

**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

**Modelo beta misto bayesiano para descrever a influência de diferentes
cepas do fungo *Colletotrichum truncatum* sobre a severidade da
antracnose na sojicultura**

Antonio Leopoldo Cardoso Sabino

Dissertação apresentada para obtenção do título de Mestre em Ciências. Área de concentração: Estatística e Experimentação Agronômica

**Piracicaba
2022**

Antonio Leopoldo Cardoso Sabino
Bacharel em Estatística

**Modelo beta misto bayesiano para descrever a influência de diferentes
cepas do fungo *Colletotrichum truncatum* sobre a severidade da
antracnose na sojicultura**

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:

Prof. Dr. **SILVIO SANDOVAL ZOCCHI**

Dissertação apresentada para obtenção do título de Mes-
tre em Ciências. Área de concentração: Estatística e
Experimentação Agronômica

Piracicaba
2022

**Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP**

Sabino, Antonio Leopoldo Cardoso

Modelo beta misto bayesiano para descrever a influência de diferentes cepas do fungo *Colletotrichum truncatum* sobre a severidade da antracnose na sojicultura / Antonio Leopoldo Cardoso Sabino. -- versão revisada de acordo com a resolução CoPGr 6018 de 2011. -- Piracicaba, 2022 .

92 p.

Dissertação (Mestrado) -- USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. Soja 2. Patógeno 3. *MCMC* 4. Taxas . I. Título.

AGRADECIMENTOS

Agradeço primeiramente a Deus, pelo apoio nos momentos difíceis, por abrir espaços frente às dificuldades e por ser meu guia, desde o princípio.

Aos meus pais, Maria Edileuza Sabino Cardoso e José Cardoso, por sempre me motivarem a seguir em frente e por todo sacrifício que fizeram para que um dia eu chegasse onde estou.

Ao Prof. Dr. Silvio Sandoval Zocchi por todo apoio, confiança, paciência, orientação e amizade ao longo deste trabalho.

Ao Prof. Dr. Cristian Marcelo Villegas Lobos, pela grande ajuda com a programação computacional e teoria sobre os modelos bayesianos.

Ao Prof. Dr. Idemauro Antonio Rodrigues de Lara pelo apoio, confiança e motivação durante este caminho.

A Dra. Flávia Rogério pela cessão do conjunto de dados usados neste trabalho, por todo apoio e disponibilidade em tirar dúvidas e realizar discussões técnicas dos resultados obtidos, que foram de extrema importância para a tomada de decisões durante as análises.

Aos demais professores do Programa de Pós-Graduação em Estatística e Experimentação Agronômica (PPGEEA) pelos valiosos conhecimentos passados. Minha mais profunda admiração por todos.

À Escola Superior de Agricultura “Luiz de Queiroz” e a Universidade de São Paulo como um todo, pela excelente estrutura acadêmica, respeito e zelo pelos seus alunos e suporte com materiais de qualidade que ajudaram na confecção deste trabalho.

Aos funcionários do Departamento de Ciências Exatas (LCE/ESALQ) por todo auxílio, atenção, disponibilidade e profissionalismo.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), pela disponibilidade de bolsa de estudos, que permitiu a condução e execução deste trabalho.

Muito obrigado a todos que, de forma direta ou indiretamente, ajudaram ao longo dessa jornada.

SUMÁRIO

Resumo	6
Abstract	7
Lista de Figuras	8
Lista de Tabelas	9
1 Introdução	11
2 Revisão Bibliográfica	13
3 Material e Métodos	29
3.1 Material	29
3.2 Métodos	31
3.2.1 Modelo para dados de uma amostra aleatória simples de Y	31
3.2.2 Modelo para dados provenientes de experimentos inteiramente ao acaso com um fator e ϕ não variável	33
3.2.3 Modelo para dados provenientes de experimentos inteiramente ao acaso com um fator e ϕ variável	36
3.2.4 Modelo misto para dados provenientes de um experimento com dois fatores aninhados, ϕ variável e mais de um componente de variância	38
3.2.5 Modelo misto para dados provenientes de um experimento com dois fatores aninhados, ϕ variável e um componente de variância	41
3.2.6 Função de verossimilhança, distribuições a priori, a posteriori, marginais e condicionais completas para o modelo considerando dados de uma amostra aleatória simples de Y	42
3.2.7 Função de verossimilhança, distribuições a priori, a posteriori, marginais e condicionais completas para o modelo considerando dados provenientes de experimentos inteiramente ao acaso com um fator e ϕ não variável	43
3.2.8 Função de verossimilhança, distribuições a priori, a posteriori, marginais e condicionais completas para o modelo considerando dados provenientes de experimentos inteiramente ao acaso com um fator e ϕ variável	45
3.2.9 Função de verossimilhança, distribuições a priori, a posteriori, marginais e condicionais completas para o modelo misto considerando dados provenientes de um experimento com dois fatores aninhados, com ϕ variável e mais de um componente de variância	46

3.2.10	Função de verossimilhança, distribuições a priori, a posteriori, marginais e condicionais completas para o modelo misto considerando dados provenientes de um experimento com dois fatores aninhados, com ϕ variável e um componente de variância	48
3.2.11	Convergência e ajuste dos modelos	50
4	Resultados e Discussão	51
5	Conclusões	71
5.1	Trabalhos futuros	71
	Referências	73
	Apêndices	79

RESUMO

Modelo beta misto bayesiano para descrever a influência de diferentes cepas do fungo *Colletotrichum truncatum* sobre a severidade da antracnose na sojicultura

A cultura da soja é muito importante na economia mundial, pois sua matéria prima é utilizada para produção de diversos produtos. As plantas de soja, entretanto, são suscetíveis a diversas doenças causadas por fungos, como a antracnose, que pode chegar a comprometer a safra inteira. Essa doença é causada por cepas do fungo *Colletotrichum truncatum*, pertencentes a grupos genéticos distintos, que podem levar a índices de severidade diversos. Nas análises estatísticas usuais de dados de severidade, entretanto, frequentemente considera-se, equivocadamente, que seguem uma distribuição normal ou são realizadas, previamente, transformações dos dados, como logit, probit ou complemento log-log, dentre outras. Como a severidade é uma variável contínua entre 0 e 100%, (ou entre 0 e 1), a distribuição beta pode ser mais apropriada e metodologias como a proposta por Cribari-Neto e Zeileis (2010) podem ser consideradas. Em experimentos de comparação de grupos genéticos de cepas de antracnose quando ao índice de severidade podemos considerar o efeito de cepas como fixo ou aleatório dependendo do interesse do pesquisador. Apresentamos, aqui, uma abordagem bayesiana para a análise de dados provenientes de experimentos inteiramente ao acaso do tipo, com dois fatores, grupo genético e cepas. A metodologia foi implementada utilizando-se a interface RStan (Stan Development Team, 2020) e ilustrada por meio de um conjunto de dados reais e inéditos. A análise dos mesmos revela que diferentes conclusões podem ser obtidas ao considerarmos efeitos de cepa fixos ou aleatórios.

Palavras-chave: Soja, Patógeno, *MCMC*, Texas.

ABSTRACT

Bayesian mixed beta model to describe the influence of different strains of the fungus *Colletotrichum truncatum* on the severity of anthracnose in soybeans

Soybean farming is very important in the world economy, as soy is used for the production of various products. Soybean plants, however, are susceptible to several diseases caused by fungi, such as anthracnose, which can compromise the entire crop. This disease is caused by strains of the fungus *Colletotrichum truncatum*, belonging to different genetic groups, which can lead to different levels of severity. In the usual statistical analysis of severity data, however, it is often mistakenly considered that they follow a normal distribution or that data transformations are carried out in advance, such as logit, probit or log-log complement, among others. As the severity is a continuous variable between 0 and 100%, (or between 0 and 1), the beta distribution can be more appropriate and methodologies like the one proposed by Cribari-Neto e Zeileis (2010) can be considered. In experiments comparing genetic groups of anthracnose strains regarding the severity index, we can consider the effect of strains as fixed or random depending on the researcher's interest. We present here a Bayesian approach to the analysis of data from entirely randomized experiments of this type, with two factors, genetic group and strain. The methodology was then implemented using the RStan interface (Stan Development Team, 2020) and illustrated using a set of unpublished real data. Their analysis reveals that different conclusions can be reached when considering fixed or random strain effects.

Keywords: Soybean, Pathogen, *MCMC*, Rates.

LISTA DE FIGURAS

Figura 1.	Gráficos das funções de densidade de probabilidade de uma variável $Y \sim Beta(p, q)$, para diferentes valores dos parâmetros p e q (à esquerda) e de uma variável $Y \sim Beta(\mu, \phi)$ para diferentes valores dos parâmetros μ e ϕ (à direita).	15
Figura 2.	Plantas da cultura da soja, inoculadas com a cepa MT5-12, após 30 dias da semeadura.	30
Figura 3.	Imagens que foram utilizadas como argumentos na função symptomatic_area	30
Figura 4.	Gráfico de dispersão da severidade na escala da função de ligação complemento log-log em função dos grupos de cepas.	51
Figura 5.	Gráfico das densidade estimada das simulações dos valores da distribuição preditiva a posteriori (y_{rep}) para o A) modelo 1, B) modelo 2 e y valores observados	53
Figura 6.	Gráfico dos intervalos de 95% das simulações dos valores da distribuição preditiva a posteriori (y_{rep}) para o A) modelo 1, B) modelo 2 e y valores observados, os valores estão na escala complemento log-log.	54
Figura 7.	Intervalos de 95% de credibilidade com a mediana dos parâmetros $\beta_1, \dots, \beta_{14}$ do modelo 2	56
Figura 8.	Intervalos de 95% de credibilidade com a mediana dos parâmetros $\lambda_1, \dots, \lambda_{14}$ do modelo 2	56
Figura 9.	Intervalos de 95% de credibilidade com a mediana dos contrastes do modelo 2	59
Figura 10.	Gráfico das densidade estimadas das simulações dos valores da distribuição preditiva a posteriori do modelo M5.	64
Figura 11.	Intervalos de 95% de credibilidade com a mediana dos parâmetros β_1, \dots, β_4 do modelo M5	66
Figura 12.	Intervalos de 95% de credibilidade para os contrastes do modelo M5	68

LISTA DE TABELAS

Tabela 1.	Funções de ligação, $g(\cdot)$, e respectivas funções inversas, $g^{-1}(\cdot)$	16
Tabela 2.	Proporções dos pixels das imagens das plantas de soja com coloração não verde (severidade) segundo grupo genético, o tratamento aplicado, o vaso e a média.	31
Tabela 3.	Observações da variável Y num experimento inteiramente ao acaso com I tratamentos e J repetições	33
Tabela 4.	Observações da variável Y num experimento inteiramente ao acaso com I tratamentos e $J_i, i = 1, \dots, I$ repetições para o i -ésimo tratamento.	35
Tabela 5.	Observações da variável Y num experimento aninhado com I grupos, $J_i, i = 1, \dots, I$, cepas para o i -ésimo grupo e K_{ij} repetições para a j -ésima cepa do i -ésimo grupo, $j = 1, \dots, J_i$	38
Tabela 6.	Modelos e distribuições a priori dos parâmetros dos modelos considerando dados provenientes de experimentos inteiramente ao acaso com um fator	52
Tabela 7.	Critério de informação amplamente aplicável (WAIC), Critério de informação leave-one-out (LOOIC) e o p-valor de verificação preditiva a posterior (p_d). Menores valores dos critérios estão destacados em vermelho.	53
Tabela 8.	Estimativas a posteriori dos parâmetros do modelo 2. médias, desvios padrões, medianas e intervalos de 95% de credibilidade (IC 95%).	55
Tabela 9.	Estimativas a posteriori dos valores esperados do modelo 2	57
Tabela 10.	Resultados dos contrastes dos parâmetros do modelo 2, mediana e intervalo de 95% de credibilidade (IC 95%).	58
Tabela 11.	Distribuições a priori dos parâmetros dos modelos considerando dados provenientes de um experimento com dois fatores aninhados.	63
Tabela 12.	Critério de informação amplamente aplicável (WAIC), Critério de informação leave-one-out (LOOIC) e o p-valor de verificação preditiva a posterior (p_d).	63
Tabela 13.	Estimativas a posteriori dos parâmetros do modelo M5 e respectivos intervalos de 95% de credibilidade (IC 95%).	65
Tabela 14.	Resultados dos contrastes do modelo M5. Média, mediana e intervalo de 95% de credibilidade (IC 95%).	68

1 INTRODUÇÃO

A cultura da soja (*Glycine max* (L.) Merr) tem grande importância socioeconômica em todo mundo, possuindo grandes áreas de cultivo, sendo seus grãos utilizados nas mais diversas áreas tais como: agropecuária, alimentícia, farmacêutica, energética entre outras. Os grãos de soja são utilizados também para produção de leite de soja (extrato proteico de soja), farelo de soja, óleo de soja, missô (pasta de soja), rações para animais, biocombustível, medicamentos entre outros produtos (FIESP, 2022).

