão aleatória, com estimação Bayesiana, [Zaabza, Gara e Rekik \(2018\)](#) obteve valores de σ_a^2 e σ_{pe}^2 entre 1,04 e 11,87 kg/dia² e entre 6,08 e 38,65 kg/dia², respectivamente, de

acordo com o estágio e a ordem de lactação, a partir de 960.000 registros de três lactações de vacas da raça Holandesa na Tunísia. A partir de modelo animal de repetibilidade e utilizando REML para estimação, [Petrini et al. \(2016\)](#) obteve valores de h^2 e c^2 iguais a 0,12 e 0,29 para PROD e 0,12 e 0,30 para ECS, respectivamente, considerando 36.457 dados de 4.203 vacas da mesma população do presente estudo.

O efeito permanente de ambiente pode ser ocasionado por um impacto constante e acumulativo das condições de manejo e clima sob as quais as vacas estão submetidas ([ZAABZA; GARA; REKIK, 2018](#)). [Boonkum et al. \(2011\)](#) e mais recentemente [Cheruiyot et al. \(2019\)](#) verificaram um aumento na variância devido ao efeito permanente de ambiente sob condições de estresse térmico. Adicionalmente, tal efeito pode ter origem em condições experimentadas pelo indivíduo em um determinado estágio de sua vida ([KRUUK; HADFIELD, 2007](#)). Especificamente em bovinocultura leiteira, esse efeito pode ser associado a lesões no úbere, por exemplo ([MRODE, 2014](#)).

Adicionalmente, pode existir um confundimento entre o efeito permanente de ambiente e demais efeitos genéticos não inclusos no modelo de análise, como efeitos epigenéticos e efeitos não aditivos ([KRUUK; HADFIELD, 2007](#)). Em [Aliloo et al. \(2016\)](#), utilizando dados fenotípicos e genotípicos de 7.510 vacas da raça Holandesa, foi verificado que a variância devido ao efeito permanente de ambiente reduziu de 150,320; 255 e 171 no modelo aditivo para 118,360 litros², 200 e 143 kg² no modelo aditivo-dominante para produção de leite, gordura e proteína na lactação, respectivamente. No entanto, ressalta-se que a variância devido à dominância estimada tem sido baixa, com proporção da variância fenotípica variando entre 0,012 e 0,038 para características de produção; 0,000 e 0,154 para saúde; e 0,000 e 0,053 para reprodução em [Aliloo et al. \(2016\)](#) e também em [Jiang et al. \(2017\)](#) e [Sun et al. \(2014\)](#), respectivamente.

Efeitos a longo prazo na produção ocasionados por fatores ambientais e de manejo têm sido associados à efeitos epigenéticos que, por sua vez, podem ser contabilizados na variação devido ao efeito permanente de ambiente ([SINGH et al., 2012](#)). [Sejrsen et al. \(1982\)](#) demonstrou que o plano nutricional de uma novilha durante os seus nove meses de vida pode causar alterações no tecido mamário com conseqüente redução da produção de leite. Além disso, a nutrição durante a gestação da novilha pode influenciar as duas primeiras lactações ([FORD; PARK, 2001](#); [PARK et al., 1989](#)). Posteriormente foi sugerido que estes efeitos sobre a lactação seriam de origem epigenética ([PARK, 2005](#)), hipótese sustentada por diferenças nos

níveis de um modulador epigenético (5' metil-desoxicitidina) no tecido mamário de acordo com o manejo nutricional (CHOI *et al.*, 1998). Efeitos epigenéticos sobre a produção de leite e sobre a resposta imune foram ainda reportados em Ju *et al.* (2020), Singh *et al.* (2010), Singh *et al.* (2012) e Thompson-Crispi *et al.* (2014).

Os valores altos de repetibilidade (próximos a 0,40) para PROD, CAS e MSC indicam que as medidas tomadas ao longo do tempo referem-se à mesma característica.

As estimativas dos componentes de variância, coeficiente de herdabilidade (h^2) para precocidade (PREC), são exibidos na Tabela 6.

Tabela 6 – Estimativas (média *a posteriori*) dos parâmetros (componentes de variância e parâmetros genéticos) obtidos com *prioris* não informativa (NI) e informativa (INF) com seus respectivos intervalos ($HPD_{95\%}$), critérios de convergência e erros de Monte Carlo (EMC) para a característica binária de precocidade (PREC), período 2012 a 2016.

Prioris	Parâmetros ¹	Estimativas	$HPD_{95\%}$ ²		EMC	Critérios de convergência ³		
			LI	LS		$ ZG $	RL	HW
NI	$\hat{\sigma}_{\gamma}^2$	0,622	0,352	0,938	0,002496	0,867	1,158	0,074
	$\hat{\sigma}_{\varepsilon}^2$	1,000	0,944	1,064	0,000376	1,110	1,029	0,122
	$\hat{\sigma}_P^2$	1,626	1,328	1,943	0,002492	0,672	1,121	0,082
	h^2	0,377	0,264	0,486	0,000948	0,696	1,139	0,055
INF	$\hat{\sigma}_{\gamma}^2$	0,782	0,577	1,008	0,001667	0,882	1,049	0,082
	$\hat{\sigma}_{\varepsilon}^2$	1,000	0,946	1,061	0,000365	1,158	0,982	0,174
	$\hat{\sigma}_P^2$	1,787	1,565	2,021	0,001650	0,592	0,999	0,078
	h^2	0,436	0,368	0,504	0,000541	0,926	1,066	0,103

Fonte: Da autora.

¹ $\hat{\sigma}_{\gamma}^2$ -variância genética aditiva; $\hat{\sigma}_{pe}^2$ -variância atribuída aos efeitos permanente de ambiente; $\hat{\sigma}_{\varepsilon}^2$ -variância residual; $\hat{\sigma}_P^2$ -variância fenotípica, h^2 -herdabilidade.

² LI - limite inferior; LS - limite superior de $HPD_{95\%}$.

³ $|ZG|$ - módulo da estatística de Geweke; RL-Fator de Raftery e Lewis; HW-valor-p do teste de Heidelberger e Welch.

A média *a posteriori* da herdabilidade para a característica de precocidade foi alta, com valor igual a 0,377 e $HPD_{95\%}$ entre 0,264 e 0,486 (Tabela 6) para *priori* não informativa e h^2 igual a 0,436 e $HPD_{95\%}$ entre 0,368 e 0,504. O estudo da variação genética para características associadas à reprodução é relevante. Contudo, à exceção de precocidade e idade ao primeiro parto, tais características geralmente apresentam herdabilidades baixas. No Irã, Faraji-Arough *et al.* (2015) estimaram herdabilidade moderada para idade ao primeiro parto, com valor igual a 0,268 e intervalo de credibilidade de 0,260 a 0,275, a partir de 720.666 registros de lactação.

Por sua vez, [Zaabza et al. \(2016\)](#), estudando o intervalo de partos em vacas de rebanhos da Tunísia, obtiveram herdabilidade de 0,047, a partir de dados de 31.348 lactações. Ambos os trabalhos utilizaram vacas da raça Holandesa e modelo linear animal multi-característica sob abordagem Bayesiana.

As estimativas de herdabilidade alta para CAS e PREC indicam a existência de variabilidade aditiva com potencial de ser explorada em programas de seleção genética. Para PROD e MSC, as estimativas de herdabilidade de baixa magnitude ressaltam a importante influência de fatores não-genéticos sobre a variação fenotípica e, portanto, a necessidade de um maior volume de informações (principalmente de genealogia) para a predição acurada do valor genético nessas características.

As baixas estimativas de herdabilidade podem ser explicadas pela menor variação genética aditiva e ao aumento da variação devida ao efeito ambiental. Alimentação, manejo e condições climáticas as quais as vacas estão submetidas podem afetar ao fenótipo ([BERNABUCCI et al., 2015](#); [HAMMAMI et al., 2008](#); [NARDONE et al., 2010](#)).

Para a definição do melhor modelo para as características avaliadas com *priori* não informativa e informativa obteve-se os DIC. Os valores são apresentados na [Tabela 7](#).

Tabela 7 – Critério de informação do desvio (DIC) do modelo animal com *priori* não informativa (NI) e informativa (INF) para das características avaliadas de produção (PROD), teor de caseína (CAS), incidência de mastite subclínica (MSC) e precocidade (PREC), período 2012 a 2016.

Característica	<i>Prioris</i>	DIC	$\Delta = DIC_{NI} - DIC_{INF} $
PROD	NI	296.640,779	3,799
	INF	296.636,980	
CAS	NI	-34.330,746	23,546
	INF	-34.307,200	
MSC	NI	117.191,403	10^{-6}
	INF	117.191,403	
PREC	NI	3.912,577	457,461
	INF	3.455,115	

Da autora.

Para as características de PROD e MSC, a diferença absoluta entre o DIC obtido nas análises sob *priori* não informativa e informativa é menor do que de 5 unidades, indicando que pelo critério DIC não existe diferença na utilização das *prioris*. No entanto, comparando-se as amplitudes dos intervalos $HPD_{95\%}$ e o EMC, o modelo com *priori* informativa apresenta maior precisão. Por sua vez, a diferença absoluta entre os DIC para as características CAS e

PREC foram maiores do que 10 unidades, de forma que para CAS o melhor modelo foi com *priori* não informativa e para PREC com *priori* informativa, respectivamente.

A introdução das *prioris* informativas na estimação dos componentes de variância e parâmetros genéticos ocasionou em redução na amplitude dos intervalos de credibilidade e do erro de Monte Carlo para as características em estudo, sugerindo aumento na precisão comparativamente ao uso de *priori* não informativa.

De modo geral, a inclusão de *prioris* informativas para os parâmetros de interesse pode fazer com que as estimativas *a posteriori* sejam mais coerentes com a realidade da variável de interesse (ROSSI, 2011). Isto é corroborado por Paulino, Turkman e Murteira (2003) que mencionam que o uso da *priori* informativa permite que as amplitudes dos intervalos de credibilidade sejam mais estreitos para os parâmetros de interesse. Fonseca *et al.* (2006), Carneiro Júnior *et al.* (2005) e Silva *et al.* (2008), ao incorporar *prioris* informativas, encontraram estimativas superiores em comparação às obtidas com as *prioris* não informativas.

Com base nisso, as tendências genéticas das características foram analisadas com base nos valores genéticos preditos a partir das análises com *prioris* informativas.

4.3 TENDÊNCIA GENÉTICA

Os critérios de convergência de todas as características foram confirmadas por meio da visualização gráfica (Apêndice C) bem como pelos critérios de convergência, indicando que há evidências suficientes de que a convergência das cadeias foram atingidas. Os valores da estatística de Geweke ($|ZG|$) foram menores do que 1,96, não apresentando indícios de não convergência. Pelo critério de Raftery e Lewis (*RL*), verificou-se que os valores dos fatores foram próximos de 1, evidenciando a independência entre iterações. Constatou-se ainda que a série é estacionária pelo valor-p do teste de Heidelberger e Welch (*HW*).

Após da verificação dos critérios de convergência, as estimativas dos parâmetros de todas as características em estudo encontram-se resumidas na Tabela 8.

Tabela 8 – Média *a posteriori* dos parâmetros da regressão linear dos valores genéticos sobre o ano de nascimento, com seus respectivos intervalos de credibilidade ($HPD_{95\%}$) e critérios de convergência.

Caraterística	Parâmetro ¹	Estimativa	$HPD_{95\%}$ ²		Critérios de convergência ³		
			LI	LS	$ ZG $	RL	HW
PROD	$\hat{\beta}_0$	5,651	-90,350	101,000	0,311	0,999	0,967
	$\hat{\beta}_1$	-0,003	-0,051	0,045 <i>NS</i>	0,311	1,000	0,967
	τ	9,283	2,701	16,770	1,350	1,000	0,754
CAS	$\hat{\beta}_0$	-27,154	-40,440	-14,230	1,882	1,000	0,623
	$\hat{\beta}_1$	0,014	0,007	0,020	1,879	1,000	0,600
	τ	496,454	143,200	896,000	0,575	1,000	0,612
MSC	$\hat{\beta}_0$	-13,682	-48,990	20,850	0,146	1,000	0,925
	$\hat{\beta}_1$	0,007	-0,011	0,024 <i>NS</i>	0,145	1,000	0,925
	τ	69,899	20,150	126,100	0,374	1,000	0,587
PREC	$\hat{\beta}_0$	-64,622	-113,600	-16,380	0,644	1,000	0,810
	$\hat{\beta}_1$	0,032	0,008	0,056	0,648	1,000	0,810
	τ	36,026	10,450	65,070	0,693	1,000	0,278

Fonte: Da autora.

¹ $\hat{\beta}_0$ -intercepto; $\hat{\beta}_1$ -coeficiente angular estimado; τ -precisão

² LI - limite inferior; LS - limite superior de $HPD_{95\%}$.

³ $|ZG|$ - módulo da estatística de Geweke; RL -Fator de Raftery e Lewis; HW -valor-p do teste de Heidelberger e Welch.