O Brasil é o maior produtor mundial com área cultivada de 127,842 milhões de hectares e produção de 135,409 milhões de toneladas, que corresponde a cerca de 37,31% de toda safra mundial. Neste país, o estado do Mato Grosso ocupa o primeiro lugar, com quantidade produzida de 35,947 milhões de toneladas, e o estado do Paraná, por sua vez, ocupa o segundo lugar, produzindo 19,872 milhões de toneladas, seguido pelos estados do Rio Grande do Sul e Goiás com produções de 20,164 e 13,720 milhões de toneladas, respectivamente (EMBRAPA, 2022).

Esta cultura, no entanto, pode ter seu rendimento reduzido significativamente por várias doenças, sendo a antracnose uma das mais impactantes, pois os danos causados na área foliar, hastes, vagens, necrose dos pecíolos podem causar a morte da planta. Em certos casos, alta intensidade da doença pode comprometer e devastar 100% da produtividade (SHARMA; GUPTA; RAMTEKE, 2011). Esta doença é causada por um patógeno que provém de várias espécies do fungo do gênero *Colletotrichum*, tais como: *C. truncatum*, *C. destructivum*, *C. coccodes*, *C. chlorophyti*, *C. gloeosporioides*, *C. incanum*, *C. plurivorum*, *C. sojiae*, *C. musicola* e *C. brevisporum*.

A espécie *Colletotrichum truncatum* é considerada a mais relevante (BOUFLEUR et al., 2021). Por apresentar adaptações a diversas condições ambientais, esta espécie apresenta, ainda, grande variabilidade genética, em cepas diferentes. No Brasil, inúmeras cepas têm sido principalmente encontradas em Mato Grosso e Goiás (FLÁVIA et al., 2019). Este fato também tem sido relatado em outros países como China e Índia (DIAO et al., 2015; SHARMA; PINNAKA; SHENOY, 2014). Estudos recentes encontraram semelhanças genéticas entre cepas, a partir dessas semelhanças grupos genéticos diferentes foram formados (FLÁVIA et al., 2019).

Como as cepas do patógeno *Colletotrichum truncatum* infectam o limbo foliar com grande severidade em diferentes culturas. Logo essas cepas causam níveis de severidade diferentes e alguns trabalhos recentes os pesquisadores utilizam modelos, em que se assume que os dados de severidade segue o comportamento de uma distribuição normal, como exemplo na cultura: do feijão-mungo (PANDEY et al., 2021), da soja (SILVA JÚNIOR et al., 2021)

e do espinafre (LIU et al., 2021).

Como foi mostrado anteriormente os trabalhos relacionados a proporção de severidade causada pela antracnose, geralmente os pesquisadores optam por fazer uma análise de variância, considerando que a variável severidade assume um comportamento de uma distribuição normal. Essa suposição de normalidade pode as vezes não ser o mais correto, pois o intervalo de variação da severidade é limitado ao $(0,1)$ e pelo motivo disso, pode ocorrer um viés nas estimativas obtidas dada a suposição de normalidade.

Um trabalho recente na área da fitopatologia faz uma abordagem diferente, considerando que a variável resposta tem um comportamento de uma parametrização da distribuição beta proposta por Ferrari e Cribari-Neto (2004), pois o suporte da distribuição é o mesmo da severidade. Esse trabalho utiliza o modelo de regressão beta para estimar a proporção de ascósporos na cultura do pequi (MARTÍNEZ-MINAYA et al., 2021).

Tal abordagem será utilizada no presente trabalho, entretanto sobre o comportamento nas áreas foliares da cultura da soja e com outros tipos de modelos. Desse modo, os objetivos do presente trabalho foram, a elaboração e a validação de modelos beta, utilizando a distribuição beta proposta por Ferrari e Cribari-Neto (2004) e com esses modelos avaliar os dados de severidade nas áreas foliares da cultura da soja em função dos grupos genéticos das cepas do fungo *Colletotrichum truncatum* e testar a hipótese se esses grupos provocam níveis de severidade diferentes.

2 REVISÃO BIBLIOGRÁFICA

Por definição a família de distribuições beta são compostas de todas as distribuições com função densidade de probabilidade definida por (JOHNSON; KOTZ; BALAKRISHNAN, 1995):

$$f(x; a, b, p, q) = \frac{1}{B(p, q)} \frac{(x - a)^{p-1} (b - x)^{q-1}}{(b - a)^{p+q-1}} I_{(a,b)}(x), \quad (1)$$

em que $a, b, q, p > 0$, $I_{(a,b)}(x)$ é uma função indicadora, que assume o valor 1 para $a < x < b$ e 0, caso contrário.

Fazendo a transformação de variável

$$Y = \frac{X - a}{b - a},$$

a equação (1) passa a ter função densidade de probabilidade, que é definida por:

$$f(y; p, q) = \frac{1}{B(p, q)} y^{p-1} (1 - y)^{q-1} I_{(0,1)}(y), \quad (2)$$

em que $I_{(0,1)}(y)$ é uma função indicadora, que assume o valor 1 para $0 < y < 1$ e 0, caso contrário, e $B(p, q)$ é a função beta, dada por:

$$B(p, q) = \int_0^1 t^{p-1} (1 - t)^{q-1} dt = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p + q)},$$

em que $\Gamma(\cdot)$ é a função gama, definida como:

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} \exp\{-x\} dx$$

para $\alpha > 0$. Neste caso, temos que a esperança e variância de Y são, respectivamente:

$$E(Y) = \frac{p}{p + q} \quad (3)$$

e

$$Var(Y) = \frac{pq}{(p + q)^2(p + q + 1)} = \frac{E(Y)(1 - E(Y))}{p + q + 1}. \quad (4)$$

Conforme ilustra a Figura 1 de acordo com os valores de p e q , a distribuição beta pode assumir diferentes formatos como U, \cap , J, J invertido, dentre outros, e quando $p = q = 1$, $f(y) = I_{(0,1)}(y)$, ou seja, Y tem distribuição uniforme no intervalo $(0, 1)$ (JOHNSON; KOTZ; BALAKRISHNAN, 1995).

De acordo com a pesquisa histórica realizada por Johnson, Kotz e Balakrishnan (1995) a distribuição beta foi utilizada em vários tipos de modelagem, como por exemplo em

meados da década de 80 Janardan e Padmanabhan (1986) modelou variáveis hidrológicas, Golenko-Ginzburg (1988) fez uma análise da técnica de avaliação e revisão de programa (*Program Evaluation and Review Technique - PERT*) assumindo que o tempo dessa técnica segue uma distribuição beta, dados de insolação para uso em sistemas fotovoltaicos foram analisados com equação (1) (RAHMAN; KHALLAT; SALAMEH, 1988), Harrop-Williams (1989) modela a proporção de porosidade com a equação (2).

Na década de 90 McNally (1990) utilizou a distribuição beta para modelar a taxa de perda de embriões, Agrawal e Yang (1991) modelaram as proporções em misturas de gases com a mesma distribuição, Graham e Hollands (1990) estimaram a taxa de absorbância da luz solar pela atmosfera, Maffett e Wackerman (1991), Karavias e Myers (1992) estimaram a taxa de absorção de gás.

De modo que a esperança de Y seja igual a um parâmetro μ , em que $0 < \mu < 1$ e a variância de Y seja inversamente proporcional a um parâmetro ϕ , em que $\phi > 0$, utilizamos, a partir daqui, a parametrização apresentada por Ferrari e Cribari-Neto (2004) em que $p = \mu\phi$ e $q = (1 - \mu)\phi$ e escrevemos $Y \sim Beta(\mu, \phi)$. Assim, substituindo $p = \mu\phi$ e $q = (1 - \mu)\phi$ em (2), (3) e (4), teremos que a função densidade de probabilidade de Y , a esperança e a variância de Y serão dadas, respectivamente, por

$$f(y; \mu, \phi) = \frac{1}{B(\mu\phi, (1 - \mu)\phi)} y^{\mu\phi-1} (1 - y)^{(1-\mu)\phi-1} I_{(0,1)}(y), \quad (5)$$

$$E(Y) = \frac{\mu\phi}{\mu\phi + (1 - \mu)\phi} = \frac{\mu\phi}{\phi(\mu + 1 - \mu)} = \mu \quad (6)$$

e

$$Var(Y) = \frac{\mu(1 - \mu)}{\mu\phi + (1 - \mu)\phi + 1} = \frac{\mu(1 - \mu)}{\mu\phi + \phi - \mu\phi + 1} = \frac{\mu(1 - \mu)}{\phi + 1}. \quad (7)$$

Podemos notar, a partir da equação (7) e dos gráficos da Figura 1 (à direita) que com essa parametrização, supondo que μ é um valor fixo, quanto maior for o valor de ϕ , menor será a variância e portanto, maior será a precisão, que, em estatística, é definida como o recíproco da variância.

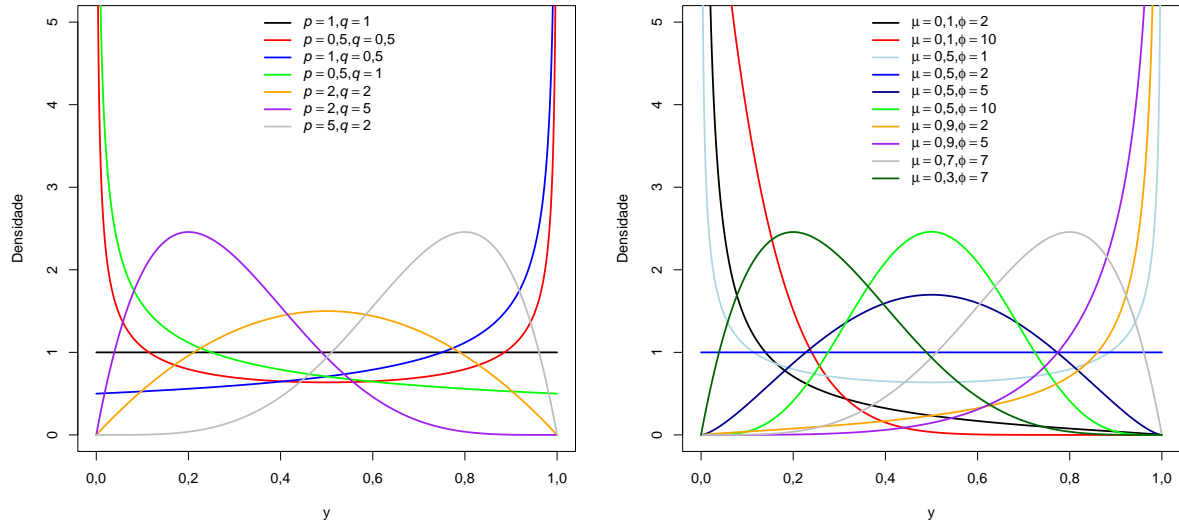


Figura 1. Gráficos das funções de densidade de probabilidade de uma variável $Y \sim \text{Beta}(p, q)$, para diferentes valores dos parâmetros p e q (à esquerda) e de uma variável $Y \sim \text{Beta}(\mu, \phi)$ para diferentes valores dos parâmetros μ e ϕ (à direita).

A função (5) possibilitou a estimação do parâmetro de escala com uma estrutura de regressão, com o auxílio de um preditor linear η e uma função de ligação $g(\cdot)$, possibilitando a criação de modelos de regressão para a distribuição beta (FERRARI; CRIBARI-NETO, 2004).

Seja $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$ uma amostra aleatória, $u = 1, \dots, n$, em que cada Y_u segue uma densidade (5) e $\mathbf{y} = (y_1, y_2, \dots, y_n)$ um conjunto de observações de \mathbf{Y} , com valor esperado μ_u e parâmetro de precisão ϕ desconhecido. De acordo com Ferrari e Cribari-Neto (2004), o modelo de regressão beta será dado por:

$$\begin{aligned} Y_u &\sim \text{Beta}(\mu_u, \phi), \quad u = 1, \dots, n, \\ \mu_u &= g^{-1}(\eta_u), \\ \eta_u &= \mathbf{x}_u^T \boldsymbol{\beta} \end{aligned} \quad (8)$$

em que: $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ é um vetor de parâmetros desconhecidos, com comprimento p ; $\mathbf{x}_u = (x_{u1}, \dots, x_{up})^T$ é um vetor de regressores, com comprimento p e $g(\cdot)$ é uma função monótona e diferenciável, com domínio $(0,1)$. Geralmente a função de ligação utilizada é a logito, que é definida como:

$$g(\mu_u) = \log \left(\frac{\mu_u}{1 - \mu_u} \right),$$

outras funções de ligação podem ser vista na Tabela 1.

Tabela 1. Funções de ligação, $g(\cdot)$, e respectivas funções inversas, $g^{-1}(\cdot)$.

Função de ligação	$g(\mu)$	$g^{-1}(\eta)$
logito	$\log\{\mu/(1 - \mu)\}$	$\exp(\eta)/\{1 + \exp(\eta)\}$
log	$\log(\mu)$	$\exp(\eta)$
probit	$\Phi^{-1}(\mu)$	$\Phi(\eta)$
complemento log-log	$\log\{-\log(1 - \mu)\}$	$1 - \exp\{-\exp(\eta)\}$
log-log	$-\log\{-\log(\mu)\}$	$\exp\{-\exp(-\eta)\}$
Cauchy	$\text{tg}\{\pi(\mu - 0,5)\}$	$\{\text{arctg}(\eta)/\pi\} + 0,5$

Nota: $\Phi^{-1}(\cdot)$ é a inversa da função de distribuição acumulada da normal padrão.

Fonte: Nelder e Wedderburn (1972).

O modelo (8) foi utilizado como, por exemplo, em Ferrari e Cribari-Neto (2004), modelando a proporção da renda gasta em alimentação em função do nível de renda e o número de pessoas no domicílio, Silva, Migon e Correia (2011) analisa as porcentagens mensais de desemprego em função das pequenas e médias empresas, que foram analisadas no período de março de 2002 a dezembro de 2009, Schmidt, Moraes e Migon (2015) estuda a proporção média do desempenho de alunos durante as edições da OBMEP de 2006 a 2013, em função das escolas, Xiang et al. (2021) estima a proporção obtida em testes de concentração de memória de informação abençoada (*Blessed Information-Memory-Concentration - BIMC*) em função das covariáveis idade, sexo, escolaridade e genótipos APOE.

Uma extensão do modelo (8) foi introduzido inicialmente na literatura por Smithson e Verkuilen (2006). Este novo modelo, além de estimar a média, permite a adição de regressoras na estimação do parâmetro ϕ , dessa forma o modelo pode assumir que determinados conjuntos de observações tem variâncias diferentes. Desse modo, com esse modelo pode-se evitar de assumir equivocadamente que o parâmetro ϕ único para todas as observações (OSPINA, 2007).

Segundo Smithson e Verkuilen (2006) o modelo de regressão beta com dispersão variável é definido da seguinte forma: seja Y_u uma variável com densidade (5), com valor

esperado μ_u , precisão ϕ_u , o modelo é dado por

$$\begin{aligned}
 Y_u &\stackrel{ind.}{\sim} \text{Beta}(\mu_u, \phi_u), \quad u = 1, \dots, n, \\
 \mu_u &= g_1^{-1}(\eta_{1u}), \\
 \eta_{1u} &= \mathbf{x}_u^T \boldsymbol{\beta}, \\
 \phi_u &= g_2^{-1}(\eta_{2u}), \\
 \eta_{2u} &= \mathbf{w}_u^T \boldsymbol{\lambda}.
 \end{aligned} \tag{9}$$

em que: $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ e $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_h)^T$ são vetores de parâmetros desconhecidos, com respectivos comprimento p e h ; $\mathbf{x}_u = (x_{u1}, \dots, x_{up})^T$ e $\mathbf{w}_u = (w_{u1}, \dots, w_{uh})^T$ são vetores de regressores, com respectivos comprimento p e h ; $g_1(\cdot)$ e $g_2(\cdot)$ são funções de ligação; η_{1u} e η_{2u} são preditores lineares associados, respectivamente, ao valor esperado e à precisão de Y_u .

Com a possibilidade da inclusão de uma estrutura de regressão para o parâmetro de precisão, vários trabalhos foram feitos, por exemplo: o modelo (9) foi utilizado para modelar três exemplos, em que foi assumido que as observações possuem a mesma estrutura de regressão do valor esperado da distribuição (SMITHSON; VERKUILEN, 2006).

O modelo de regressão beta com efeito fixo mostrado anteriormente é bastante utilizado, mas para esse tipo de modelo existe o pressuposto que as observações são independentes, ou seja, não possui uma dependência entre as mesmas. Às vezes esse pressuposto não é satisfeito, nesse caso os modelos de regressão beta com efeito fixo não são recomendados (ZERBETO, 2018).

A existência de correlação entre as observações é comum, de acordo com Zerbeto (2018) dados relacionados a educação, saúde e ciências sociais, apresentam relação entre as observações, devido a isso o modelo (8) foi estendido, com adição do efeito aleatório, criando assim a classe de modelos de regressão beta misto. Segundo Bonat, Jr e Zeviani (2015), seja Y_u uma variável com distribuição beta, com valor esperado μ_u e precisão ϕ o modelo é dado por

$$\begin{aligned}
 Y_u &\stackrel{ind.}{\sim} \text{Beta}(\mu_u, \phi), \quad u = 1, \dots, n, \\
 \mu_u &= g^{-1}(\eta_u), \\
 \eta_u &= \mathbf{x}_u^T \boldsymbol{\beta} + \mathbf{z}_u^T \mathbf{b},
 \end{aligned} \tag{10}$$

em que: μ_u é o valor esperado de Y_u ; $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ é um vetor de parâmetros desconhecidos, com comprimento p ; $\mathbf{x}_u = (x_{u1}, \dots, x_{up})^T$ e $\mathbf{z}_u = (z_{u1}, \dots, z_{us})^T$ são vetores de regressores, com respectivos comprimentos p e s ; \mathbf{b} é o vetor de efeitos aleatórios, em que cada componente do vetor segue uma distribuição normal com média 0 e variância σ_b^2 , com comprimento s e $g(\cdot)$ é uma função de ligação .

No modelo (10) foi adicionado componentes aleatórios, que vem sendo utilizado para analisar dados que são correlacionados, em que a variável resposta é uma taxa ou proporção, onde se assume uma estrutura de correlação na matriz de variâncias e covariâncias do efeito aleatório. Para elucidar a aplicação do modelo podemos citar alguns trabalhos tais como Verkuilen e Smithson (2012) estuda a taxa de habilidade de 55 alunos de graduação para realizar 10 atividades, Manco (2013) modela a porcentagem de volume de gás ainda presente nos olhos em função do tempo após injetado e concentração, ele faz outra análise em que estima a taxa de abandono escolar em escolas públicas em função de quatro covariáveis.

De modo a obter estimativas dos parâmetros dos modelos (8), (9) e (10), que serão estimados sob uma abordagem bayesiana, na estimação dos parâmetros via inferência bayesiana é assumido que eles são variáveis aleatórias e as informações sobre os mesmos podem ser modeladas por funções densidade de probabilidade a posteriori (GELMAN et al., 2013; LESAFFRE; LAWSON, 2012).

Utilizando o teorema de bayes é possível obter a função densidade a posteriori, a partir da função de densidade a priori e da função de verossimilhança. De modo, que $\boldsymbol{\theta}$ é o vetor de parâmetros, com comprimento p , $f(\boldsymbol{\theta})$ é a função densidade de probabilidade a priori de $\boldsymbol{\theta}$ e $f(\mathbf{y}|\boldsymbol{\theta})$ é a distribuição dos dados observados condicionado aos parâmetros, chamada também de função de verossimilhança, no caso quando vista como função de $\boldsymbol{\theta}$, as vezes escrita como $L(\boldsymbol{\theta}|\mathbf{y}) = f(\mathbf{y}|\boldsymbol{\theta})$ (BOX; TIAO, 2011). No nosso caso usaremos a notação de $f(\mathbf{y}|\boldsymbol{\theta})$ sendo a função de verossimilhança, então função densidade de probabilidade a posteriori é dada por:

$$f(\boldsymbol{\theta}|\mathbf{y}) = \frac{f(\mathbf{y}|\boldsymbol{\theta})f(\boldsymbol{\theta})}{\int_{\Theta} f(\mathbf{y}|\boldsymbol{\theta})f(\boldsymbol{\theta}) d\boldsymbol{\theta}}, \quad \boldsymbol{\theta} \in \Theta. \quad (11)$$

A integral múltipla que se encontra no denominador da equação (11), segundo Gelman et al. (2013) é denominada constante normalizadora da função densidade de probabilidade a posteriori. Box e Tiao (2011) definem proporcionalmente a função densidade de probabilidade a posteriori, que resulta na equação dada por:

$$f(\boldsymbol{\theta}|\mathbf{y}) \propto f(\mathbf{y}|\boldsymbol{\theta})f(\boldsymbol{\theta}), \quad (12)$$

em que $f(\mathbf{y}|\boldsymbol{\theta})$ é a função de verossimilhança e $f(\boldsymbol{\theta})$ é a distribuição a priori.

Neste caso, a função de verossimilhança para o modelo correspondente (8) será

$$f(\mathbf{y}|\boldsymbol{\theta}) = \prod_{u=1}^n f(y_u; \boldsymbol{\theta}), \quad (13)$$

em que $\mathbf{y} = (y_1, \dots, y_n)^T$ é um vetor de dimensão n de observações de uma variável aleatória Y , $f(\cdot)$ é uma função densidade de probabilidade e $\boldsymbol{\theta}$ é o vetor de parâmetros.

As distribuições a priori são importantes, pois nelas onde é incluída a experiência (conhecimento prévio) sobre os parâmetros a serem estimados. Contudo escolher tais distribuições é uma tarefa difícil, pois assumir tal informação a priori pode ajudar a estimação dos parâmetros ou não. No caso em que se tem uma informação a priori verídica sobre um respectivo parâmetro, isso pode ajudar na estimação dos parâmetros do modelo.

Existem dois tipos de prioris: as informativas que é o caso, onde o pesquisador tem uma informação prévia sobre os parâmetros e as minimamente informativas que é o cenário em que o pesquisador não tem nenhuma experiência sobre os parâmetros. Um priori informativa seria quando o pesquisador sabe aproximadamente qual o valor esperado ou a variância da distribuição e uma priori minimamente informativa conhecido na literatura é a priori de Jeffreys (JEFFREYS, 1998).

Dada a função de verossimilhança $f(\mathbf{y}|\boldsymbol{\theta})$ e as distribuições a priori, podemos encontrar a função densidade de probabilidade da distribuição a posteriori. Considerando que cada parâmetro presente no vetor $\boldsymbol{\theta}$ é independente dos outros, ou seja, a função densidade de probabilidade a priori conjunta é dada pelos produtos das densidades a priori, que é definida por:

$$f(\boldsymbol{\theta}) = f(\theta_1) \cdots f(\theta_p), \quad (14)$$

logo a distribuição a posteriori, removendo a constante normalizadora pode ser reescrita

$$f(\boldsymbol{\theta}|\mathbf{y}) = \frac{f(\boldsymbol{\theta}, \mathbf{y})}{f(\mathbf{y})} = \frac{f(\mathbf{y}|\boldsymbol{\theta})f(\boldsymbol{\theta})}{f(\mathbf{y})} \propto f(\mathbf{y}|\boldsymbol{\theta})f(\boldsymbol{\theta}) = f(\mathbf{y}|\boldsymbol{\theta})f(\theta_1) \cdots f(\theta_p).$$

Para obter amostras da distribuição a posteriori nos casos que sua obtenção é difícil analiticamente, frequentemente é utilizado um método que gera amostras a partir de simulações de Monte Carlo via cadeia de Markov (*Markov Chain Monte Carlo – MCMC*). Esse método é um grande portfólio de mecanismos computacionais para geração de amostras de uma distribuição de probabilidade (NEAL, 2011).