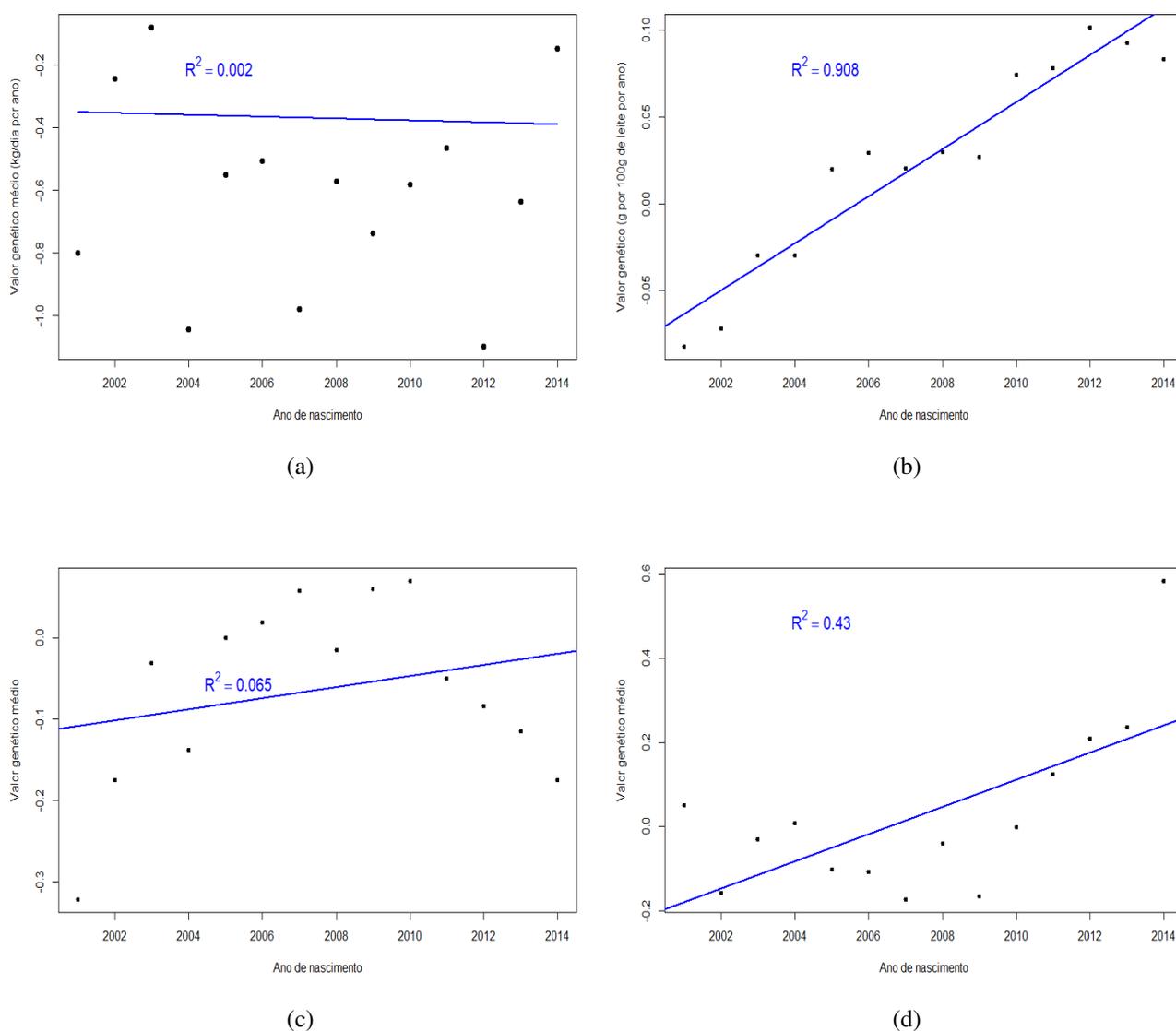
NS-não significativo

Analisando-se a [Tabela 8](#), pode-se verificar por meio do $HPD_{95\%}$ que os parâmetros da regressão linear dos valores genéticos para PROD e MSC sobre o ano de nascimento dos animais avaliados não foram significativos, visto que estes intervalos contém o valor zero. Isto indica que não há ganho genético nestas características. No entanto, para CAS e PREC, características de herdabilidade alta, estes parâmetros foram significativos, apesar dos valores preditos não se ajustarem à reta ([Figura 2](#)).

Os gráficos da tendência genética são apresentados na [Figura 2](#) para PROD, CAS, MSC e PREC. Observa-se que para PROD as médias dos valores genéticos não apresentam um comportamento crescente ou decrescente ao longo do período observado enquanto que para CAS as médias dos valores genéticos aumentaram no mesmo período. Isto era esperado visto que o critério de seleção geralmente utilizado para produção é o volume de leite acumulado na lactação (e não produção no dia do controle leiteiro). Em relação à caseína, é provável que o ganho genético seja resultado da seleção para teor de proteína no leite, sendo que ambas possuem correlação genética alta (próxima a 0,99) e herdabilidade alta (aproximada 0,40)

(PETRINI, 2016). O ganho genético anual para CAS foi igual à 0,014 g/100g de leite por ano, ou seja, 0,59% da média da característica.

Figura 2 – Tendências genéticas preditos e valores genéticos para as características contínuas de produção de leite (PROD) (a), teor de caseína (CAS) (b), incidência de mastite subclínica (MSC) (c) e a precocidade (PREC) (d), período de 2001 a 2014.



Fonte: Da autora.

Similarmente à PROD, para MSC não foi verificado um comportamento linear para as médias dos valores genéticos em relação ao ano de nascimento, indicativo da ausência de progresso genético nessa característica. No entanto, houve um decréscimo na média do valor genético a partir do ano de 2010. Para PREC a tendência genética foi positiva, com aumento

do valor genético principalmente a partir também de 2010. Esse resultado sugere uma possível mudança recente quanto aos critérios de seleção nessas propriedades, incorporando características relacionadas à saúde e reprodução.

5 CONCLUSÕES

Por meio dos resultados encontrados, é possível concluir que os modelos lineares e de limiar sob abordagem Bayesiana foram adequados na obtenção dos componentes de variância das características estudadas, principalmente a partir do uso de *prioris* informativas. Utilizando tamanho efetivo da amostra igual a 5.000 para cada parâmetro desconhecido do modelo não foram encontrados problemas de convergência das cadeias e os erros de Montecarlo foram baixos.

Os componentes de variância genética estimados para teor de caseína e precocidade indicam variabilidade genética, justificando a inclusão das mesmas em programas de melhoramento genético de bovinos da raça Holandesa visando à melhoria da qualidade do leite e da fertilidade do rebanho.

Por sua vez, os coeficientes de herdabilidade baixos para produção de leite no dia do controle e incidência de mastite subclínica associados à alta proporção da variância fenotípica devido ao efeito permanente de ambiente demonstram a alta influência do ambiente sobre o fenótipo dessas. Neste caso, faz-se importante a informação de pedigree e de características geneticamente correlacionadas para a predição acurada do mérito genético dos animais para essas características.

As tendências genéticas obtidas para as características avaliadas indicam que a seleção realizada vem promovendo um progresso genético lento, talvez pela pouca intensidade aplicada na seleção dos animais ou mesmo pelo uso de material genético não avaliado sob as condições de manejo utilizadas. Assim, é fundamental que os criadores adotem técnicas adequadas de avaliação genética e seleção para a obtenção de maiores ganhos genéticos.

Futuramente, deve-se investigar o ajuste de modelos que incluam outros componentes da variação genética, como o efeito de dominância, de forma a melhorar a estimação do componente genético aditivo e verificar a proporção da variância fenotípica associada a cada um desses componentes. A incorporação de outras fontes de informação (como a informação genômica) também podem auxiliar à avaliação genética.

Os resultados obtidos neste estudo podem ser utilizados para o planejamento de programas de melhoramento genético que visem à melhoria da produção, composição do leite, saúde e fertilidade das vacas leiteiras da raça Holandesa, com o objetivo de aumentar o ganho genético, a eficiência produtiva e, conseqüentemente, a rentabilidade do sistema de produção.

REFERÊNCIAS

- ABDEL-RADY, A.; SAYED, M. Epidemiological Studies on Subclinical Mastitis in Dairy Cows in Assiut Governorate. **Veterinary World**, v. 2, n. 10, 2009. Disponível em: <<https://pdfs.semanticscholar.org/8122/40e3221f029bfcf0bb71bb6e2b5e5886b480.pdf>>. Acesso em: 01 Jun. 2019.
- ALIJANI, S. *et al.* Estimation of variance components for some production traits of Iranian Holstein dairy cattle using bayesian and AI-REML methods. **Pakistan Veterinary Journal**, v. 32, n. 4, p. 562–566, 2012. Disponível em: <https://www.researchgate.net/profile/Hossein_Daghigh_Kia/publication/249964346_Pakistan_Veterinary_Journal_Estimation_of_variance_components_for_some_production_traits_of_Iranian_Holstein_dairy_cattle_using_Bayesian_and_AI-REML_methods/links/00b7d51e7762f4a13a000000.pdf>. Acesso em: 10 Oct. 2019.
- ALILLOO, H. *et al.* Accounting for dominance to improve genomic evaluations of dairy cows for fertility and milk production traits. **Genetics Selection Evolution**, v. 48, n. 1, p. 8, 2016. Disponível em: <<https://doi.org/10.1186/s12711-016-0186-0>>. Acesso em: 11 Oct. 2019.
- ARANGO, J.; ECHEVERRI, J. Asociación del valor genético del toro con caracteres productivos en vacas lecheras en Colombia. **Archivos de Zootecnia**, v. 63, n. 242, p. 227–237, 2014. Disponível em: <<http://dx.doi.org/10.4321/S0004-05922014000200001>>. Acesso em: 11 Oct. 2019.
- ARAÚJO, C. de A. *et al.* Perfil tecnológico de sistemas de produção de leite de bovino no São Alagoano. **Diversitas Journal**, v. 4, n. 1, p. 31–38, 2019. Disponível em: <<https://doi.org/10.17648/diversitas-journal-v4i1.660>>. Acesso em: 14 May. 2019.
- AYALEW, W.; ALIY, M.; NEGUSSIE, E. Estimation of genetic parameters of the productive and reproductive traits in Ethiopian Holstein using multi-trait models. **Asian-Australasian Journal of Animal Sciences**, v. 30, n. 11, p. 1550, 2017. Disponível em: <<https://doi.org/10.5713/ajas.17.0198>>. Acesso em: 14 May. 2019.
- BEIJO, L. A.; VIVANCO, M. J. F.; MUNIZ, J. A. Análise bayesiana no estudo do tempo de retorno das precipitações pluviiais máximas em Jaboticabal (SP). **Ciência e Agrotecnologia**, v. 33, n. 1, p. 261–270, 2009. Disponível em: <<https://doi.org/10.1590/S1413-70542009000100036>>. Acesso em: 14 May. 2019.
- BERNABUCCI, U. *et al.* Effect of summer season on milk protein fractions in Holstein cows. **Journal of Dairy Science**, v. 98, n. 3, p. 1815–1827, 2015. Disponível em: <<https://doi.org/10.3168/jds.2014-8788>>. Acesso em: 12 Set. 2019.
- BERRY, D.; CROMIE, A. Associations between age at first calving and subsequent performance in Irish ring calving Holstein–Friesian dairy cows. **Livestock Science**, v. 123, n. 1, p. 44–54, 2009. Disponível em: <<https://doi.org/10.1016/j.livsci.2008.10.005>>. Acesso em: 10 Set. 2019.
- BERRY, D. P. *et al.* Negative influence of high maternal milk production before and after

conception on offspring survival and milk production in dairy cattle. **Journal of Dairy Science**, v. 91, n. 1, p. 329–337, 2008. Disponível em: <<https://doi.org/10.3168/jds.2007-0438>>. Acesso em: 18 Set. 2019.

BIRD, A. DNA methylation patterns and epigenetic memory. **Genes & Development**, v. 16, n. 1, p. 6–21, 2002. Disponível em: <<https://doi.org/10.1101/gad.947102>>. Acesso em: 03 Feb. 2020.

BITTANTE, G.; CIPOLAT-GOTET, C.; CECCHINATO, A. Genetic parameters of different measures of cheese yield and milk nutrient recovery from an individual model cheese-manufacturing process. **Journal of Dairy Science**, v. 96, n. 12, p. 7966–7979, 2013. Disponível em: <<https://doi.org/10.3168/jds.2012-6517>>. Acesso em: 03 Feb. 2020.

BLASCO, A. **Bayesian data analysis for animal scientists**. [S.l.]: Springer, 2017.

BOAKE, C. R. Repeatability: its role in evolutionary studies of mating behavior. **Evolutionary Ecology**, v. 3, n. 2, p. 173–182, 1989. Disponível em: <<https://link.springer.com/content/pdf/10.1007/BF02270919.pdf>>. Acesso em: 13 Jan. 2020.

BONFATTI, V. *et al.* Genetic analysis of detailed milk protein composition and coagulation properties in Simmental cattle. **Journal of Dairy Science**, v. 94, n. 10, p. 5183–5193, 2011. Disponível em: <<https://doi.org/10.3168/jds.2011-4297>>. Acesso em: 23 Jul. 2019.