Para obter as amostras das distribuições a posteriori apresentadas, é necessário determinar qual *MCMC* é o método adequado, para cada caso. Alguns algoritmos *MCMC* conhecidos na literatura são o Amostrador de Gibbs (*Gibbs Sampling*) e o algoritmo de Metropolis-Hastings (ROBERT; CASELLA, 2011).

O primeiro, introduzido por Geman e Geman (1984), pode ser utilizado quando as distribuições condicionais completas são conhecidas. Seu algoritmo é apresentado a seguir:

¹ Seja $\boldsymbol{\theta}$ o vetor de parâmetros de comprimento d e $\boldsymbol{\theta}^T = (\theta_1^{(t)}, \theta_2^{(t)}, \dots, \theta_d^{(t)})$ o vetor de parâmetros na t -ésima iteração.

¹Nota: Medidas que indicam a convergência serão apresentados posteriormente.

1. Especificar o vetor de valores iniciais dos parâmetros, denotado por:

$$\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, \dots, \theta_d^{(0)})^T, \quad (15)$$

e inicializar o contador da cadeia na primeira iteração, isto é, fazer $t = 1$.

2. Gerar valores para $\theta^{(t)}$, $t = 1, \dots, d$, a partir dos valores $\theta^{(t-1)}$ por meio de simulações das distribuições condicionais completas, da seguinte forma

$$\begin{aligned} \theta_1^{(t)} &\sim f(\theta_1 | \theta_2^{(t-1)}, \theta_3^{(t-1)}, \dots, \theta_d^{(t-1)}, \mathbf{y}) \\ \theta_2^{(t)} &\sim f(\theta_2 | \theta_1^{(t)}, \theta_3^{(t-1)}, \dots, \theta_d^{(t-1)}, \mathbf{y}) \\ &\vdots \\ \theta_d^{(t)} &\sim f(\theta_d | \theta_1^{(t)}, \theta_2^{(t)}, \dots, \theta_{d-1}^{(t)}, \mathbf{y}). \end{aligned} \quad (16)$$

3. Atualizar o contador de t para $t + 1$ e voltar ao passo 2 até que a convergência seja alcançada.

No caso, que as distribuições condicionais completas não são conhecidas, o amostrador de Gibbs não poderá ser utilizado e teremos que recorrer a algoritmos *MCMC*, tais como o Metropolis-Hastings e o Monte Carlo Hamiltoniano.

O algoritmo Metropolis foi inicialmente proposto por Metropolis et al. (1953) nos meados da década de 1950, e estendido após duas décadas, por Hastings (1970) originando, assim, o algoritmo denominado Metropolis-Hastings (*MH*). O algoritmo *MH* simula valores da distribuição a posteriori do vetor de parâmetros $\boldsymbol{\theta}$, que são aceitos com probabilidade α , definida por:

$$\alpha = \min \left\{ 1, \frac{f(\boldsymbol{\theta}^{(t)} | \mathbf{y}) q(\boldsymbol{\theta}^{(t-1)} | \boldsymbol{\theta}^{(t)})}{f(\boldsymbol{\theta}^{(t-1)} | \mathbf{y}) q(\boldsymbol{\theta}^{(t)} | \boldsymbol{\theta}^{(t-1)})} \right\}, \quad (17)$$

em que: $f(\cdot)$ é a distribuição de interesse; $q(\cdot)$ é uma distribuição conhecida, geralmente simétrica, como, por exemplo a distribuição normal, que tem por objetivo aumentar a velocidade de convergência do algoritmo.

De acordo com Chib e Greenberg (1995), o algoritmo é definido pelos seguintes passos:

1. Especificar o vetor de valores iniciais, denotado por:

$$\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, \dots, \theta_d^{(0)})^T, \quad (18)$$

e inicializar o contador da cadeia na primeira iteração isto é, fazer $t = 1$.

2. Gerar novos valores do vetor $\boldsymbol{\theta}^{(t)}$ por meio da função de densidade de probabilidade $q(\boldsymbol{\theta}^{(t)}|\boldsymbol{\theta}^{(t-1)})$, depois um valor Υ é gerado de uma distribuição uniforme $u \sim U(0, 1)$.
3. A equação (17) é utilizada na avaliação da probabilidade de aceitação e rejeição, da seguinte maneira,

$$\boldsymbol{\theta}^t \leftarrow \begin{cases} \boldsymbol{\theta}^t, & \alpha \geq \Upsilon \\ \boldsymbol{\theta}^{t-1}, & \text{caso contrário.} \end{cases} \quad (19)$$

Note que, quando $q(\boldsymbol{\theta}^{(t-1)}|\boldsymbol{\theta}^{(t)}) = q(\boldsymbol{\theta}^{(t)}|\boldsymbol{\theta}^{(t-1)})$ o que acontece quando, por exemplo, $q(\boldsymbol{\theta}^{(t)}|\boldsymbol{\theta}^{(t-1)}) \sim N_d(\boldsymbol{\theta}^{(t-1)}, \mathbf{I}_d \sigma_q^2)$, em que: \mathbf{I}_d é a matriz identidade com dimensões $d \times d$ e σ_q^2 é a variância de $q(\cdot)$, a equação (20) é simplificada para:

$$\alpha = \min \left(1, \frac{f(\boldsymbol{\theta}^{(t)}|\mathbf{y})}{f(\boldsymbol{\theta}^{(t-1)}|\mathbf{y})} \right), \quad (20)$$

definindo o denominado algoritmo Metropolis.

4. Atualizar o contador de t para $t + 1$ e voltar ao passo 2 até que a convergência seja alcançada.

Os métodos *MCMC*, mostrados até aqui, podem demorar até convergir para a distribuição alvo pelo motivo de o passeio aleatório local que acontece nas simulações, quando a série amostrada fica vagando a caminho da distribuição, podendo fazer com que a simulação demore até a convergência.

Por volta de meados da década de 1980 um trabalho marcante da Duane et al. (1987) fez a junção do *MCMC* com uma abordagem da dinâmica molecular, assim, criando o algoritmo Monte Carlo Híbrido ou também chamado Monte Carlo Hamiltoniano (*Hamiltonian Monte Carlo - HMC*). Este foi primeiramente utilizado nas simulações da teoria do campo da rede da cromodinâmica quântica, e pelos meados da década de 1990, começou a ser incluído nas aplicações estatísticas (NEAL, 2011).

Gelman et al. (2013) sugerem outro algoritmo *MCMC* denominado Monte Carlo Híbrido ou também chamado Monte Carlo Hamiltoniano (*Hamiltonian Monte Carlo - HMC*), a fim de minimizar o comportamento do passeio aleatório local do algoritmo *MH*, assim conseguindo chegar a da densidade a posteriori $f(\boldsymbol{\theta}|\mathbf{y})$ rapidamente, com uma taxa de aceitação maior que o algoritmo *MH*.

Em aplicações estatísticas, o algoritmo, utiliza a dinâmica hamiltoniana, da forma que utilizam a nomenclatura posição para o vetor de parâmetros $\boldsymbol{\theta}$ e $\boldsymbol{\rho}$ chamado de momento que é um vetor de valores com mesmo comprimento do vetor posição.

Com isso, a função densidade de probabilidade da distribuição a posteriori $f(\boldsymbol{\theta}|\mathbf{y})$ adicionando uma função de densidade de probabilidade independente $f(\boldsymbol{\rho})$, em que $\boldsymbol{\rho} = (\rho_1, \dots, \rho_I)^T$ é um vetor de variáveis independentes, com mesmo comprimento do vetor $\boldsymbol{\theta}$ e $\rho_i \sim N(0, M_{ii})$, $i = 1, \dots, I$. Desta maneira a função de densidade de probabilidade conjunta é $f(\boldsymbol{\theta}, \boldsymbol{\rho}|\mathbf{y}) = f(\boldsymbol{\theta}|\mathbf{y})f(\boldsymbol{\rho})$.

Como temos interesse apenas, em obter amostras de $\boldsymbol{\theta}$, o $\boldsymbol{\rho}$ é apenas uma variável auxiliar, que ajuda a aumentar a velocidade de busca do algoritmo no espaço paramétrico. A estrutura do algoritmo *HMC* será descrita em resumo, pelos seguintes passos (GELMAN et al., 2013):

1. Especificar os valores iniciais de ϵ e dos vetores, denotados por:

$$\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, \dots, \theta_d^{(0)})^T, \quad (21)$$

$$\boldsymbol{\rho}^{(0)} = (\rho_1^{(0)}, \dots, \rho_d^{(0)})^T \quad (22)$$

e inicializar o contador da cadeia na primeira iteração, isto é, fazer $t = 1$.

2. Os valores do vetor dimensão $\boldsymbol{\rho}^t$ são gerados por sua respectiva função densidade de probabilidade, que é uma distribuição $N_I(\mathbf{0}, \mathbf{M})$, em que $\mathbf{0}$ é um vetor de zeros e \mathbf{M} é a matriz massa, assim chamada por semelhança ao modelo dinâmico hamiltoniano da Física.
3. Simultaneamente $(\boldsymbol{\theta}^t, \boldsymbol{\rho}^t)$ são atualizados L vezes por um algoritmo de integração “*leapfrog*”, que utiliza as equações hamiltonianas (23), (24) e (25) a cada salto, que é dimensionado por um fator ϵ (período de tempo ou tamanho do salto), dessa forma os vetores $\boldsymbol{\theta}^t$ e $\boldsymbol{\rho}^t$ são atualizados a cada etapa *leapfrog*. Essas etapas são definidas respectivamente por:
 - a. Atualizar $\boldsymbol{\rho}^t$ utilizando a derivada do logaritmo da função densidade de probabilidade a posteriori de $\boldsymbol{\theta}^{t-1}$:

$$\boldsymbol{\rho}^t \leftarrow \boldsymbol{\rho}^{t-1} + \frac{1}{2}\epsilon \frac{\partial \log(f(\boldsymbol{\theta}^{t-1}|\mathbf{y}))}{\partial \boldsymbol{\theta}}. \quad (23)$$

- b. Atualizar $\boldsymbol{\theta}^t$ utilizando o $\boldsymbol{\rho}^t$

$$\boldsymbol{\theta}^t \leftarrow \boldsymbol{\theta}^{t-1} + \epsilon \mathbf{M}^{-1} \boldsymbol{\rho}^t, \quad (24)$$

em que \mathbf{M}^{-1} é a inversa da matriz de variâncias e covariância do $\boldsymbol{\rho}$.

- c. Atualizar novamente $\boldsymbol{\rho}^t$ utilizando $\boldsymbol{\theta}^t$.

$$\boldsymbol{\rho}^t \leftarrow \boldsymbol{\rho}^t + \frac{1}{2}\epsilon \frac{\partial \log(f(\boldsymbol{\theta}^t|\mathbf{y}))}{\partial \boldsymbol{\theta}}. \quad (25)$$

4. Os valores atualizados $\boldsymbol{\theta}^t$ e $\boldsymbol{\rho}^t$ serão aceitos da seguinte forma: primeiro é gerado um valor Υ de uma distribuição uniforme $u \sim U(0, 1)$, depois é calculado um valor α , que é definido por:

$$\alpha = \min \left(1, \frac{f(\boldsymbol{\theta}^t | \mathbf{y}) f(\boldsymbol{\rho}^t)}{f(\boldsymbol{\theta}^{t-1} | \mathbf{y}) f(\boldsymbol{\rho}^{t-1})} \right), \quad (26)$$

logo, a condição para aceitar o vetor de parâmetros proposto, é dado por:

$$\boldsymbol{\theta}^t \leftarrow \begin{cases} \boldsymbol{\theta}^t, & \alpha \geq \Upsilon \\ \boldsymbol{\theta}^{t-1}, & \text{caso contrário.} \end{cases} \quad (27)$$

5. Atualizar o contador de t para $t + 1$ e voltar ao passo 2 até que a convergência seja alcançada.

Em algumas situações no algoritmo *HMC*, determinar uma boa quantidade de saltos e intervalos entre cada salto, pode ser uma tarefa difícil. Esse problema pode ocasionar uma demora na convergência para a distribuição alvo, ou até não convergir. Uma extensão do algoritmo *HMC* que foi proposta por Hoffman e Gelman (2014) foi o chamado amostrador *No-U-Turn* (*No-U-Turn Sampler - NUTS*), ele faz a escolha dos saltos automaticamente, evitando o problema existente no método *HMC*. Uma descrição melhor desse outro algoritmo pode ser vista em Hoffman e Gelman (2014), Nishio e Arakawa (2019).