BOONKUM, W. *et al.* Genetic effects of heat stress on milk yield of thai Holstein crossbreds. **Journal of Dairy Science**, v. 94, n. 1, p. 487–492, 2011. Disponível em: <<https://doi.org/10.3168/jds.2010-3421>>. Acesso em: 26 Jul. 2019.

BOURDON, R. M. **Understanding animal breeding**. [S.l.]: Prentice Hall Upper Saddle River, NJ, 2000. v. 2.

BRASIL, R. B. *et al.* Estrutura e estabilidade das micelas de caseína do leite bovino. **Ciência animal**, Universidade Estadual do Ceará, p. 71–80, 2015. Disponível em: <<http://200.137.217.156/bitstream/ri/12468/5/Artigo%20-%20Rafaela%20Belchior%20Brasil%20-%202015.pdf>>. Acesso em: 23 May. 2019.

CARDOSO, V. *et al.* Breeding goals and economic values for pasture based milk production systems in the southeast region of Brazil. **Revista Brasileira de Zootecnia**, Viçosa, v. 33, n. 2, p. 320–327, 2004. Disponível em: <<https://doi.org/10.1590/S1516-35982005000600014>>. Acesso em: 20 Feb. 2020.

CARNEIRO JÚNIOR, J. M. *et al.* Influência da informação a priori na avaliação genética animal utilizando dados simulados. **R. Bras. Zootec**, v. 34, n. 6, p. 1905–1913, 2005. Disponível em: <<https://doi.org/10.1590/S1516-35982005000600014>>. Acesso em: 02 Dez. 2019.

CARRARA, E. R. *et al.* Genetic and phenotypic trends for milk fatty acids in a Holstein cattle population reared under tropical conditions. **Livestock Science**, v. 228, p. 84–92, 2019. Disponível em: <<https://doi.org/10.1016/j.livsci.2019.08.006>>. Acesso em: 15 Oct. 2019.

CECCHINATO, A. *et al.* Genetic parameters of coagulation properties, milk yield, quality, and acidity estimated using coagulating and noncoagulating milk information in brown swiss and Holstein-friesian cows. **Journal of Dairy Science**, v. 94, n. 8, p. 4205–4213, 2011. Disponível em: <<https://doi.org/10.3168/jds.2010-3913>>. Acesso em: 09 Oct. 2019.

CHANDAN, R. Properties of milk and its components. **Dairy-Based Ingredients**, p. 1–10, 1997.

CHERUIYOT, E. *et al.* Genotype-by-environment (temperature-humidity) interaction of milk production traits in Australian Holstein cattle. **Journal of Dairy Science**, 2019. Disponível em: <<https://doi.org/10.3168/jds.2019-17609>>. Acesso em: 31 Jan. 2020.

CHOI, Y. *et al.* Effects of compensatory growth on the expression of milk protein gene and biochemical changes of the mammary gland in Holstein cows. **The Journal of Nutritional Biochemistry**, v. 9, n. 7, p. 380–387, 1998. Disponível em: <[https://doi.org/10.1016/S0955-2863\(98\)00027-8](https://doi.org/10.1016/S0955-2863(98)00027-8)>. Acesso em: 31 Ago. 2019.

COSTA, A. *et al.* Heritability and repeatability of milk lactose and its relationships with traditional milk traits, somatic cell score and freezing point in Holstein cows. **Animal**, v. 13, n. 5, p. 909–916, 2018. Disponível em: <<https://doi.org/10.1017/S1751731118002094>>. Acesso em: 25 Ago. 2019.

DOHM, M. Repeatability estimates do not always set an upper limit to heritability. **Functional Ecology**, v. 16, n. 2, p. 273–280, 2002. Disponível em: <<https://doi.org/10.1046/j.1365-2435.2002.00621.x>>. Acesso em: 23 Jan. 2020.

ELER, J. P. **Teorias e métodos em melhoramento genético animal: bases do melhoramento genético animal**. Faculdade de Zootecnia e Engenharia de Alimentos da USP, 2017. Disponível em: <<http://www.livrosabertos.sibi.usp.br/portaldelivrosUSP/catalog/book/162>>. Acesso em: 05 May. 2019.

ETTEMA, J.; SANTOS, J. Impact of age at calving on lactation, reproduction, health, and income in first-parity Holsteins on commercial farms. **Journal of Dairy Science**, v. 87, n. 8, p. 2730–2742, 2004. Disponível em: <[https://doi.org/10.3168/jds.S0022-0302\(04\)73400-1](https://doi.org/10.3168/jds.S0022-0302(04)73400-1)>. Acesso em: 05 Feb. 2020.

FALCONER, D. **Introdução à genética quantitativa**. [S.l.]: UFV, 1987.

FALCONER, D.; MACKAY, T. **Introduction to Quantitative Genetics**. [S.l.: s.n.], 1996.

FARAJI-AROUGH, H. *et al.* Bayesian inference of (co) variance components and genetic parameters for economic traits in Iranian Holsteins via gibbs sampling. **Iranian Journal of Applied Animal Science**, v. 5, n. 1, 2015. Disponível em: <http://ijas.iaurasht.ac.ir/article_513421_558cc4a49784e0d74ec896d539da8b41.pdf>. Acesso em: 17 Oct. 2019.

FARAWAY, J. J. **Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models**. [S.l.]: Chapman and Hall/CRC, 2016.

FERRAZ, J. B. S.; ELER, J. P. **Melhoramento genético para aumento de produtividade em gado de corte no Brasil: a história, o presente e o futuro**. [S.l.]: Bovinocultura de corte. v. 2, 2010.

FONSECA, R. d. *et al.* Influência da informação à priori na escolha de modelos para avaliação genética de suínos. **Ciência e Agrotecnologia**, v. 30, n. 3, p. 538–546, 2006. Disponível em: <<https://doi.org/10.1590/S1413-70542006000300021>>. Acesso em: 01 Feb. 2020.

FORD, J.; PARK, C. Nutritionally directed compensatory growth enhances heifer development and lactation potential. **Journal of Dairy Science**, v. 84, n. 7, p. 1669 – 1678, 2001. ISSN 0022-0302. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0022030201746024>>.

GAMERMAN, D.; LOPES, H. F. **Markov chain Monte Carlo: stochastic simulation for Bayesian inference**. [S.l.]: Chapman and Hall/CRC, 2006.

GARCIA, C. H. **Aplicação de equações de modelos mistos em testes clonais de Eucalyptus SPP**. 2004. Dissertação (Mestrado em Estatística e Experimentação Agronômica) — Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2004. Disponível em: <<https://doi.org/10.11606/D.11.2004.tde-06012005-144944>>. Acesso em: 19 Jun. 2019.

GASQUE, R. G. **Enciclopedia Bovina**. 1. ed. [S.l.]: Universidad Nacional Autónoma de México Facultad de Medicina Veterinaria y Zootecnia Ciudad Universitaria México, 2008.

GAYA, C. *et al.* Milk casein content from milk recorded Holstein cows in Parana state/porcentagem de caseína no leite de vacas Holandesas em controle leiteiro no estado do Paraná. **Veterinaria e Zootecnia**, v. 20, n. 2, p. 56–58, 2013. Disponível em: <<https://go.gale.com/ps/anonymou?id=GALE%7CA387607060&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=01025716&p=IFME&sw=w>>. Acesso em: 19 Jun. 2019.

GELMAN, A. *et al.* **Bayesian data analysis**. [S.l.]: Chapman and Hall/CRC, 2013.

GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments. **Bayesian Statistics**, v. 4, p. 641–649, 1992. Disponível em: <<https://ideas.repec.org/p/fip/fedmsr/148.html>>. Acesso em: 19 Jul. 2019.

GIANOLA, D.; FERNANDO, R. L. Bayesian methods in animal breeding theory. **Journal of Animal Science**, v. 63, n. 1, p. 217–244, 1986. Disponível em: <<https://doi.org/10.2527/jas1986.631217x>>. Acesso em: 19 Abr. 2019.

GUTIERREZ, J. P. G. **Iniciación a la valoración genética animal metodología adaptada al EEES**. [S.l.], 2010.

HADFIELD, J.; NAKAGAWA, S. General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. **Journal of Evolutionary Biology**, v. 23, n. 3, p. 494–508, 2010. Disponível em: <<https://doi.org/10.1111/j.1420-9101.2009.01915.x>>. Acesso em: 02 Jan. 2020.

- HAMMAMI, H. *et al.* Genetic parameters for Tunisian Holsteins using a test-day random regression model. **Journal of Dairy Science**, v. 91, n. 5, p. 2118–2126, 2008. Disponível em: <<https://doi.org/10.3168/jds.2007-0382>>. Acesso em: 31 Oct. 2019.
- HARMON, R. Physiology of mastitis and factors affecting somatic cell counts. **Journal of Dairy Science**, v. 77, n. 7, p. 2103–2112, 1994. Disponível em: <[https://doi.org/10.3168/jds.S0022-0302\(94\)77153-8](https://doi.org/10.3168/jds.S0022-0302(94)77153-8)>. Acesso em: 14 Jan. 2020.
- HEIDELBERGER, P.; WELCH, P. D. Simulation run length control in the presence of an initial transient. **Operations Research**, v. 31, n. 6, p. 1109–1144, 1983. Disponível em: <<https://www.jstor.org/stable/170841>>. Acesso em: 13 Jul. 2019.
- HENDERSON, C. Best linear unbiased estimation and prediction under a selection model. **Biometrics**, p. 423–447, 1975. Disponível em: <<https://doi.org/10.1093/ansci/1973.Symposium.10>>. Acesso em: 15 Abr. 2019.
- HENDERSON, C. *et al.* The estimation of environmental and genetic trends from records subject to culling. **Biometrics**, v. 15, n. 2, p. 192–218, 1959. Disponível em: <<https://www.jstor.org/stable/2527669>>. Acesso em: 13 Abr. 2019.
- HENDERSON, C. R. Sire evaluation and genetic trends. **Journal of Animal Science**, v. 1973, n. Symposium, p. 10–41, 1973. Disponível em: <<https://www.jstor.org/stable/2529430>>. Acesso em: 14 Abr. 2019.
- HERINGSTAD, B. *et al.* Bayesian analysis of liability of clinical mastitis in norwegian cattle with a threshold model: Effects of data sampling method and model specification. **Journal of Dairy Science**, v. 84, n. 11, p. 2337–2346, 2001. Disponível em: <[https://doi.org/10.3168/jds.S0022-0302\(01\)74682-6](https://doi.org/10.3168/jds.S0022-0302(01)74682-6)>. Acesso em: 15 Dez. 2019.
- HOFFMANN, R. **Análise de regressão: uma introdução à econometria**. 5. ed. Piracicaba: [s.n.], 2016. Disponível em: <<http://www.livrosabertos.sibi.usp.br/portaldelivrosUSP/catalog/book/73>>. Acesso em: 15 Dez. 2019.
- HOSSEIN-ZADEH, N. G.; ARDALAN, M. Bayesian estimates of genetic parameters for metritis, retained placenta, milk fever, and clinical mastitis in Holstein dairy cows via GIBBS sampling. **Research in Veterinary Science**, v. 90, n. 1, p. 146–149, 2011. Disponível em: <<https://doi.org/10.1016/j.rvsc.2010.04.018>>. Acesso em: 19 Oct. 2019.
- HUEY, R. B.; DUNHAM, A. E. Repeatability of locomotor performance in natural populations of the lizard *Sceloporus Merriami*. **Evolution**, v. 41, n. 5, p. 1116–1120, 1987. Disponível em: <<https://doi.org/10.1111/j.1558-5646.1987.tb05880.x>>. Acesso em: 21 Jan. 2020.
- HUTCHISON, J. *et al.* Genomic evaluation of age at first calving. **Journal of Dairy Science**, v. 100, n. 8, p. 6853–6861, 2017. Disponível em: <<https://doi.org/10.3168/jds.2016-12060>>. Acesso em: 21 Jan. 2020.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Produção da pecuária municipal**, Rio de Janeiro, v. 46, p. 1–8, 2018. Disponível em: <<https://biblioteca.ibge.gov.br/>>

[visualizacao/periodicos/84/ppm_2018_v46_br_informativo.pdf](#)>. Acesso em: 21 May. 2020.

IKONEN, T. *et al.* Genetic and phenotypic correlations between milk coagulation properties, milk production traits, somatic cell count, casein content, and pH of milk. **Journal of Dairy Science**, v. 87, n. 2, p. 458–467, 2004. Disponível em: <[https://doi.org/10.3168/jds.S0022-0302\(04\)73185-9](https://doi.org/10.3168/jds.S0022-0302(04)73185-9)>. Acesso em: 19 Ago. 2019.

IRANO, N. *et al.* Genetic association between milk yield, stayability, and mastitis in Holstein cows under tropical conditions. **Tropical Animal Health and Production**, v. 46, n. 3, p. 529–535, 2014. Disponível em: <<https://doi.org/10.1007/s11250-013-0524-y>>. Acesso em: 27 Oct. 2019.

JEFFREYS, H. **Theory of Probability**. [S.l.: s.n.], 1961.

JIANG, J. *et al.* Dissection of additive, dominance, and imprinting effects for production and reproduction traits in Holstein cattle. **BMC Genomics**, v. 18, n. 1, p. 425, 2017. Disponível em: <<https://doi.org/10.1186/s12864-017-3821-4>>. Acesso em: 09 Jun. 2019.

JU, Z. *et al.* Genome-wide methylation and transcriptome of blood neutrophils reveal the roles of DNA methylation in affecting transcription of protein-coding genes and mirnas in e. coli-infected mastitis cows. **BMC Genomics**, v. 21, n. 1, p. 102, 2020. Disponível em: <<https://doi.org/10.1186/s12864-020-6526-z>>. Acesso em: 09 Jul. 2019.

KAWASAKI, K.; LAFONT, A.-G.; SIRE, J.-Y. The Evolution of Milk Casein Genes from Tooth Genes before the Origin of Mammals. **Molecular Biology and Evolution**, v. 28, n. 7, p. 2053–2061, 2011. Disponível em: <<https://doi.org/10.1093/molbev/msr020>>. Acesso em: 03 Jul. 2019.

KELLEHER, M. *et al.* Additive genetic, non-additive genetic and permanent environmental effects for female reproductive performance in seasonal calving dairy females. **Irish Journal of Agricultural and Food Research**, v. 55, n. 1, p. 10–23, 2016. Disponível em: <<https://doi.org/10.1515/ijafr-2016-0002>>. Acesso em: 04 Jul. 2019.

KRUUK, L.; HADFIELD, J. How to separate genetic and environmental causes of similarity between relatives. **Journal of Evolutionary Biology**, v. 20, n. 5, p. 1890–1903, 2007. Disponível em: <<https://doi.org/10.1111/j.1420-9101.2007.01377.x>>. Acesso em: 01 Jul. 2019.

LACERDA, J. J. D. *et al.* Parâmetros e tendências genéticas para características de crescimento em bovinos da raça Nelore no estado da Bahia utilizando inferência bayesiana. **Revista Brasileira de Saúde e Produção Animal**, v. 15, n. 1, p. 10 – 19, 2014. Disponível em: <<https://doi.org/10.1590/S1519-99402014000100003>>. Acesso em: 25 Ago. 2019.

LAUREANO, M. *et al.* Estimativas de herdabilidade e tendências genéticas para características de crescimento e reprodutivas em bovinos da raça Nelore: Estimates of heritability and genetic trends for growth and reproduction traits in Nelore cattle. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 63, n. 1, p. 143–152, 2011. Disponível em: <<https://doi.org/10.1590/S0102-09352011000100022>>. Acesso em: 14 Dez. 2019.

LIGGES, U. *et al.* **Package ‘R2OpenBUGS’**, 2017. Disponível em: <<https://cran.biodisk.org/web/packages/R2OpenBUGS/R2OpenBUGS.pdf>>. Acesso em: 17 Dez. 2019.

LINK, W. A.; BARKER, R. J. **Bayesian inference: with ecological applications**. [S.l.]: Academic Press, 2010.

LOPES, P. S. *et al.* **Teoria do melhoramento animal**. [S.l.: s.n.], 2005. 89 p.

M-ROCHA, J. F. *et al.* Estimación de parámetros genéticos para edad al primer parto e intervalo entre partos en poblaciones bovinas de la raza Blanco Orejinegro (BON) en Colombia. **Revista Colombiana de Ciencias Pecuarias**, Colombia, v. 25, n. 2, p. 220–228, 2012. Disponível em: <<https://www.redalyc.org/pdf/2950/295023555007.pdf>>. Acesso em: 22 Oct. 2019.

MARTINEZ, C. A.; PERDOMO, C. M.; ELZO, M. A. Cattle genetic evaluation: a historical perception. **Revista Colombiana de Ciencias Pecuarias**, Colombia, v. 25, n. 2, p. 293–311, 2012. Disponível em: <<https://redalyc.org/pdf/2950/295023555014.pdf>>. Acesso em: 17 Nov. 2019.

MISZTAL, I. Computational techniques in animal breeding. **University of Georgia, Athens, GA**, 2016. Disponível em: <<http://nce.ads.uga.edu/~ignacy/course2002/notes.pdf>>. Acesso em: 01 Abr. 2019.

MISZTAL, I. *et al.* **Manual for BLUPF90 family of programs**. Georgia, USA: Athens: University of Georgia, 2019. Disponível em: <http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all2.pdf>. Acesso em: 17 Nov. 2019.

MRODE, R. A. **Linear Models for the Prediction of Animal Breeding Values**. 3. ed. London: Cabi, 2014.

MUÑOZ, M. F. *et al.* **Modelación aplicada a las ciencias animales: II. Evaluaciones genéticas**. [S.l.]: Fondo Editorial Biogénesis, 2012. 1–137 p.

MUÑOZ, R.; GONZÁLES, C. **Guía Práctica de Evaluación Genética Animal con Virtual BLUP**. [S.l.]: Servicios Informáticos y Tecnologías Web, E.I.R.L., 2016.

NARDONE, A. *et al.* Effects of climate changes on animal production and sustainability of livestock systems. **Livestock Science**, v. 130, n. 1-3, p. 57–69, 2010. Disponível em: <<https://doi.org/10.1016/j.livsci.2010.02.011>>. Acesso em: 22 Oct. 2019.

NILFOROOSHAN, M.; EDRISS, M. Effect of age at first calving on some productive and longevity traits in Iranian Holsteins of the isfahan province. **Journal of Dairy Science**, v. 87, n. 7, p. 2130–2135, 2004. Disponível em: <[https://doi.org/10.3168/jds.S0022-0302\(04\)70032-6](https://doi.org/10.3168/jds.S0022-0302(04)70032-6)>. Acesso em: 29 Oct. 2019.

NOGUEIRA, D.; SÁFADI, T.; FERREIRA, D. Avaliação de critérios de convergência para o método de Monte Carlo via cadeias de Markov. **Revista Brasileira de Estatística**, v. 65, n. 224, 2004.

NOGUEIRA, D. A. *et al.* Análises clássica e bayesiana de um modelo misto aplicado ao melhoramento animal: uma ilustração. **Ciência e Agrotecnologia**, Lavras, v. 27, p. 1614–1624, 2003. Disponível em: <http://pointer.esalq.usp.br/departamentos/lce/arquivos/aulas/2012/LCE5872/ANALISES_CLASSICA_E_BAYESIANA_DE_UM_MODELO_MISTO_melhoramento_animal.PDF>. Acesso em: 13 Mar. 2019.

NTZOUFRAS, I. **Bayesian modeling using WinBUGS**. [S.l.]: John Wiley & Sons, 2011.

ØDEGÅRD, J. *et al.* A simple algorithm to estimate genetic variance in an animal threshold model using bayesian inference. **Genetics Selection Evolution**, v. 42, n. 1, p. 29, 2010. Disponível em: <<https://doi.org/10.1186/1297-9686-42-29>>. Acesso em: 13 Mar. 2019.

OLIVEIRA, C. *et al.* Prevalência e etiologia da mastite bovina na bacia leiteira de Rondon do Pará, estado do Pará. **Pesquisa Veterinária Brasileira**, v. 31, n. 2, p. 104–110, 2011. Disponível em: <<https://doi.org/10.1590/S0100-736X2011000200002>>. Acesso em: 24 Dez. 2019.

OLIVEIRA, E. *et al.* Associação genética entre ocorrência de mastite clínica e produção de leite em vacas Holandesas. **Ciência Rural**, v. 45, n. 12, p. 2187–2192, 2015. Disponível em: <<https://doi.org/10.1590/0103-8478cr20131639>>. Acesso em: 23 Dez. 2019.

PAIVA, J. *et al.* Herdabilidade de características de crescimento em bovinos da raça Nelore utilizando métodos da máxima verossimilhança restrita e inferência bayesiana. **Archivos de Zootecnia**, v. 68, n. 263, p. 440–446, 2019. Disponível em: <<https://doi.org/10.21071/az.v68i263.4206>>. Acesso em: 26 Ago. 2019.

PARK, C. *et al.* Role of compensatory growth in lactation: a stair-step nutrient regimen modulates differentiation and lactation of bovine mammary gland. **Growth, Development, and Aging: GDA**, v. 53, n. 4, p. 159–166, 1989. Disponível em: <<https://doi.org/10.1016/j.livsci.2014.12.014>>. Acesso em: 28 Oct. 2019.

PARK, C. S. Role of compensatory mammary growth in epigenetic control of gene expression. **The FASEB Journal**, v. 19, n. 12, p. 1586–1591, 2005. Disponível em: <<https://doi.org/10.1096/fj.05-3816hyp>>. Acesso em: 26 Oct. 2019.

PAULINO, C. D. M.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística Bayesiana**. Lisboa, Portugal: [s.n.], 2003. 446 p.

PEÑA, E. C. Q.; RUIZ, A. L. **Predicción de valores de cría de animales domésticos**. 2. ed. Peru: Editorial Bravo Impresores, 2018. 365 p.

PENASA, M. *et al.* Bayesian inference of genetic parameters for test-day milk yield, milk quality traits, and somatic cell score in Burlina cows. **Journal of Applied Genetics**, v. 51, n. 4, p. 489–495, 2010. Disponível em: <<https://doi.org/10.1007/BF03208878>>. Acesso em: 24 Oct. 2019.

_____. Genetics of milk fatty acid groups predicted during routine data recording in Holstein dairy cattle. **Livestock Science**, v. 173, p. 9–13, 2015. Disponível em:

<https://doi.org/10.1016/j.livsci.2014.12.014>. Acesso em: 28 Oct. 2019.

PEREZ, B. C. *et al.* Count bayesian models for genetic analysis of in vitro embryo production traits in Guzerá cattle. **Animal**, v. 11, n. 9, p. 1440–1448, 2017. Disponível em: <https://doi.org/10.1017/S175173111700012X>. Acesso em: 28 Oct. 2019.

PÉREZ-CABAL, M. *et al.* Genetic evaluation of susceptibility to clinical mastitis in Spanish Holstein cows. **Journal of Dairy Science**, v. 92, n. 7, p. 3472–3480, 2009. Disponível em: <https://doi.org/10.3168/jds.2008-1978>. Acesso em: 01 Nov. 2019.

PETRINI, J. **Incorporação de informações genômicas no desenvolvimento de índices econômicos para a seleção de bovinos leiteiros**. 2016. Tese (Doutorado em ciências) — Escola Superior de Agricultura Luiz de Queiroz/Universidade de São Paulo, Piracicaba, 2016. Disponível em: https://teses.usp.br/teses/disponiveis/11/11139/tde-09062016-181056/publico/Juliana_Petrini_versao_revisada.pdf. Acesso em: 08 Sep. 2019.

PETRINI, J. *et al.* Genetic parameters for milk fatty acids, milk yield and quality traits of a Holstein cattle population reared under tropical conditions. **Journal of Animal Breeding and Genetics**, v. 133, n. 5, p. 384–395, 2016. Disponível em: <https://doi.org/10.1111/jbg.12205>. Acesso em: 08 Nov. 2019.

PIRLO, G.; MIGLIOR, F.; SPERONI, M. Effect of age at first calving on production traits and on difference between milk yield returns and rearing costs in Italian Holsteins. **Journal of Dairy Science**, v. 83, n. 3, p. 603–608, 2000. Disponível em: [https://doi.org/10.3168/jds.S0022-0302\(00\)74919-8](https://doi.org/10.3168/jds.S0022-0302(00)74919-8). Acesso em: 07 Set. 2019.

PLUMMER, M. *et al.* CODA: Convergence diagnosis and output analysis for mcmc. **R News**, v. 6, n. 1, p. 7–11, 2006. Disponível em: <https://journal.r-project.org/archive/>. Acesso em: 07 Dez. 2019.

PRESTES, D.; FILAPPI, A.; CECIM, M. Susceptibilidade à mastite: fatores que a influenciam-uma revisão. **Revista da FZVA**, v. 9, n. 1, 2002. Disponível em: <http://revistaseletronicas.pucrs.br/ojs/index.php/fzva/article/view/2153/1662>. Acesso em: 02 Jan. 2020.

PRIETO, R. A. **Rendirse no es una opción**. [S.l.]: Amat Editorial, 2016. 190 p.

QUIJANO, J. H.; ECHEVERRI, J. J. **Genética cuantitativa aplicada al mejoramiento animal**. [S.l.]: Universidad Nacional de Colombia, 2015.

RAFTERY, A. E.; LEWIS, S. **How many iterations in the Gibbs sampler?** [S.l.], 1992. 1–14 p. Disponível em: <https://pdfs.semanticscholar.org/0daf/54c4b59fd2c362de822de0ffdab84f49c6fd.pdf>. Acesso em: 02 May. 2019.

_____. Comment: One long run with diagnostics: Implementation strategies for Markov chain. **Statistical Science**, v. 7, n. 4, p. 493–497, 1992b. Disponível em: <https://www.stat.washington.edu/~raftery/Research/PDF/gelman-comment1992.pdf>. Acesso em: 03 May. 2020.

RESENDE, J. *et al.* Produção e produtividade de leite no mundo. In: RENTERO, N. (Ed.). **Anuario Leite**. EMBRAPA, 2019. Disponível em: <<https://www.embrapa.br/gado-de-leite>>. Acesso em: 02 Feb. 2019.