Os métodos *MCMC* que foram mostrados podem demorar até atingir indícios de convergência entre as cadeias simuladas, por isso é necessário alguns artifícios que ajudam a melhorar isto. Dessa forma Lesaffre e Lawson (2012) sugerem formas de melhorar a convergência do algoritmo, a primeira delas é não utilizar todos os valores gerados em cada simulação, pois podem resultar em amostras não estacionárias. Esse artifício é chamado espaçamento (*thinning*) que é um valor inteiro positivo que determina de quanto a quanto os valores simulados devem ser utilizados, por exemplo se o espaçamento escolhido for 10, ou seja, após o primeiro valor contabilizado, o décimo próximo valor será contabilizado e assim sucessivamente. Com isso podemos obter amostras estacionárias das distribuições a posteriori.

Os valores iniciais das simulações podem influenciar nas médias a posteriori dos parâmetros, uma maneira de solucionar esse problema é utilizar um técnica chamada *Burn-in*, que retira uma determinada quantidade de valores iniciais, por exemplo se o *Burn-in* escolhido é 100, em uma simulação com mil iterações, os 100 primeiros valores gerados são desconsiderados.

Existem outras medidas que servem para verificação de convergência entre as cadeias simuladas, pelos algoritmos *MCMC*, que são fator de redução de escala potencial estimada

(potential scale reduction estimated - \widehat{R}) e o tamanho efetivo da amostra (*effective sample size* - n_{eff}). O \widehat{R} é obtido por meio das simulações para cada valor θ que são descritas como, θ_{ij} , $i = 1, \dots, n$, $j = 1, \dots, m$, em que n é a número de iterações, $n > 1$ e m é o número de cadeias, a partir dos θ_{ij} pode-se calcular as variações entre e dentro das cadeias (GELMAN et al., 2013). Primeiramente são calculados os valores de B e W , que são definidos como

$$B = \frac{n}{m-1} \sum_{j=1}^m (\bar{\theta}_{.j} - \bar{\theta}_{..})^2, \quad (28)$$

$$W = \frac{1}{m} \sum_{j=1}^m s_j^2, \quad (29)$$

em que: $\bar{\theta}_{.j} = \frac{1}{n} \sum_{i=1}^n \theta_{ij}$; $\bar{\theta}_{..} = \frac{1}{n} \sum_{j=1}^m \bar{\theta}_{.j}$ e $s_j^2 = \frac{1}{n-1} \sum_{i=1}^n (\theta_{ij} - \bar{\theta}_{.j})^2$.

Por meio de uma ponderação da média das equações (28) e (29), é possível encontrar a estimava da variância de $(\theta|y)$, que é definida por:

$$\widehat{var}^+(\theta|y) = \frac{n-1}{n} W + \frac{1}{n} B. \quad (30)$$

Utilizando as equações (29) e (30) se encontra o fator de redução de escala potencial, que é dado por:

$$\widehat{R} = \sqrt{\frac{\widehat{var}^+(\theta|y)}{W}}. \quad (31)$$

Segundo Lesaffre e Lawson (2012) o valor substancialmente ótimo para o \widehat{R} é 1, podemos concluir que entre as m cadeias e suas respectivas n simulações está indicando a convergência. Para estimar a segunda medida de convergência, que é chamada de tamanho efetivo da amostra é necessário utilizar a estimativa de autocorrelação dentro das cadeia ρ , entre as m cadeias e suas respectivas n iterações, dessa forma o tamanho efetivo da amostra é definido por:

$$n_{eff} = \frac{mn}{1 + 2 \sum_{t=1}^{\infty} \rho_t}, \quad (32)$$

em que ρ_t é a autocorrelação da cadeia θ na defasagem t . Utilizando a soma parcial positiva como a estimativa do $\sum_{t=1}^{\infty} \rho_t$ a equação (32) pode ser reescrita da seguinte forma:

$$\hat{n}_{eff} = \frac{mn}{1 + 2 \sum_{t=1}^T \hat{\rho}_t}, \quad (33)$$

em que T é o primeiro número inteiro positivo ímpar para o qual $\hat{\rho}_{T+1} + \hat{\rho}_{T+2}$ é negativo e $\hat{\rho}_t$ é a estimativa da autocorrelação da cadeia θ na defasagem t , que é definida por:

$$\hat{\rho}_t = 1 - \frac{V_t}{2\widehat{var}^+} \quad (34)$$

e

$$V_t = \frac{1}{m(n-1)}(\theta_{i,j} - \theta_{i-t,j})^2. \quad (35)$$

Além dessas medidas de verificação de convergência mostrada até aqui, o Lesaffre e Lawson (2012) sugerem alguns gráficos que podem indicar visualmente, que as cadeias convergiram, são os gráficos de: linhas, autocorrelação, gráfico de dispersão.

O primeiro é o gráfico de linhas das simulações das amostras a posteriori, também conhecido como *trace plot*, evidencia o quão rápido a cadeia converge para distribuição alvo e também se houve mistura entre as cadeias simuladas. Caso aconteça misturas entre todas, é um indício de convergência nas simulações.

Já o gráfico de autocorrelação pode ilustrar se as amostras obtidas pelo método *MCMC* estão gerando uma distribuição estacionária, no caso onde existe altas correlações em todas as defasagens, indica que as simulações não atingiram uma distribuição estacionária, ou seja, evidenciando que não há indícios de convergência. Entretanto, existem maneiras de melhorar a convergência, uma delas é aumentar o número de iterações e a outra aumenta *thinning*.

Para verificar se os parâmetros possuem uma dependência a posteriori, é utilizado o gráfico de dispersão, também conhecido como *Scatter plot*. Mesmo após a análise gráfica da convergência e também pelas medidas de indícios de convergência das simulações \hat{R} , é necessário alguma medida estatística que forneça a informação de que o modelo se ajustou bem aos dados. Uma forma é calcular uma estatística que nos forneça essa informação.

Gelman et al. (2013) sugerem uma medida chamada verificação preditiva a posteriori (*Posterior Predictive Check - PPC*), que é uma comparação entre a quantidade de teste realizados para os dados observados, $T(\mathbf{y}, \boldsymbol{\theta})$, e a quantidade de testes realizados para os valores preditos, $T(\mathbf{y}^{rep}, \boldsymbol{\theta})$. A partir dos testes podemos estimar um p -valor que é definido como a probabilidade de que os valores preditos dos dados podem ser mais extremos do que os dados observados, conforme medido pela quantidade de teste:

$$p_d = P(T(\mathbf{y}^{rep}, \boldsymbol{\theta}) \geq T(\mathbf{y}, \boldsymbol{\theta}) | \mathbf{y}),$$

em que: $T(\cdot)$ é a medida estatística, que pode ser a média, mediana entre outras; \mathbf{y} é o vetor de valores observados; \mathbf{y}^{rep} é o vetor de valores preditos e $\boldsymbol{\theta}$ é o vetor de estimativas a posteriori dos parâmetros.

O p_d obtido de amostras *MCMC*, é a proporção das K simulações, em que a quantidade do teste $T(\mathbf{y}^{rep(k)}, \boldsymbol{\theta}^k)$ é igual ou superior ao $T(\mathbf{y}, \boldsymbol{\theta}^k)$, que é dado por:

$$p_d = \frac{1}{K} \sum_{k=1}^K I(T(\mathbf{y}^{rep(k)}, \boldsymbol{\theta}^k) \geq T(\mathbf{y}, \boldsymbol{\theta}^k)), \quad k = 1, \dots, K, \quad (36)$$

em que: $I(\cdot)$ é a função indicadora, que resulta no valor um se a expressão entre parênteses for verdade e zero caso contrário. No caso de se utilizar a estatística de teste $T(\cdot)$ é a média teremos o p_d definido como:

$$p_d = \frac{1}{K} \sum_{k=1}^K I \left(\frac{\sum_{i=1}^n y_i^{rep(k)}}{n} \geq \frac{\sum_{i=1}^n y_i}{n} \right), \quad k = 1, \dots, K, \quad i = 1, \dots, n, \quad (37)$$

em que: y_i foi i -ésimo valor observado; $y_i^{rep(k)}$ foi i -ésimo valor predito na k -ésima simulação. De acordo com Gelman (2013) o valor $p_d = 0,5$ é o ideal, indicando que o modelo está bem ajustado aos dados.

Quando existe mais de um modelo bem ajustado aos dados, é importante escolher apenas um modelo, de preferência o que seja mais parcimonioso. Uma maneira de comparar modelos é utilizar critérios de informações, pois eles revelam qual modelo apresenta a menor perda em informação quando comparado aos dados. Dentre os critérios de seleção de modelos Bayesianos, o mais conhecido é o critério de informação de desvio (*Deviance Information Criterion - DIC*) proposto por Spiegelhalter et al. (2014). Este critério é uma generalização do critério de informação de Akaike (*Akaike Information Criterion - AIC*) (AKAIKE, 1974), que é definido por:

$$DIC = -2\widehat{elpd}_{dic} = -2(\log f(\mathbf{y}|\hat{\boldsymbol{\theta}}) - p_{dic}), \quad (38)$$

$$\log f(\mathbf{y}|\hat{\boldsymbol{\theta}}) = \sum_{i=1}^n \log f(y_i|\hat{\boldsymbol{\theta}}), \quad (39)$$

$$p_{dic} = 2 \left(\sum_{i=1}^n \log f(y_i|\hat{\boldsymbol{\theta}}) - \frac{1}{K} \sum_{k=1}^K \sum_{i=1}^n \log f(y_i|\boldsymbol{\theta}^k) \right), \quad (40)$$

em que: $elpd_{dic}$ é o logaritmo esperado da densidade preditiva pontual; p_{dic} é o número efetivo de parâmetros; $\log f(\mathbf{y}|\boldsymbol{\theta})$ é o logaritmo da função de verossimilhança; \mathbf{y} é o vetor de observações; $\boldsymbol{\theta}$ é o vetor de parâmetros desconhecidos e $\hat{\boldsymbol{\theta}}$ é a estimativa média a posteriori dos parâmetros.

Uma melhoria do *DIC* que foi proposto por Watanabe (2010) denominado critério de informação amplamente aplicável (*Widely Applicable Information Criterion - WAIC*). Para calcular esse critério é necessário o valor esperado do logaritmo da densidade preditiva pontual (*expected log pointwise predictive density - elpd*) para uma nova observação i , em que $i = 1, \dots, n$, que é dado por:

$$elpd = \sum_{i=1}^n \int f_t(\tilde{y}_i) \log f(\tilde{y}_i|\mathbf{y}) d\tilde{y}_i, \quad (41)$$

em que $f_t(\tilde{y}_i)$ é a distribuição do verdadeiro processo de geração de dados para \tilde{y}_i e $\log f(\tilde{y}_i|\mathbf{y})$ é o logaritmo da densidade preditiva a posteriori.