RESENDE, M.; SILVA, F. e.; AZEVEDO, C. **Estatística matemática, biométrica e computacional: modelos mistos, multivariados, categóricos e generalizados (reml/blup), inferência bayesiana, regressão aleatória, seleção genômica, qtl-gwas, estatística espacial e temporal, competição, sobrevivência**. [S.l.: s.n.], 2014.

RIBEIRO, L. H. M. *et al.* Bayesian modelling of the effects of nitrogen doses on the morphological characteristics of braquiaria grass. **Revista Agro@ambiente On-line**, v. 12, n. 4, p. 245–257, 2018. Disponível em: <[10.18227/1982-8470ragro.v12i4.5166](https://doi.org/10.18227/1982-8470ragro.v12i4.5166)>. Acesso em: 21 Feb. 2020.

ROBERT, C. **The Bayesian choice: from decision-theoretic foundations to computational implementation**. 2. ed. [S.l.]: Springer Science & Business Media, 2007.

ROSSI, R. M. **Introdução aos métodos Bayesianos na análise de dados zootécnicos com uso do WinBUGS e R**. [S.l.]: Eduem: Maringá, 2011.

SAMAD, M. **Animal Husbandry and Veterinary Science**. Mymensingh, Bangladesh: Mrs Mahfuza Bulbul, 2008. 1783 p.

SAMORÉ, A. B. *et al.* Genetics of casein content in brown swiss and Italian Holstein dairy cattle breeds. **Italian Journal of Animal Science**, v. 11, n. 2, p. 195–202, 2012. Disponível em: <<https://doi.org/10.4081/ijas.2012.e36>>. Acesso em: 08 Ago. 2019.

SANTOS, G.; LOPES, M. A. Custos de produção de fêmeas bovinas leiteiras do nascimento ao primeiro parto. **Ciência Animal Brasileira**, Goiânia, v. 15, p. 11–19, 03 2014. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1809-68912014000100002&nrm=iso>. Acesso em: 05 Feb. 2020.

SCHAEFFER, L. Cumulative permanent environmental effects for repeated records animal models. **Journal of Animal Breeding and Genetics**, v. 128, n. 2, p. 95–99, 2011. Disponível em: <<https://doi.org/10.1111/j.1439-0388.2010.00894.x>>. Acesso em: 01 Feb. 2020.

SCHUKKEN, Y. H. *et al.* Monitoring udder health and milk quality using somatic cell counts. **Veterinary Research**, v. 34, n. 5, p. 579–596, 2003. Disponível em: <<https://doi.org/10.1051/vetres:2003028>>. Acesso em: 11 Jan. 2020.

SEEGERS, H.; FOURICHON, C.; BEAUDEAU, F. Production effects related to mastitis and mastitis economics in dairy cattle herds. **Veterinary Research**, USA, v. 34, n. 5, p. 475–491, 2003. Disponível em: <<https://doi.org/10.1051/vetres:2003027>>. Acesso em: 02 Feb. 2020.

SEJRSEN, K. *et al.* Influence of nutrition on mammary development in pre-and postpubertal heifers. **Journal of Dairy Science**, v. 65, n. 5, p. 793–800, 1982. Disponível em: <[https://doi.org/10.3168/jds.S0022-0302\(82\)82268-6](https://doi.org/10.3168/jds.S0022-0302(82)82268-6)>. Acesso em: 29 Jan. 2020.

SEJRSEN, K.; PURUP, S. Influence of prepubertal feeding level on milk yield potential of dairy heifers: a review. **Journal of Animal Science**, v. 75, n. 3, p. 828–835, 1997. Disponível em: <<https://doi.org/10.2527/1997.753828x>>. Acesso em: 30 Jan. 2020.

SILVA, F. F. *et al.* Comparação bayesiana de modelos de previsão de diferenças esperadas nas progênes no melhoramento genético de gado Nelore. **Pesquisa Agropecuária Brasileira**, v. 43, n. 1, p. 37–45, 2008. Disponível em: <<https://doi.org/10.1590/S0100-204X2008000100006>>. Acesso em: 29 Abr. 2019.

_____. Bayesian inference of mixed models in quantitative genetics of crop species. **Theoretical and Applied Genetics**, v. 126, n. 7, p. 1749–1761, 2013. Disponível em: <<https://doi.org/10.1007/s00122-013-2089-6>>. Acesso em: 21 Abr. 2019.

SILVA, H. T. *et al.* Alternative count bayesian models for genetic evaluation of litter traits in pigs. **Livestock Science**, v. 225, p. 140–143, 2019. Disponível em: <<https://doi.org/10.1016/j.livsci.2019.05.006>>. Acesso em: 24 Abr. 2019.

SILVA, M. A. G. **Modelos mistos na seleção de famílias de cana de açúcar aparentadas sob o enfoque clássico e Bayesiano**. 2012. Dissertação (Magister Scientiae) — Universidade Federal de Viçosa, Viçosa, Minas Gerais, 2012. Disponível em: <<https://doi.org/10.2527/1997.753828x>>. Acesso em: 21 May. 2019.

SINGH, K. *et al.* Epigenetic regulation of milk production in dairy cows. **Journal of Mammary Gland Biology and Neoplasia**, v. 15, n. 1, p. 101–112, 2010. Disponível em: <<https://doi.org/10.1007/s10911-010-9164-2>>. Acesso em: 04 Jul. 2019.

_____. Epigenetics: a possible role in acute and transgenerational regulation of dairy cow milk production. **Animal**, v. 6, n. 3, p. 375–381, 2012. Disponível em: <<https://doi.org/10.1017/S1751731111002564>>. Acesso em: 10 Jul. 2019.

SMITH, A. F. Bayesian methods in reliability. In: SANTERR, P.; BADOUX, R. (Ed.). **Topics in safety, reliability and quality**. London: [s.n.], 1991.

SMITH, B. J. *et al.* BOA: an R package for MCMC output convergence assessment and posterior inference. **Journal of Statistical Software**, v. 21, n. 11, p. 1–37, 2007. Disponível em: <<http://www.jstatsoft.org/>>. Acesso em: 10 Ago. 2019.

SORENSEN, D.; GIANOLA, D. **Likelihood, Bayesian, and MCMC methods in quantitative genetics**. [S.l.]: Springer Science & Business Media, 2002.

SPIEGELHALTER, D. J. *et al.* Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v. 64, n. 4, p. 583–639, 2002. Disponível em: <<https://doi.org/10.1111/1467-9868.00353>>. Acesso em: 20 Dez. 2019.

STONE, J. V. **Bayes' rule: A tutorial introduction to Bayesian analysis**. [S.l.]: Sebtel Press, 2013.

SUN, C. *et al.* Improvement of prediction ability for genomic selection of dairy

cattle by including dominance effects. **PloS One**, v. 9, n. 8, 2014. Disponível em: <<https://doi.org/10.1371/journal.pone.0103934>>. Acesso em: 03 May. 2019.

R Core Team. **R: a language and environment for statistical computing.**, Viena, Austria, 2019. Disponível em: <<https://www.R-project.org/>>. Acesso em: 17 May. 2019.

THOMAS, A.; O'HARA, R. B. **OpenBUGS**, 2004. Disponível em: <<http://www.openbugs.net/w/Downloads%3E>>. Acesso em: 10 Ago. 2019.

THOMPSON-CRISPI, K. *et al.* Bovine mastitis: frontiers in immunogenetics. **Frontiers in Immunology**, v. 5, p. 493, 2014. Disponível em: <<https://doi.org/10.3389/fimmu.2014.00493>>. Acesso em: 02 Ago. 2019.

TIEZZI, F. *et al.* Heritability and repeatability of milk coagulation properties predicted by mid-infrared spectroscopy during routine data recording, and their relationships with milk yield and quality traits. **Animal**, v. 7, n. 10, p. 1592–1599, 2013. Disponível em: <<https://doi.org/10.1017/S1751731113001195>>. Acesso em: 15 Jan. 2020.

TOFFANIN, V. *et al.* Genetic parameters for milk mineral content and acidity predicted by mid-infrared spectroscopy in Holstein–Friesian cows. **Animal**, v. 9, n. 5, p. 775–780, 2015. Disponível em: <<https://doi.org/10.1017/S1751731114003255>>. Acesso em: 15 Dez. 2019.

TOZER, P.; HEINRICHS, A. J. What affects the costs of raising replacement dairy heifers: a multiple-component analysis. **Journal of Dairy Science**, v. 84, n. 8, p. 1836–1844, 2001. Disponível em: <[https://doi.org/10.3168/jds.S0022-0302\(01\)74623-1](https://doi.org/10.3168/jds.S0022-0302(01)74623-1)>. Acesso em: 05 Nov. 2019.

VALLIMONT, J. *et al.* Heritability estimates associated with alternative definitions of mastitis and correlations with somatic cell score and yield. **Journal of Dairy Science**, v. 92, n. 7, p. 3402–3410, 2009. Disponível em: <<https://doi.org/10.3168/jds.2008-1229>>. Acesso em: 10 Jan. 2020.

VANDERICK, S. **Improvement of genetic evaluation systems for maternally influenced traits and multi-breed livestock populations.** 2017. Tese (Sciences agronomiques et ingénierie biologique) — Université de Liège, Liège, Belgique, 2017. Disponível em: <<http://hdl.handle.net/2268/208997>>. Acesso em: 10 Jan. 2020.

VARSHNEY, J.; NARESH, R. Evaluation of a homeopathic complex in the clinical management of udder diseases of riverine Buffaloes. **Homeopathy**, v. 93, n. 01, p. 17–20, 2004. Disponível em: <[10.1016/j.homp.2003.11.007](https://doi.org/10.1016/j.homp.2003.11.007)>. Acesso em: 12 Jan. 2020.

VAYEGO, S. **Uso de Modelos Mistos na Avaliação Genética de Linhagens de Matrizes de Frango de Corte.** 2007. 104 p. Tese (Tese (Doutorado em Melhoramento Genético Animal)) — Universidade Federal do Paraná, Curitiba, 2007.

VAZQUEZ, A. *et al.* an r package for fitting generalized linear mixed models in animal breeding. **Journal of Animal Science**, v. 88, n. 2, p. 497–504, 2010. Disponível em: <<https://doi.org/10.2527/jas.2009-1952>>. Acesso em: 10 Abr. 2019.

VENTURA, H. T. *et al.* Comparing multi-trait poisson and gaussian bayesian models for genetic evaluation of litter traits in pigs. **Livestock Science**, v. 176, p. 47–53, 2015. Disponível em: <<https://doi.org/10.1016/j.livsci.2015.03.030>>. Acesso em: 10 Abr. 2019.

VILLEMEREUIL, P. On the relevance of bayesian statistics and MCMC for animal models. **Journal of Animal Breeding and Genetics**, v. 136, n. 5, p. 339–340, 2019. Disponível em: <<https://onlinelibrary.wiley.com/doi/pdfdirect/10.1111/jbg.12426>>. Acesso em: 21 Jul. 2019.

WHITE, T. L.; HODGE, G. R. **Predicting breeding values with applications in forest tree improvement**. [S.l.]: Dordrecht: Kluwer Academic, 1989.

WOLFFE, A. P.; GUSCHIN, D. Chromatin structural features and targets that regulate transcription. **Journal of Structural Biology**, v. 129, n. 2-3, p. 102–122, 2000. Disponível em: <<https://doi.org/10.1006/jsbi.2000.4217>>. Acesso em: 21 Jul. 2019.

ZAABZA, H. B. *et al.* Estimation of variance components of milk, fat, and protein yields of Tunisian Holstein dairy cattle using bayesian and REML methods. **Archives Animal Breeding**, v. 59, n. 2, p. 243–248, 2016. Disponível em: <<https://doi.org/10.5194/aab-59-243-2016>>. Acesso em: 14 May. 2019.

_____. Multiple-trait estimation of genetic parameters of yield traits of dairy cattle in Tunisia using an animal model. **Journal of New Sciences**, v. 42, 2017. Disponível em: <<https://doi.org/10.5713/ajas.17.0198>>. Acesso em: 14 Oct. 2019.

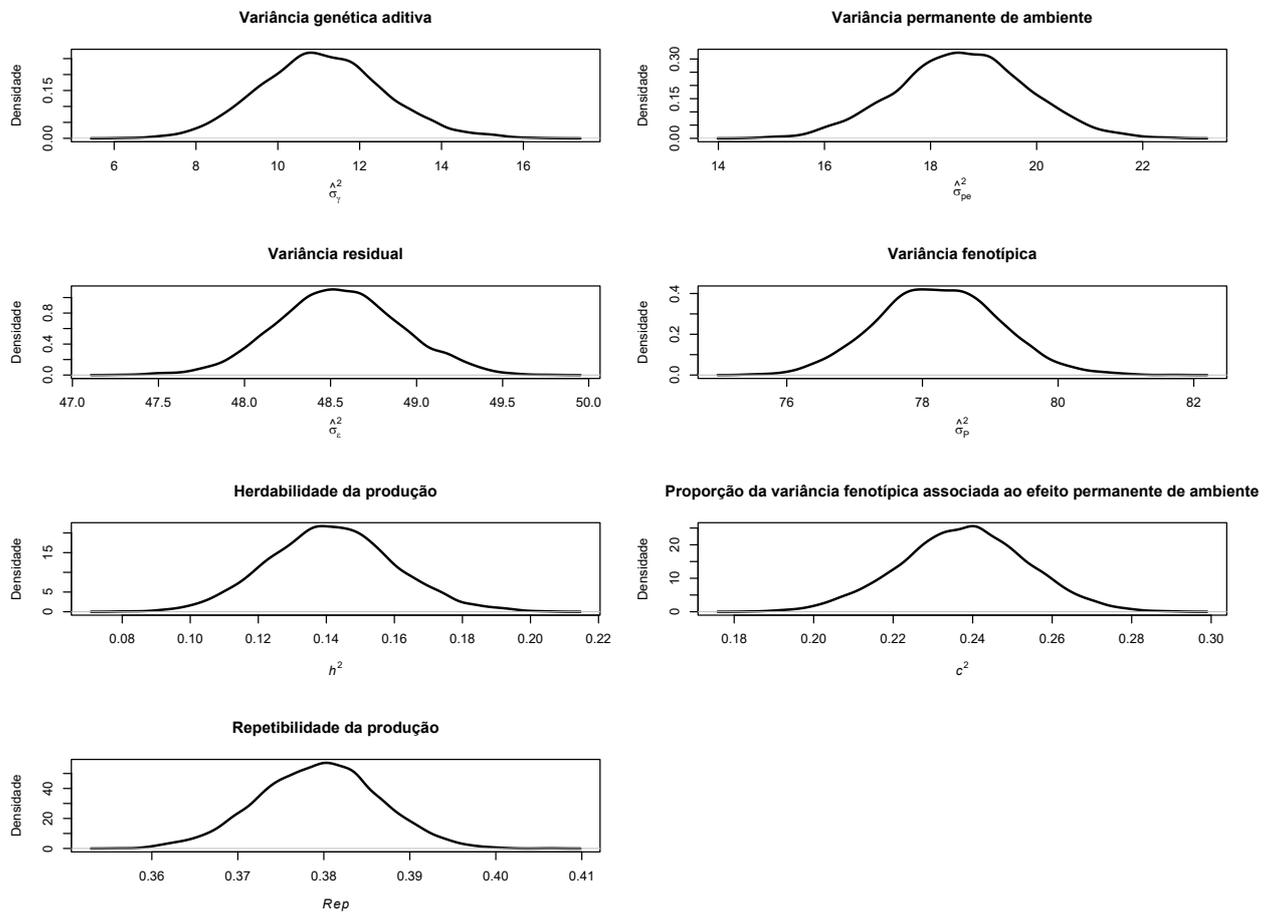
ZAABZA, H. B.; GARA, A. B.; REKIK, B. Genetic analysis of milk production traits of Tunisian Holsteins using random regression test-day model with legendre polynomials. **Asian-Australasian Journal of Animal Sciences**, v. 31, n. 5, p. 636, 2018. Disponível em: <<https://doi.org/10.5713/ajas.17.0332>>. Acesso em: 27 Oct. 2019.

ZOCCAL, R. Menos vacas ordenhadas e mais produção de leite. In: RENTERO, N. (Ed.). **Anuario Leite**. EMBRAPA, 2018. Disponível em: <<http://ainfo.cnptia.embrapa.br/digital/bitstream/item/198698/1/Anuario-LEITE-2019.pdf>>. Acesso em: 27 May. 2019.

APÊNDICES

APÊNDICE A - Gráficos de densidades das cadeias *a posteriori* dos parâmetros $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_p^2$, h^2 , c^2 e Rep resultantes do modelo de repetibilidade (linear) para as características contínuas de produção de leite e caseína com *priori* não informativa e informativa

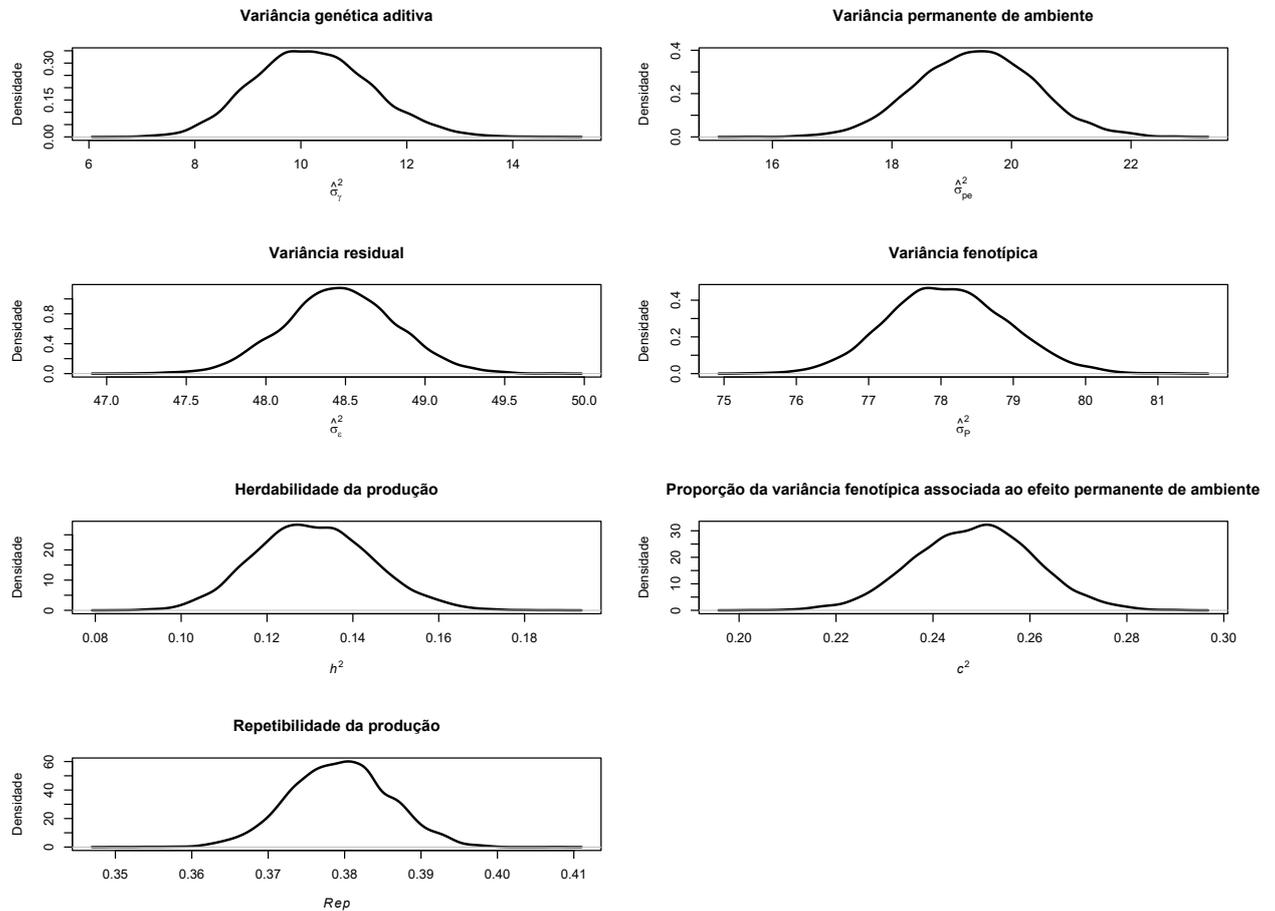
Figura 3 – Gráficos de densidade das cadeias *a posteriori* dos parâmetros estimados $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_p^2$, h^2 , c^2 e Rep resultante da característica de produção de leite, utilizando *a priori* não informativa, período 2012 a 2016.



Fonte: Da autora.

$\hat{\sigma}_\gamma^2$ -variância genética aditiva; $\hat{\sigma}_{pe}^2$ -variância atribuída aos efeitos de ambiente permanente; $\hat{\sigma}_\varepsilon^2$ -variância residual; $\hat{\sigma}_p^2$ -variância fenotípica; h^2 -herdabilidade; c^2 -proporção da variância fenotípica associada ao efeito permanente de ambiente; Rep -Repetibilidade.

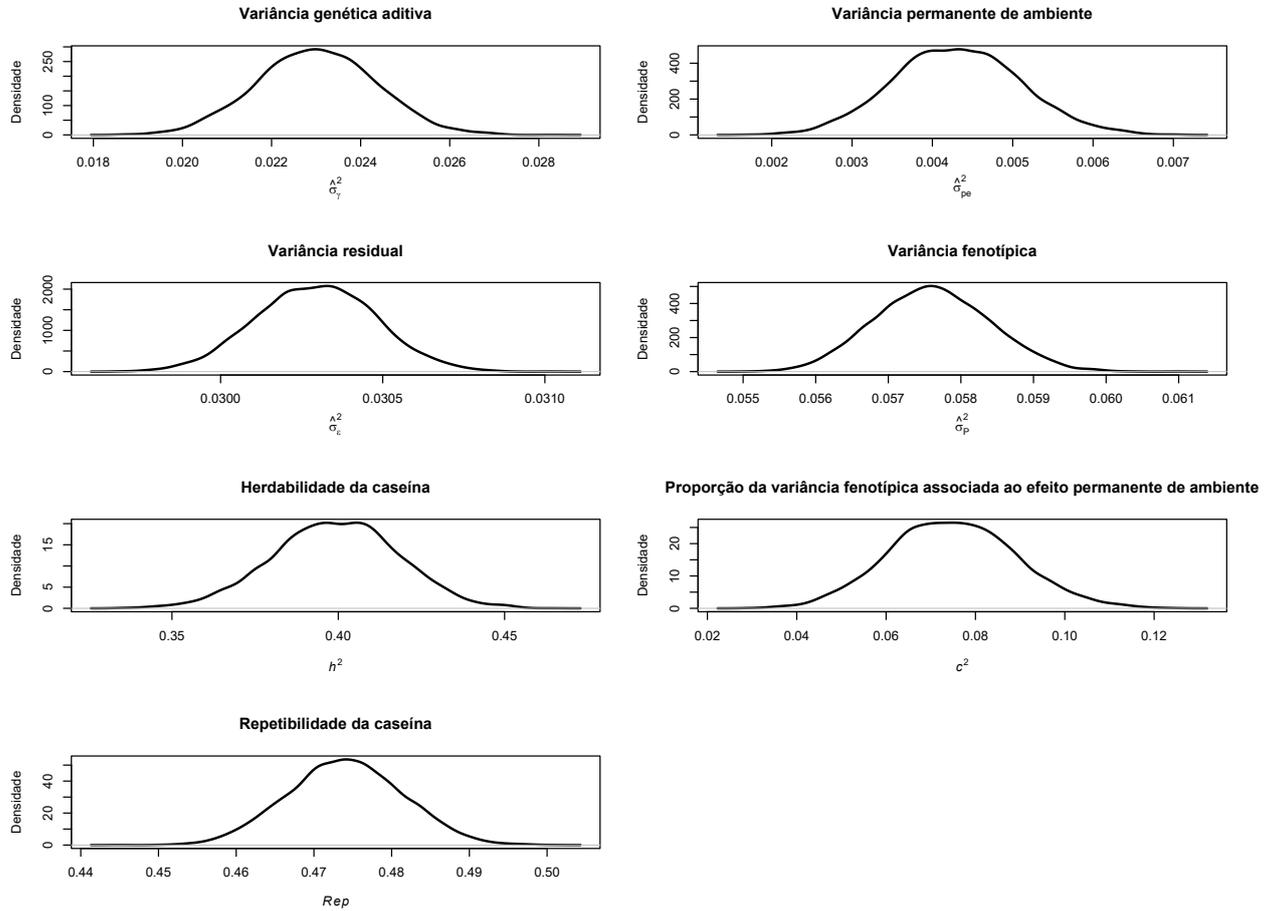
Figura 4 – Gráficos de densidade das cadeias *a posteriori* dos parâmetros estimados $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_P^2$, h^2 , c^2 e Rep resultante da característica de produção de leite, utilizando *a priori* informativa, período 2012 a 2016.



Fonte: Da autora.

$\hat{\sigma}_\gamma^2$ -variância genética aditiva; $\hat{\sigma}_{pe}^2$ -variância atribuída aos efeitos de ambiente permanente; $\hat{\sigma}_\varepsilon^2$ -variância residual; $\hat{\sigma}_P^2$ -variância fenotípica; h^2 -herdabilidade; c^2 -proporção da variância fenotípica associada ao efeito permanente de ambiente; Rep -Repetibilidade.