Depois o logaritmo da densidade preditiva pontual (*log pointwise predictive density* - lpd), que é dado por:

$$lpd = \sum_{i=1}^n \log f(y_i | \mathbf{y}) = \sum_{i=1}^n \log \int \cdots \int f(y_i | \boldsymbol{\theta}) f(\boldsymbol{\theta} | \mathbf{y}) d\boldsymbol{\theta} \quad (42)$$

e sua estimativa é obtida por meio de reamostragem, onde $\boldsymbol{\theta}^k$ amostras simuladas da função de densidade de probabilidade a posteriori são calculadas pela seguinte expressão:

$$\widehat{lpd} = \sum_{i=1}^n \log \left(\frac{1}{K} \sum_{k=1}^K f(y_i | \boldsymbol{\theta}^k) \right), \quad k = 1, \dots, K. \quad (43)$$

Podemos agora estimar a medida logaritmo da precisão esperada, que segundo Watanabe (2010) é dada por:

$$e\widehat{lpd}_{waic} = \widehat{lpd} - \hat{p}_{waic}, \quad (44)$$

$$\hat{p}_{waic} = 2 \sum_{i=1}^n \left(\log \left(\frac{1}{K} \sum_{k=1}^K f(y_i | \boldsymbol{\theta}^k) \right) - \frac{1}{K} \sum_{k=1}^K \log f(y_i | \boldsymbol{\theta}^k) \right), \quad (45)$$

Assim, o critério *WAIC* é definido por:

$$WAIC = -2e\widehat{lpd}_{waic}, \quad (46)$$

em que o modelo que obtiver o menor valor *WAIC* é o modelo mais parcimonioso.

Outro método que utiliza a soma do logaritmo da função de verossimilhança avaliada nas amostras da distribuição a posteriori dos parâmetros avaliadas com a retirada da i -ésima observação (*leave-one-out* - *loo*). Segundo Vehtari, Gelman e Gabry (2016) a medida logaritmo da precisão preditiva *loo* pode ser estimada pela seguinte equação:

$$elpd_{loo} = \sum_{i=1}^n \log f(\tilde{y}_i | \tilde{\mathbf{y}}_{i-1}), \quad (47)$$

em que $f(y_i | \tilde{\mathbf{y}}_{i-1})$ é a função densidade de probabilidade preditiva *leave-one-out* dada a amostra sem o i -ésima observação. Sua respectiva estimativa é obtida pela seguinte expressão dada por:

$$e\widehat{lpd}_{loo} = \sum_{i=1}^n \log \left(\frac{\sum_{k=1}^K w_i^k f(\tilde{y}_i | \tilde{\mathbf{y}}_{i-1} | \boldsymbol{\theta}^k)}{\sum_{k=1}^K w_i^k} \right), \quad (48)$$

em que w_i^k é um vetor de pesos para i -ésima observação.

Logo, o critério (*leave-one-out* - *looic*) é dado por:

$$looic = -2e\widehat{lpd}_{loo}. \quad (49)$$

Neste critério, o modelo que obtiver o menor valor *loaic* é o modelo que deve ser escolhido. Outros dois critérios também utilizados na inferência bayesiana, são o *EAIC* e o *EBIC*.

O critério proposto por Brooks et al. (2002) denominado critério de informação Akaike esperado (*Expected Akaike Information Criterion - EAIC*), foi definido como o valor esperado do *AIC*, que é dada por:

$$\widehat{EAIC} = -2 \sum_{i=1}^n \log f(y_i | \boldsymbol{\theta}) + 2p, \quad (50)$$

em que, p é o número de parâmetros do modelo em estudo. O modelo que apresentar o menor valor \widehat{EAIC} é o que deve ser escolhido.

E o critério de informação bayesiana esperado (*Expected Bayesian Information Criterion - EBIC*) que foi proposto por Carlim B.; Louis (2001), que é o valor esperado do critério BIC (Bayesian Information Criterion) proposto por Schwarz (1978). Logo o *EBIC* é dado por:

$$\widehat{EBIC} = -2 \sum_{i=1}^n \log f(y_i | \boldsymbol{\theta}) + p \log(n), \quad (51)$$

em que, p é o número de parâmetros do modelo em estudo e n é o tamanho amostral. O modelo que tiver menor valor do critério é o que deve ser escolhido.

3 MATERIAL E MÉTODOS

3.1 Material

A fim de aplicar os procedimentos de estimação propostos neste trabalho, foram selecionadas cepas do fungo *C. truncatum* em de 13 diferentes áreas produtoras de soja com problemas de antracnose nos estados de Goiás e Mato Grosso nomeadas: GO2_12, GO4_07, GO4_08, GO5_11, GO5_14, GO5_25, MT2_05, MT3_01, MT3_21, MT4_05, MT4_13, MT5_12 e MT5_32.

No início de 2021, as cepas foram estudadas e agrupadas por Rogério e Massola Junior (2019) em três grupos genéticos distintos, denominados, Grupo 1, Grupo 2 e Grupo 3. As pertencentes ao Grupo 1 foram GO2_12, GO4_08, GO5_25, MT2_05 e MT3_01. As pertencentes ao Grupo 2, por sua vez, foram GO5_11, MT3_21, MT4_05 e MT5_12. Finalmente, as pertencentes ao Grupo 3 foram GO4_07, GO5_14, MT4_13 e MT5_32. Assim o intuito de verificar se as diferentes cepas ou se os diferentes grupos genéticos apresentam severidades diferentes, foi conduzido um experimento de janeiro a junho de 2021, no Departamento Fitopatologia da Escola Superior de Agricultura “Luiz de Queiroz”, ESALQ/USP.

Para este experimento, foi adotado o delineamento inteiramente ao acaso com 14 tratamentos (13 cepas + controle) e 3 repetições. As sementes pré germinadas foram inoculadas com suspensão de 1×10^6 conídios por *mL* do patógeno e, em seguida, foram semeadas em vasos alocados em câmara de crescimento. Cada vaso constituiu de uma unidade experimental contendo 3 plantas de soja.

Passados cerca de 30 dias, as partes aéreas das plantas foram fixadas em folhas de papel branco por meio de fita adesiva transparente e as folhas contendo as plantas foram então fotografadas gerando imagens como está apresentada na Figura 2, essas imagens foram cedidas pela pesquisadora Flávia Rogério (2019).

Neste estudo a variável resposta de interesse é a severidade da doença, que, segundo Amorim (1995), é a porcentagem da área do tecido com sintomas. As plantas foram retiradas dos vasos e separadas por tratamento e repetição para a mensuração da severidade em cada organismo.

Depois foram capturadas imagens de cada respectiva planta e com essas imagens e a função `symptomatic_area` do pacote `pliman` (OLIVOTO, 2021) da linguagem R (R Core Team, 2021) foi calcula a porcentagem de severidade presente na área foliar de cada planta por meio das cores presentes na imagem, para utilizar a respectiva função, existe a necessidade de declarar quatro argumentos.

O primeiro argumento é o `img`, que é a imagem da planta que será avaliada, o segundo argumento é o `img_healthy` que é uma imagem mostrando uma área que seja



Figura 2. Plantas da cultura da soja, inoculadas com a cepa MT5-12, após 30 dias da semeadura.

considerada saudável para servir como base para a função, terceiro o **img_symptoms** que é uma imagem mostrando uma área que tenha a doença para servir como base para a função e por fim **img_background** atribuir uma imagem que contenha a cor do fundo que foi utilizado como base para tirar as fotos das plantas, por exemplo se a foto da planta foi capturada sobre um fundo branco, o **img_background** deve ser uma imagem branca. As respectivas imagens que foram utilizadas nos argumentos **img_healthy** sendo a Figura 3 **healthy**, **img_symptoms** sendo a Figura 3 **symptoms** e **img_background** sendo a Figura 3 **background** da função **symptomatic_area**.

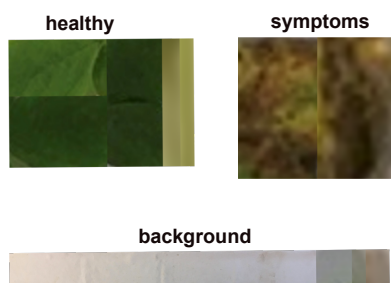


Figura 3. Imagens que foram utilizadas como argumentos na função **symptomatic_area**.

Com os respectivos argumentos presentes, a função **symptomatic_area** faz uma estimativa da porcentagem da severidade, o que pode gerar algum viés nos resultados. Para que isso fosse evitado, cada imagem foi analisada pela função mil vezes, depois calculado a média desses mil valores para cada uma, assim gerando os dados que estão expostos na Tabela 2. Um roteiro de programação da linguagem R, para analisar a severidade de uma

planta da soja pode ser visto Apêndice I.

Na Figura 2 podemos observar a imagem com nove plantas de soja inoculada com a cepa MT5_12, e pode-se notar que os galhos circulosados de vermelho tem traços do patógeno bem visível nas áreas foliares.

Tabela 2. Proporções dos pixels das imagens das plantas de soja com coloração não verde (severidade) segundo grupo genético, o tratamento aplicado, o vaso e a média.

Grupo		Vaso			Média
Genético	Tratamento	1	2	3	
Grupo 1	GO2_12	0,017245	0,016004	0,018750	0,017333
	GO5_25	0,069472	0,042412	0,047448	0,053111
	GO4_08	0,006261	0,003700	0,007028	0,005663
	MT2_05	0,032511	0,008444	0,017563	0,019506
	MT3_01	0,019640	0,026801	0,027687	0,024709
Grupo 2	MT3_21	0,108647	0,141415	0,101186	0,117082
	MT4_05	0,179279	0,058026	0,030375	0,089227
	GO5_11	0,004398	0,002144	0,000890	0,003098
	MT5_12	0,052347	0,020050	0,023854	0,032084
Grupo 3	GO5_14	0,020913	0,011177	0,022080	0,018057
	MT4_13	0,058740	0,328004	0,585577	0,324107
	GO4_07	0,001941	0,014672	0,012876	0,009829
	MT5_32	0,075947	0,021119	0,010918	0,035995
Controle		0,025313	0,021013	0,131149	0,059158

3.2 Métodos

3.2.1 Modelo para dados de uma amostra aleatória simples de \mathbf{Y}

Considere $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$ uma amostra aleatória simples de tamanho n de $Y \sim \text{Beta}(\beta, \phi)$ e $\mathbf{y} = (y_1, y_2, \dots, y_n)$ um conjunto de observações de \mathbf{Y} . Neste caso, o modelo estatístico será

$$\begin{aligned}
 Y_u &\overset{\text{ind.}}{\sim} \text{Beta}(\mu_u, \phi), \quad u = 1, \dots, n, \\
 \mu_u &= g^{-1}(\eta_u), \\
 \eta_u &= \beta,
 \end{aligned} \tag{52}$$

em que: $\mu_u = E(Y_u)$ é o valor esperado de Y_u ; η_u é o valor do preditor linear; β é parâmetro de escala; $g(\cdot)$ é uma função monótona e diferenciável, chamada função de ligação, sendo que as mais comumente utilizadas são apresentadas na Tabela 1. Note que, neste caso, $E(Y_u) = \mu_u = g^{-1}(\beta)$, ou seja, é constante para $u = 1, \dots, n$.

Se utilizarmos, por exemplo, a função logito, teremos que

$$\eta_u = \log \left(\frac{\mu_u}{1 - \mu_u} \right) = \beta \quad (53)$$

e

$$\mu_u = g^{-1}(\eta_u) = \frac{\exp(\eta_u)}{1 + \exp(\eta_u)} = \frac{\exp(\beta)}{1 + \exp(\beta)}, \quad (54)$$

constante para $u = 1, \dots, n$. Além das funções de ligação apresentadas na Tabela 1, podemos considerar, de forma geral, a família de funções de ligação apresentada por Aranda-Ordaz (1981), originalmente para dados binários, dada por

$$\eta = g(\mu) = \log \left\{ \frac{(1 - \mu)^{-c} - 1}{c} \right\} \quad (55)$$

e sua inversa

$$\mu = g^{-1}(\eta) = 1 - \{1 + c \exp(\eta)\}^{-\frac{1}{c}}, \quad (56)$$

em que $c > 0$ é um parâmetro a ser estimado. Como casos particulares dessa família, temos, para $c = 1$, a função de ligação logito e para $c \rightarrow 0^+$, a função de ligação complemento log-log.