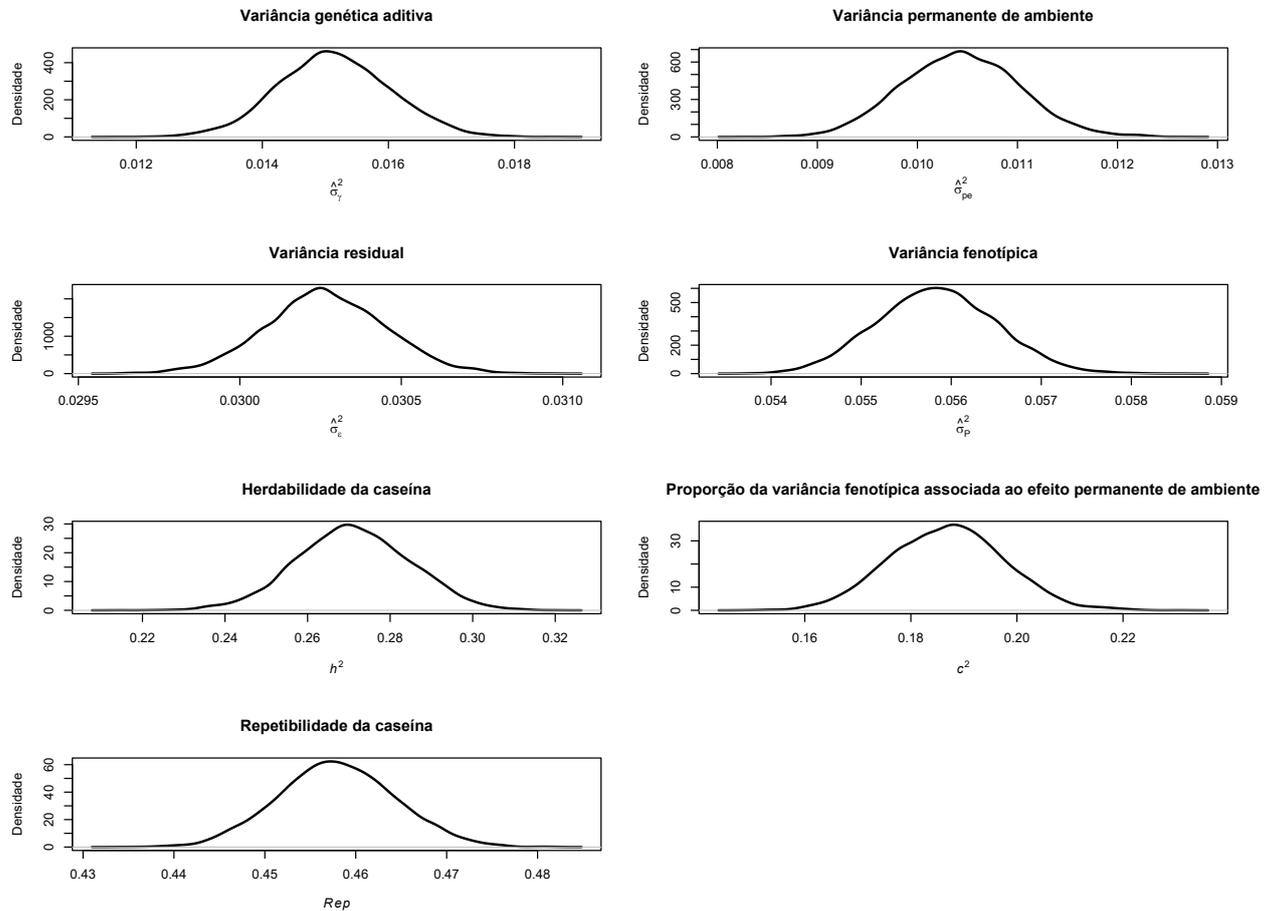
Figura 5 – Gráficos de densidade das cadeias *a posteriori* dos parâmetros estimados $\hat{\sigma}_{\gamma}^2$, $\hat{\sigma}_{\varepsilon}^2$, $\hat{\sigma}_P^2$, h^2 , c^2 e Rep resultante da característica de caseína, utilizando *a priori* não informativa, período 2012 a 2016.



Fonte: Da autora.

$\hat{\sigma}_{\gamma}^2$ -variância genética aditiva; $\hat{\sigma}_{pe}^2$ -variância atribuída aos efeitos de ambiente permanente; $\hat{\sigma}_{\varepsilon}^2$ -variância residual; $\hat{\sigma}_P^2$ -variância fenotípica; h^2 -herdabilidade; c^2 -proporção da variância fenotípica associada ao efeito permanente de ambiente; Rep -Repetibilidade.

Figura 6 – Gráficos de densidade das cadeias *a posteriori* dos parâmetros estimados $\hat{\sigma}_{\gamma}^2$, $\hat{\sigma}_{\varepsilon}^2$, $\hat{\sigma}_P^2$, h^2 , c^2 e Rep resultante da característica de caseína, utilizando *a priori* informativa, período 2012 a 2016.

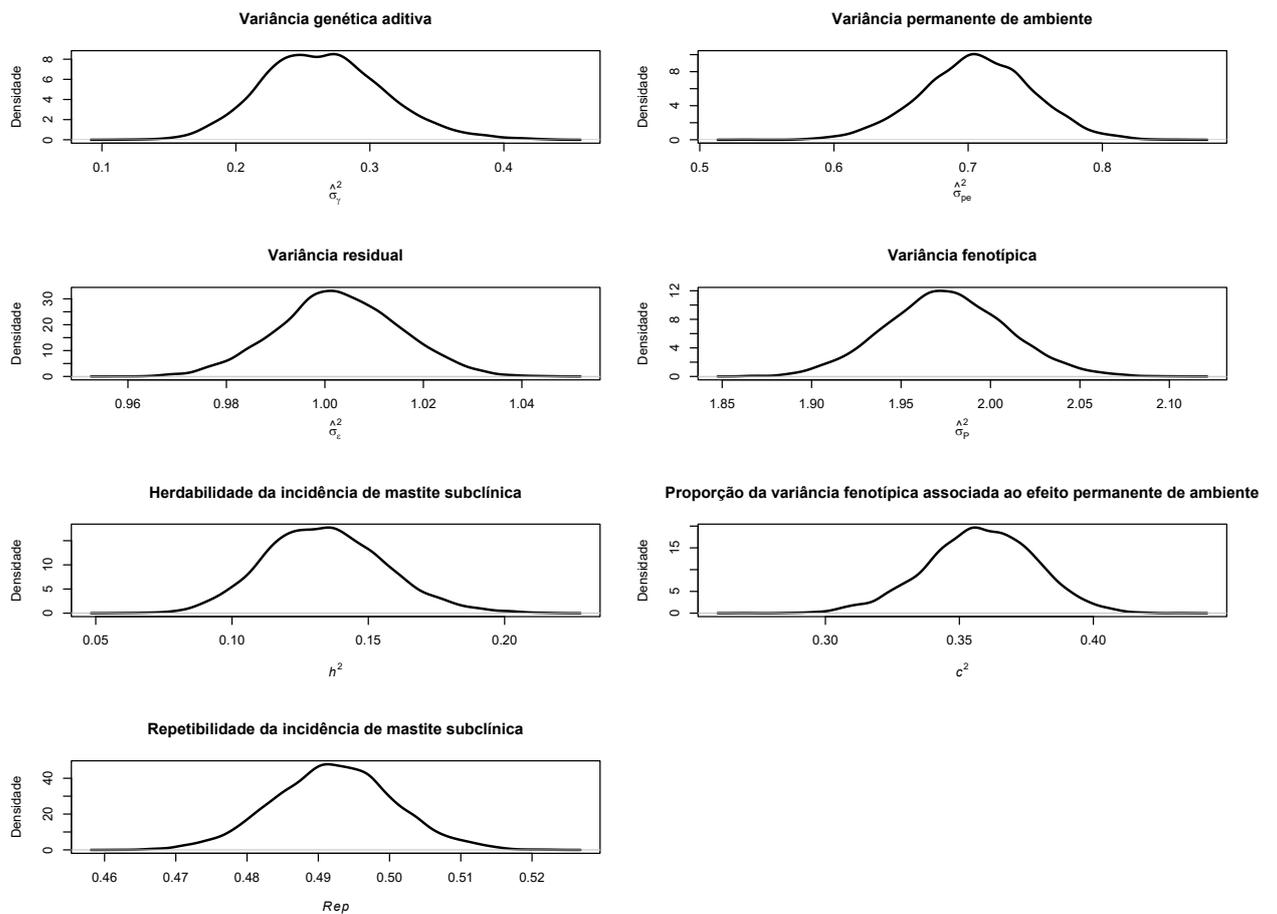


Fonte: Da autora.

$\hat{\sigma}_{\gamma}^2$ -variância genética aditiva; $\hat{\sigma}_{pe}^2$ -variância atribuída aos efeitos de ambiente permanente; $\hat{\sigma}_{\varepsilon}^2$ -variância residual; $\hat{\sigma}_P^2$ -variância fenotípica; h^2 -herdabilidade; c^2 -proporção da variância fenotípica associada ao efeito permanente de ambiente; Rep -Repetibilidade.

APÊNDICE B - Gráficos de densidades das cadeias *a posteriori* dos parâmetros $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_P^2$, h^2 , c^2 e Rep resultantes do modelo de limiar aditivo para as características binárias de incidência de mastite subclínica e precocidade com *priori* não informativa e informativa

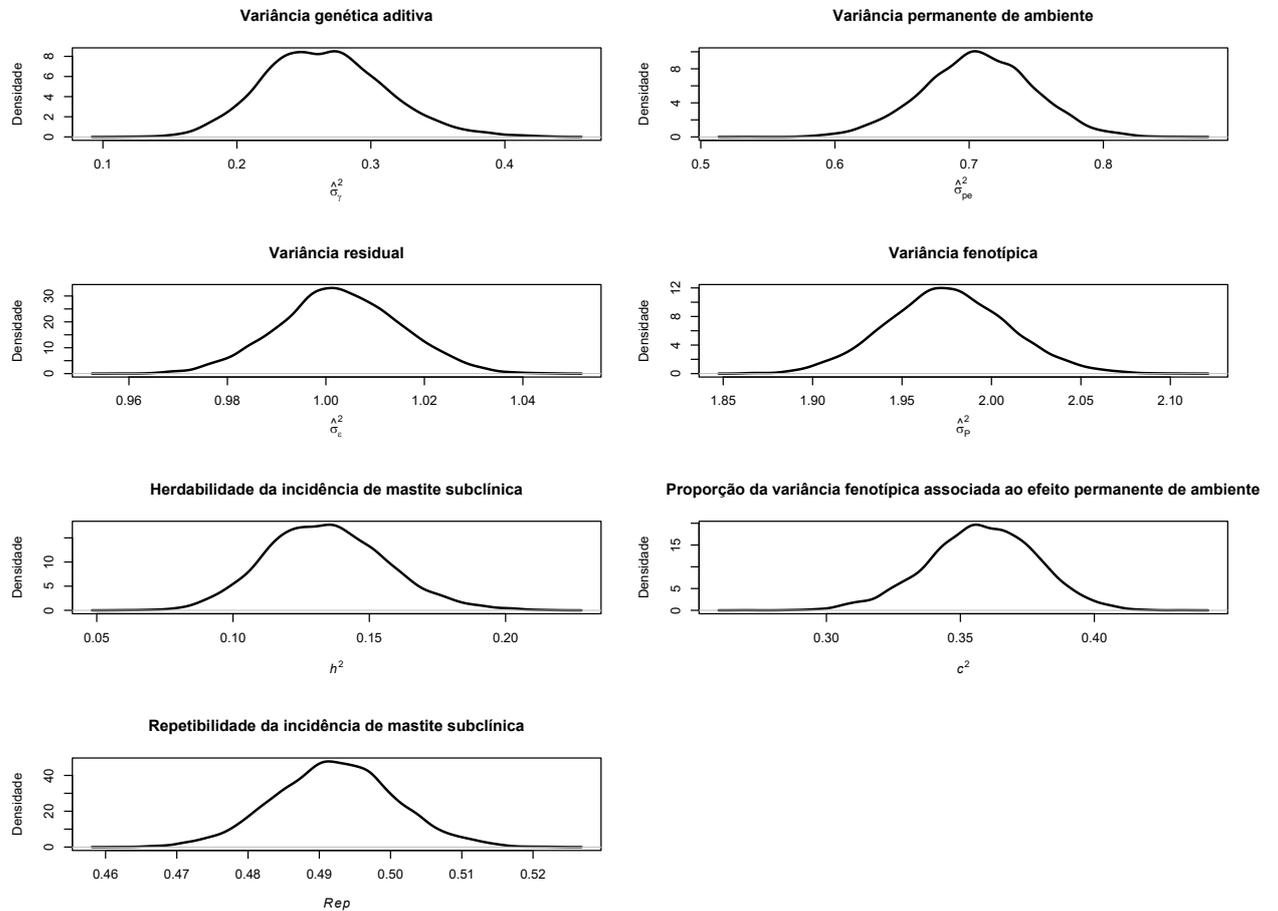
Figura 7 – Gráficos de densidade das cadeias *a posteriori* dos parâmetros estimados $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_P^2$, h^2 , c^2 e Rep resultantes da característica de incidência de mastite subclínica (MSC), utilizando *a priori* não informativa, período 2012 a 2016.



Fonte: Da autora.

$\hat{\sigma}_\gamma^2$ -variância genética aditiva; $\hat{\sigma}_{pe}^2$ -variância atribuída aos efeitos de ambiente permanente; $\hat{\sigma}_\varepsilon^2$ -variância residual; $\hat{\sigma}_P^2$ -variância fenotípica; h^2 -herdabilidade; c^2 -proporção da variância fenotípica associada ao efeito permanente de ambiente; Rep -Repetibilidade.