Voltemos ao modelo (52). Vamos escrevê-lo, agora, na sua forma matricial, com a notação apresentada por Ferrari e Cribari-Neto (2004), ou seja,

$$\begin{aligned} Y_u &\overset{ind.}{\sim} \text{Beta}(\mu_u, \phi), \quad u = 1, \dots, n \\ \boldsymbol{\mu} &= g^{-1}(\boldsymbol{\eta}), \\ \boldsymbol{\eta} &= \mathbf{X}\boldsymbol{\beta}, \end{aligned} \quad (57)$$

em que: $\beta = \beta$, escalar; \mathbf{X} é a matriz de delineamento (no caso, $\mathbf{X} = \mathbf{1}_n$, isto é, igual a um vetor de comprimento n); $g(\cdot)$ é a função de ligação e $\boldsymbol{\mu}$ e $\boldsymbol{\eta}$ são, respectivamente, os vetores de valores esperados e de preditores lineares, ambos com comprimento n , ou seja,

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix}, \quad \mathbf{X} = \mathbf{1}_n = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \beta = \beta, \quad \boldsymbol{\eta} = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix}$$

e

$$\mathbf{X}\boldsymbol{\beta} = \begin{pmatrix} \beta \\ \beta \\ \vdots \\ \beta \end{pmatrix}.$$

3.2.2 Modelo para dados provenientes de experimentos inteiramente ao acaso com um fator e ϕ não variável

Consideremos, a seguir, o modelo para a análise dos dados provenientes de experimentos inteiramente ao acaso com I tratamentos e J repetições por tratamento, como ilustra a Tabela 3.

Tabela 3. Observações da variável Y num experimento inteiramente ao acaso com I tratamentos e J repetições

Tratamento	Repetição			
	1	2	...	J
1	y_{11}	y_{12}	...	y_{1J}
2	y_{21}	y_{22}	...	y_{2J}
\vdots	\vdots	\vdots	\ddots	\vdots
I	y_{I1}	y_{I2}	...	y_{IJ}

Supondo que $Y_{ij} \sim \text{Beta}(\mu_{ij}, \phi)$, $i = 1, \dots, I$, $j = 1, \dots, J$, são variáveis independentes e que y_{ij} é uma observação de Y_{ij} , um modelo estatístico para a análise, considerando ϕ fixo, poderá ser

$$\begin{aligned}
 Y_{ij} &\stackrel{\text{ind.}}{\sim} \text{Beta}(\mu_{ij}, \phi), \quad i = 1, \dots, I, \quad j = 1, \dots, J, \\
 \mu_{ij} &= g^{-1}(\eta_{ij}), \\
 \eta_{ij} &= m + \tau_i,
 \end{aligned} \tag{58}$$

em que m é uma constante e τ_i , $i = 1, \dots, I$, é o efeito do i -ésimo tratamento. No entanto, devido a problemas de singularidade, τ_i não é estimável mas $m + \tau_i$ é e assim, geralmente adotamos a parametrização $m + \tau_i = \beta_i$, $i = 1, \dots, I$. Logo, fazendo $m + \tau_i = \beta_i$, teremos que o modelo (58)

$$\begin{aligned}
 Y_{ij} &\sim \text{Beta}(\mu_{ij}, \phi), \quad i = 1, \dots, I, \quad j = 1, \dots, J, \\
 \mu_{ij} &= g^{-1}(\eta_{ij}), \\
 \eta_{ij} &= \beta_i
 \end{aligned} \tag{59}$$

e desta forma, o valor esperado da variável resposta para o i -ésimo tratamento será $E(Y_{ij}) = \mu_{ij} = g^{-1}(\beta_i)$, $i = 1, \dots, I$. Este modelo, na forma matricial utilizando-se a notação apresentada em (57), será, portanto,

$$\begin{aligned}
 Y_{ij} &\stackrel{\text{ind.}}{\sim} \text{Beta}(\mu_{ij}, \phi), \quad i = 1, \dots, I, \quad j = 1, \dots, J, \\
 \boldsymbol{\mu} &= g^{-1}(\boldsymbol{\eta}), \\
 \boldsymbol{\eta} &= \mathbf{X}\boldsymbol{\beta},
 \end{aligned} \tag{60}$$

em que: $\boldsymbol{\beta}$ é o vetor de parâmetros desconhecidos a serem estimados, com comprimento I ; $\mathbf{X} = \mathbf{I}_I \otimes \mathbf{1}_J$ é a matriz conhecida do delineamento, com dimensões $IJ \times I$; \mathbf{I}_I é a matriz identidade com dimensões $I \times I$; $\mathbf{1}_J$ é o vetor de uns, com comprimento J ; $\boldsymbol{\eta}$ é o vetor de preditores lineares, com comprimento IJ e $\boldsymbol{\mu}$ é o vetor de valores esperados de \mathbf{Y} , com comprimento IJ , ou seja,

$$\mathbf{Y} = \begin{pmatrix} Y_{11} \\ \vdots \\ Y_{1J} \\ Y_{21} \\ \vdots \\ Y_{2J} \\ \vdots \\ Y_{I1} \\ \vdots \\ Y_{IJ} \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_{11} \\ \vdots \\ y_{1J} \\ y_{21} \\ \vdots \\ y_{2J} \\ \vdots \\ y_{I1} \\ \vdots \\ y_{IJ} \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_{11} \\ \vdots \\ \mu_{1J} \\ \mu_{21} \\ \vdots \\ \mu_{2J} \\ \vdots \\ \mu_{I1} \\ \vdots \\ \mu_{IJ} \end{pmatrix}, \quad \mathbf{X} = \mathbf{I}_I \otimes \mathbf{1}_J = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_I \end{pmatrix},$$

$$\boldsymbol{\eta} = \begin{pmatrix} \eta_{11} \\ \vdots \\ \eta_{1J} \\ \eta_{21} \\ \vdots \\ \eta_{2J} \\ \vdots \\ \eta_{I1} \\ \vdots \\ \eta_{IJ} \end{pmatrix} \quad \text{e} \quad \mathbf{X}\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_2 \\ \vdots \\ \beta_I \\ \vdots \\ \beta_I \end{pmatrix}.$$

Consideremos, agora, que haja dados faltantes ou que dois ou mais tratamentos tenham números diferentes de repetições, como ilustra a Tabela 4.

Neste caso, o modelo continua sendo o apresentado em (60), porém com \mathbf{Y} , \mathbf{y} , $\boldsymbol{\mu}$,

Tabela 4. Observações da variável Y num experimento inteiramente ao acaso com I tratamentos e J_i , $i = 1, \dots, I$ repetições para o i -ésimo tratamento.

Tratamento	Repetição			
	1	2	\dots	J_i
1	y_{11}	y_{12}	\dots	y_{1J_1}
2	y_{21}	y_{22}	\dots	y_{2J_2}
\vdots	\vdots	\vdots	\ddots	\vdots
I	y_{I1}	y_{I2}	\dots	y_{IJ_I}

\mathbf{X} , $\boldsymbol{\beta}$, $\boldsymbol{\eta}$ e $\mathbf{X}\boldsymbol{\beta}$ dados, respectivamente, por

$$\mathbf{Y} = \begin{pmatrix} Y_{11} \\ \vdots \\ Y_{1J_1} \\ Y_{21} \\ \vdots \\ Y_{2J_2} \\ \vdots \\ Y_{I1} \\ \vdots \\ Y_{IJ_I} \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_{11} \\ \vdots \\ y_{1J_1} \\ y_{21} \\ \vdots \\ y_{2J_2} \\ \vdots \\ y_{I1} \\ \vdots \\ y_{IJ_I} \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_{11} \\ \vdots \\ \mu_{1J_1} \\ \mu_{21} \\ \vdots \\ \mu_{2J_2} \\ \vdots \\ \mu_{I1} \\ \vdots \\ \mu_{IJ_I} \end{pmatrix}, \quad (61)$$

$$\mathbf{X} = \bigoplus_{i=1}^I \mathbf{1}_{J_i} = \mathbf{1}_{J_1} \oplus \dots \oplus \mathbf{1}_{J_I} = \text{diag} \{ \mathbf{1}_{J_1}, \dots, \mathbf{1}_{J_I} \},$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_I \end{pmatrix}, \quad \boldsymbol{\eta} = \begin{pmatrix} \eta_{11} \\ \vdots \\ \eta_{1J_1} \\ \eta_{21} \\ \vdots \\ \eta_{2J_2} \\ \vdots \\ \eta_{I1} \\ \vdots \\ \eta_{IJ_I} \end{pmatrix} \quad \text{e} \quad \mathbf{X}\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_2 \\ \vdots \\ \beta_I \\ \vdots \\ \beta_I \end{pmatrix},$$

em que $\mathbf{1}_{J_i}$ é um vetor de uns, com comprimento J_i , $i = 1, \dots, I$.

3.2.3 Modelo para dados provenientes de experimentos inteiramente ao acaso com um fator e ϕ variável

Do exposto até aqui, consideramos o parâmetro de precisão ϕ variável. Entretanto, pode haver situações em que a precisão varia de acordo com diferentes níveis de um fator, como, por exemplo, com diferentes tratamentos. Assim, de modo a deixar o modelo (60) mais geral, consideraremos que $Y_{ij} \stackrel{ind.}{\sim} Beta(\mu_{ij}, \phi_{ij})$, $i = 1, \dots, I$, $j = 1, \dots, J_i$, e o modelo estatístico será:

$$\begin{aligned} Y_{ij} &\stackrel{ind.}{\sim} Beta(\mu_{ij}, \phi_{ij}), \quad i = 1, \dots, I, \quad j = 1, \dots, J_i, \\ \mu_{ij} &= g_1^{-1}(\eta_{1ij}), \\ \eta_{1ij} &= \beta_i, \\ \phi_{ij} &= g_2^{-1}(\eta_{2ij}), \\ \eta_{2ij} &= \lambda_i. \end{aligned} \tag{62}$$

em que: $\mu_{ij} = g_1^{-1}(\eta_{1ij})$ é o valor esperado de Y_{ij} ; $\phi_{ij} = g_2^{-1}(\eta_{2ij})$ é o parâmetro de precisão de Y_{ij} ; $g_1(\cdot)$ e $g_2(\cdot)$ são funções de ligação; β_i e λ_i são parâmetros associados, respectivamente, ao valor esperado e à precisão de Y_{ij} e $\eta_{1ij} = \beta_i$ e $\eta_{2ij} = \lambda_i$ são preditores lineares associados, respectivamente, ao valor esperado e à precisão de Y_{ij} . Note que Cribari-Neto e Zeileis (2010) sugerem, para a precisão, as funções de ligação logarítmica $g_2(\cdot) = \log(\cdot)$ ou raiz quadrada $g_2(\cdot) = (\cdot)^{\frac{1}{2}}$. Na sua forma matricial o modelo (62) pode ser, então, escrito da seguinte maneira

$$\begin{aligned} Y_{ij} &\stackrel{ind.}{\sim} Beta(\mu_{ij}, \phi_{ij}), \quad i = 1, \dots, I, \quad j = 1, \dots, J_i, \\ \boldsymbol{\mu} &= g_1^{-1}(\boldsymbol{\eta}_1), \\ \boldsymbol{\eta}_1 &= \mathbf{X}_1 \boldsymbol{\beta}, \\ \boldsymbol{\phi} &= g_2^{-1}(\boldsymbol{\eta}_2), \\ \boldsymbol{\eta}_2 &= \mathbf{X}_2 \boldsymbol{\lambda}, \end{aligned} \tag{63}$$

em que: $\boldsymbol{\beta}$ e $\boldsymbol{\lambda}$ são vetores de parâmetros, com comprimento I_1 e I_2 ; \mathbf{X}_1 e \mathbf{X}_2 são matrizes com dimensões $\left(\sum_{i=1}^I J_i\right) \times I_1$ e $\left(\sum_{i=1}^I J_i\right) \times I_2$, geralmente formadas por combinações lineares das colunas da matriz $\mathbf{X} = \bigoplus_{i=1}^I \mathbf{1}_{J_i}$ conhecida do delineamento; $g_1(\cdot)$ e $g_2(\cdot)$ são funções de ligação; $\boldsymbol{\eta}_1$ e $\boldsymbol{\eta}_2$ são vetores de preditores lineares associados, respectivamente, ao valor esperado e à precisão, com comprimentos $\sum_{i=1}^I J_i$; $\boldsymbol{\mu}$ e $\boldsymbol{\phi}$ são vetores de parâmetros, com comprimento $\sum_{i=1}^I J_i$.