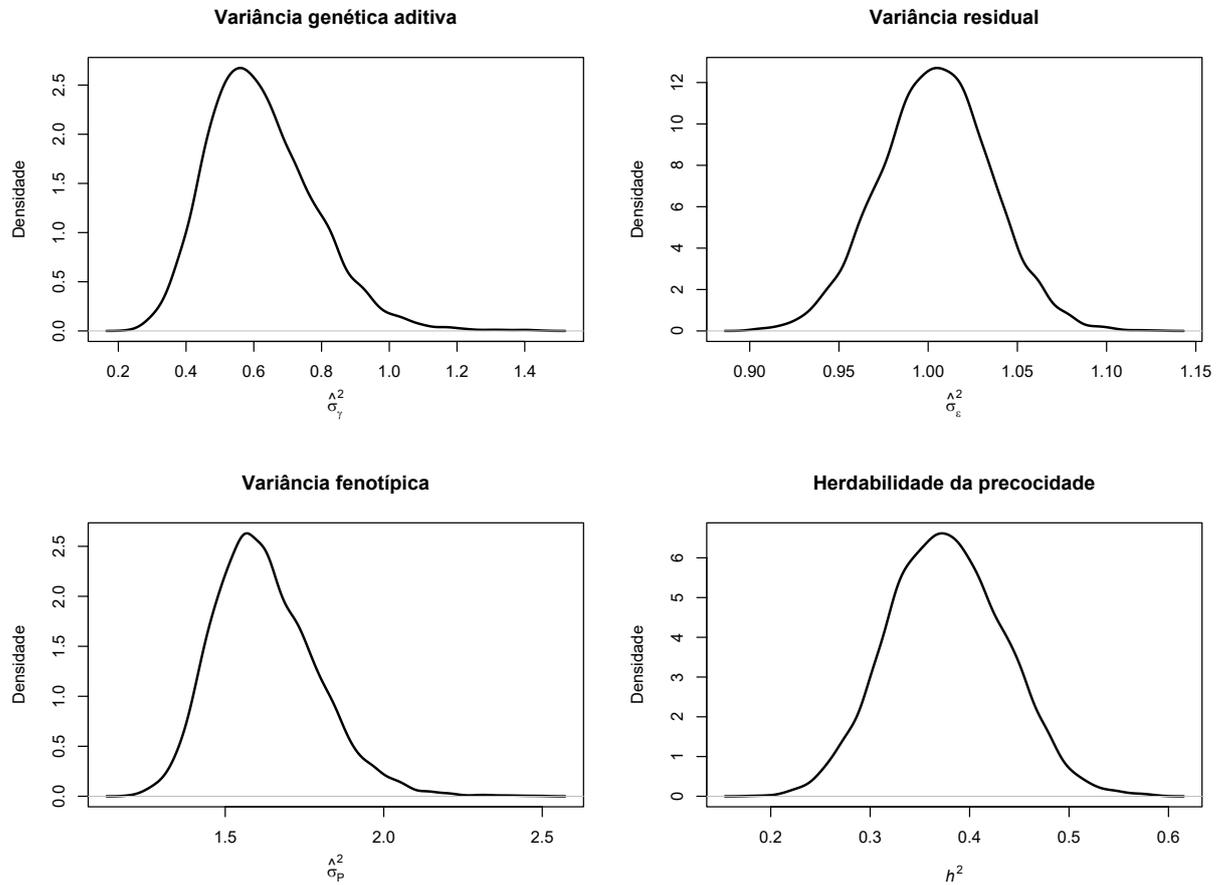
Figura 8 – Gráficos de densidade das cadeias *a posteriori* dos parâmetros estimados $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_P^2$, h^2 , c^2 e Rep resultante da característica de incidência de mastite subclínica (MSC), utilizando *a priori* informativa, período 2012 a 2016.



Fonte: Da autora.

$\hat{\sigma}_\gamma^2$ -variância genética aditiva; $\hat{\sigma}_{pe}^2$ -variância atribuída aos efeitos de ambiente permanente; $\hat{\sigma}_\varepsilon^2$ -variância residual; $\hat{\sigma}_P^2$ -variância fenotípica; h^2 -herdabilidade; c^2 -proporção da variância fenotípica associada ao efeito permanente de ambiente; Rep -Repetibilidade.

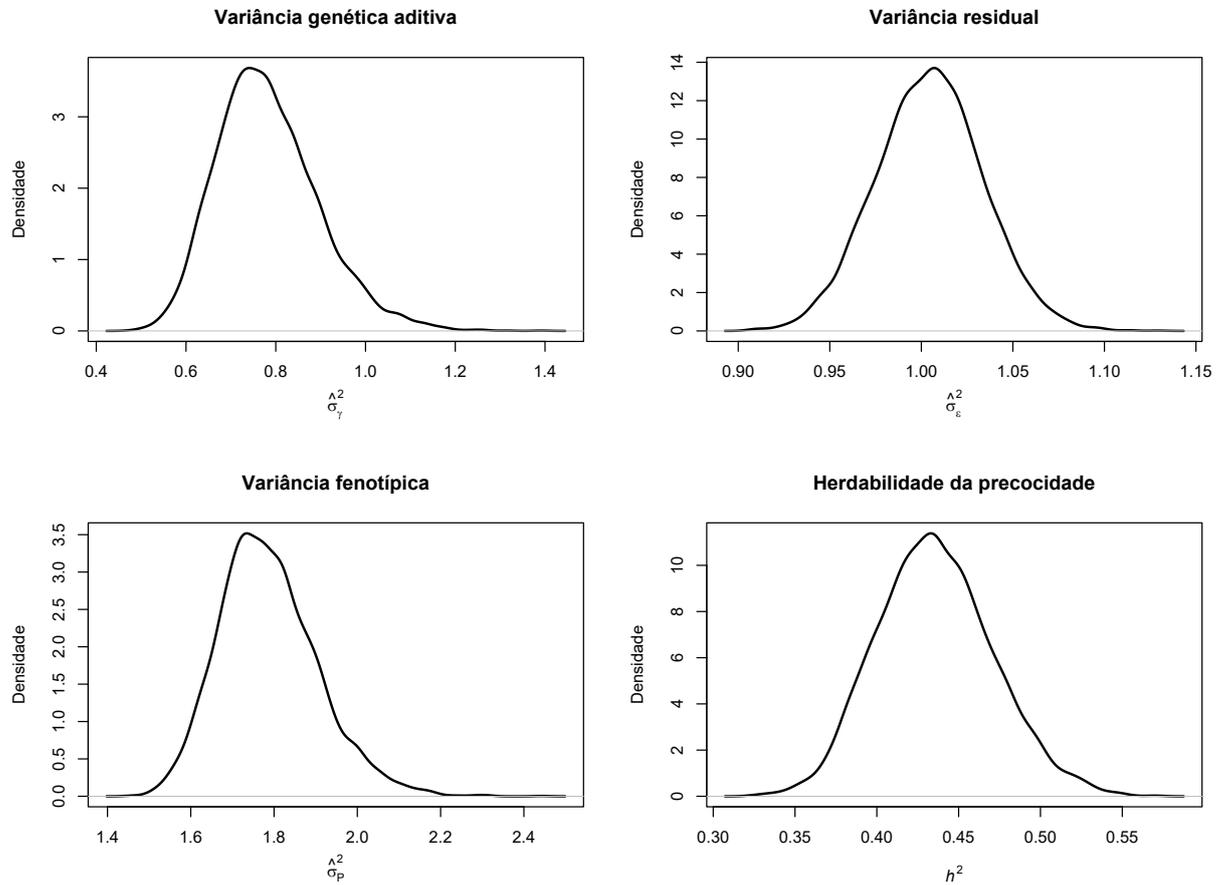
Figura 9 – Gráficos de densidade das cadeias *a posteriori* dos parâmetros estimados $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_P^2$ e h^2 resultantes da característica de precocidade (PREC) com *priori* não informativa, período 2012 a 2016.



Fonte: Da autora.

$\hat{\sigma}_\gamma^2$ -variância genética aditiva; $\hat{\sigma}_{pe}^2$ -variância atribuída aos efeitos de ambiente permanente; $\hat{\sigma}_\varepsilon^2$ -variância residual; $\hat{\sigma}_P^2$ -variância fenotípica; h^2 -herdabilidade.

Figura 10 – Gráficos de densidade das cadeias *a posteriori* dos parâmetros estimados $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_P^2$ e h^2 resultantes da característica de precocidade (PREC) com *priori* informativa, período 2012 a 2016.

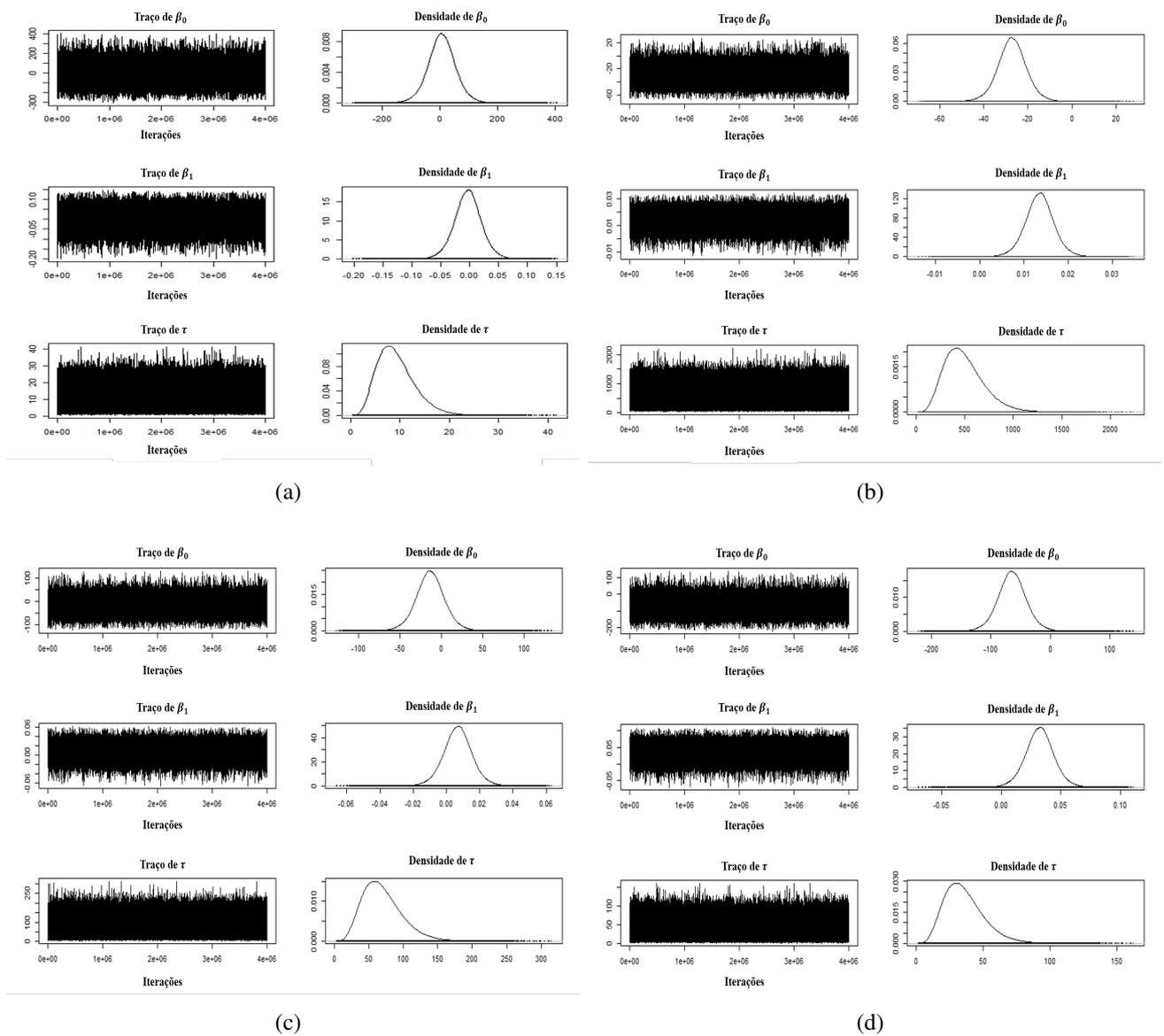


Fonte: Da autora.

$\hat{\sigma}_\gamma^2$ -variância genética aditiva; $\hat{\sigma}_{pe}^2$ -variância atribuída aos efeitos de ambiente permanente; $\hat{\sigma}_\varepsilon^2$ -variância residual; $\hat{\sigma}_P^2$ -variância fenotípica; h^2 -herdabilidade.

APÊNDICE C - Gráficos de traço e densidades das cadeias *a posteriori* dos parâmetros β_0 , β_1 e τ , resultantes do modelo de regressão linear simples para os valores genéticos para as características de produção de leite, teor de caseína, incidência de mastite subclínica e precocidade

Figura 11 – Gráficos do traço e densidade das cadeias *a posteriori* dos parâmetros β_0 , β_1 e τ resultantes do ajuste do modelo linear simples aos dados de valores genéticos para as características de produção de leite (a), teor de caseína (b), incidência de mastite (c) e precocidade (d) período de 2001 a 2014.



Fonte: Da autora.