Note que, no caso mais geral, quando há um parâmetro de escala e um de precisão

associado a cada tratamento, teremos que $I_1 = I_2 = I$ e $\mathbf{X}_1 = \mathbf{X}_2 = \mathbf{X} = \bigoplus_{i=1}^I \mathbf{1}_{J_i}$, ou seja,

$$\mathbf{Y} = \begin{pmatrix} Y_{11} \\ \vdots \\ Y_{1J_1} \\ Y_{21} \\ \vdots \\ Y_{2J_2} \\ \vdots \\ Y_{I1} \\ \vdots \\ Y_{IJ_I} \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_{11} \\ \vdots \\ y_{1J_1} \\ y_{21} \\ \vdots \\ y_{2J_2} \\ \vdots \\ y_{I1} \\ \vdots \\ y_{IJ_I} \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_{11} \\ \vdots \\ \mu_{1J_1} \\ \mu_{21} \\ \vdots \\ \mu_{2J_2} \\ \vdots \\ \mu_{I1} \\ \vdots \\ \mu_{IJ_I} \end{pmatrix}, \quad \mathbf{X}_1 = \mathbf{X}_2 = \mathbf{X} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_I \end{pmatrix}, \quad \boldsymbol{\eta}_1 = \begin{pmatrix} \eta_{111} \\ \vdots \\ \eta_{11J_1} \\ \eta_{121} \\ \vdots \\ \eta_{12J_2} \\ \vdots \\ \eta_{1I1} \\ \vdots \\ \eta_{1IJ_I} \end{pmatrix}, \quad \mathbf{X}_1 \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_2 \\ \vdots \\ \beta_I \\ \vdots \\ \beta_I \end{pmatrix}, \quad \boldsymbol{\phi} = \begin{pmatrix} \phi_{11} \\ \vdots \\ \phi_{1J_1} \\ \phi_{21} \\ \vdots \\ \phi_{2J_2} \\ \vdots \\ \phi_{I1} \\ \vdots \\ \phi_{IJ_I} \end{pmatrix}, \quad \boldsymbol{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_I \end{pmatrix},$$

$$\boldsymbol{\eta}_2 = \begin{pmatrix} \eta_{211} \\ \vdots \\ \eta_{21J_1} \\ \eta_{221} \\ \vdots \\ \eta_{22J_2} \\ \vdots \\ \eta_{2I1} \\ \vdots \\ \eta_{2IJ_I} \end{pmatrix} \quad \text{e} \quad \mathbf{X}_2 \boldsymbol{\lambda} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_2 \\ \vdots \\ \lambda_I \\ \vdots \\ \lambda_I \end{pmatrix}.$$

3.2.4 Modelo misto para dados provenientes de um experimento com dois fatores aninhados, ϕ variável e mais de um componente de variância

Apresentaremos, a seguir, o modelo para a análise de dados com distribuição beta provenientes de experimentos inteiramente ao acaso com I grupos ($i = 1, \dots, I$), J_i tratamento e K_{ij} repetições, com $i = 1, \dots, I$, $j = 1, \dots, J_i$ e $k = 1, \dots, K_{ij}$, conforme ilustra a Tabela 5.

Tabela 5. Observações da variável Y num experimento aninhado com I grupos, J_i , $i = 1, \dots, I$, cepas para o i -ésimo grupo e K_{ij} repetições para a j -ésima cepa do i -ésimo grupo, $j = 1, \dots, J_i$.

Grupo	Tratamento	Repetição			
		1	2	...	K_{ij}
1	1	y_{111}	y_{112}	...	$y_{11K_{11}}$
	2	y_{121}	y_{122}	...	$y_{12K_{12}}$
	\vdots	\vdots	\vdots	\ddots	\vdots
	J_1	y_{1J_11}	y_{1J_12}	...	$y_{1J_1K_{1J_1}}$
2	1	y_{211}	y_{212}	...	$y_{21K_{21}}$
	2	y_{221}	y_{222}	...	$y_{22K_{22}}$
	\vdots	\vdots	\vdots	\ddots	\vdots
	J_2	y_{2J_21}	y_{2J_22}	...	$y_{2J_2K_{2J_2}}$
\vdots	\vdots	\vdots	\vdots	\ddots	\vdots
I	1	y_{I11}	y_{I12}	...	$y_{I1K_{I1}}$
	2	y_{I21}	y_{I22}	...	$y_{I2K_{I2}}$
	\vdots	\vdots	\vdots	\ddots	\vdots
	J_I	y_{IJ_I1}	y_{IJ_I2}	...	$y_{IJ_IK_{IJ_I}}$

Seja Y_{ijk} uma variável com distribuição beta, com valor esperado μ_{ijk} e precisão ϕ_{ijk} e y_{ijk} uma observação de Y_{ijk} , com $i = 1, \dots, I$, $j = 1, \dots, J_i$ e $k = 1, \dots, K_{ij}$. Neste caso, dada a existência de J tratamentos por grupo, consideraremos que há, sobre o preditor linear η_{1ijk} , um efeito aleatório b_{ij} com distribuição normal com média 0 e variância σ_i^2 e teremos,

assim, o seguinte modelo:

$$\begin{aligned}
Y_{ijk} &\stackrel{ind.}{\sim} \text{Beta}(\mu_{ijk}, \phi_{ijk}), \quad i = 1, \dots, I, \quad j = 1, \dots, J_i, \quad k = 1, \dots, K_{ij} \\
\mu_{ijk} &= g_1^{-1}(\eta_{1ijk}), \\
\eta_{1ijk} &= \beta_i + b_{ij}, \\
\phi_{ijk} &= g_2^{-1}(\eta_{2ijk}), \\
\eta_{2ijk} &= \lambda_i, \\
b_{ij} &\stackrel{ind.}{\sim} N(0, \sigma_i^2),
\end{aligned} \tag{64}$$

em que: μ_{ijk} é o valor esperado de Y_{ijk} ; β_i é o efeito fixo associado ao i -ésimo grupo, $i = 1, \dots, I$; λ_i é o parâmetro associado a precisão; b_{ij} é o efeito aleatório associado ao i -ésimo grupo no j -ésimo tratamento, $j = 1, \dots, J_i$; $g_1(\cdot)$ e $g_2(\cdot)$ são funções de ligação e σ_i^2 é a variância do i -ésimo grupo da distribuição do efeito aleatório. Note que, em alguns experimentos o pesquisador possa desejar estudar o parâmetro de precisão das cepas que estão aninhada aos grupos, com isso o λ_i passar a ser λ_{ij} , ou que todas as observações possuem uma mesma precisão, sendo assim o λ_i passar a ser λ .

Reescrevendo o modelo (64) matricialmente, teremos que

$$\begin{aligned}
Y_{ijk} &\stackrel{ind.}{\sim} \text{Beta}(\mu_{ijk}, \phi_{ijk}), \quad i = 1, \dots, I, \quad j = 1, \dots, J_i, \quad k = 1, \dots, K_{ij} \\
\boldsymbol{\mu} &= g_1^{-1}(\boldsymbol{\eta}_1), \\
\boldsymbol{\eta}_1 &= \mathbf{X}_1 \boldsymbol{\beta} + \mathbf{Z} \mathbf{b}, \\
\boldsymbol{\phi} &= g_2^{-1}(\boldsymbol{\eta}_2), \\
\boldsymbol{\eta}_2 &= \mathbf{X}_2 \boldsymbol{\lambda}, \\
\mathbf{b} &\sim N_{(\sum_{i=1}^I J_i)}(\mathbf{0}, \boldsymbol{\Sigma}), \quad \boldsymbol{\Sigma} = \text{diag} \{ \mathbf{I}_{J_1} \sigma_1^2, \dots, \mathbf{I}_{J_I} \sigma_I^2 \},
\end{aligned} \tag{65}$$

em que: $\boldsymbol{\beta}$ é o vetor de parâmetros associado aos efeitos dos grupos, com comprimento I ; $\boldsymbol{\lambda}$ é o vetor de parâmetros associado a precisão, com comprimento I ; \mathbf{b} é o vetor de variáveis aleatórias, com comprimento $\sum_{i=1}^I J_i$; $\mathbf{0}$ é um vetor de zeros, com comprimento $\sum_{i=1}^I J_i$; $\mathbf{I}(\cdot)$ é uma matriz identidade com dimensão (\cdot) ; $\boldsymbol{\Sigma}$ é uma matriz de variâncias e covariâncias, com dimensão $\sum_{i=1}^I J_i \times \sum_{i=1}^I J_i$; $\mathbf{X}_1 = \mathbf{X}_2$ são matrizes de delineamento associado aos efeitos dos grupos, com dimensões $\left(\sum_{i=1}^I \sum_{j=1}^{J_i} K_{ij} \right) \times I_1$ e $\left(\sum_{i=1}^I \sum_{j=1}^{J_i} K_{ij} \right) \times I_2$; \mathbf{Z} é a matriz de delineamento associado aos efeitos das cepas, com dimensões $\left(\sum_{i=1}^I \sum_{j=1}^{J_i} K_{ij} \right) \times \sum_{i=1}^I J_i$; $\boldsymbol{\eta}_1$ e $\boldsymbol{\eta}_2$ são vetores de preditores lineares, com comprimento $\sum_{i=1}^I \sum_{j=1}^{J_i} K_{ij}$; e $\boldsymbol{\mu}$ e $\boldsymbol{\phi}$ são vetores

de parâmetros, com comprimento $\sum_{i=1}^I \sum_{j=1}^{J_i} K_{ij}$, ou seja,

$$\mathbf{y} = \begin{pmatrix} y_{111} \\ \vdots \\ y_{11K_{11}} \\ y_{121} \\ \vdots \\ y_{12K_{12}} \\ \vdots \\ y_{IJ_11} \\ \vdots \\ y_{IJ_1K_{IJ_1}} \end{pmatrix}, \quad \mathbf{Y} = \begin{pmatrix} Y_{111} \\ \vdots \\ Y_{11K_{11}} \\ Y_{121} \\ \vdots \\ Y_{12K_{12}} \\ \vdots \\ Y_{IJ_11} \\ \vdots \\ Y_{IJ_1K_{IJ_1}} \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_{111} \\ \vdots \\ \mu_{11K_{11}} \\ \mu_{121} \\ \vdots \\ \mu_{12K_{12}} \\ \vdots \\ \mu_{IJ_11} \\ \vdots \\ \mu_{IJ_1K_{IJ_1}} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_I \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_{11} \\ \vdots \\ b_{1J_1} \\ b_{21} \\ \vdots \\ b_{2J_2} \\ \vdots \\ b_{I1} \\ \vdots \\ b_{IJ_1} \end{pmatrix},$$

$$\boldsymbol{\eta}_1 = \begin{pmatrix} \eta_{111} \\ \vdots \\ \eta_{11K_{11}} \\ \eta_{121} \\ \vdots \\ \eta_{12K_{12}} \\ \vdots \\ \eta_{IJ_11} \\ \vdots \\ \eta_{IJ_1K_{IJ_1}} \end{pmatrix}, \quad \boldsymbol{\eta}_2 = \begin{pmatrix} \eta_{111} \\ \vdots \\ \eta_{11K_{11}} \\ \eta_{121} \\ \vdots \\ \eta_{12K_{12}} \\ \vdots \\ \eta_{IJ_11} \\ \vdots \\ \eta_{IJ_1K_{IJ_1}} \end{pmatrix},$$

$$\mathbf{X}_1 = \mathbf{X}_2 = \bigoplus_{i=1}^I \mathbf{1}_{\sum_{j=1}^{J_i} K_{ij}} = \mathbf{1}_{\sum_{j=1}^{J_1} K_{1j}} \oplus \cdots \oplus \mathbf{1}_{\sum_{j=1}^{J_I} K_{Ij}} = \text{diag} \left\{ \mathbf{1}_{\sum_{j=1}^{J_1} K_{1j}}, \dots, \mathbf{1}_{\sum_{j=1}^{J_I} K_{Ij}} \right\}$$

e

$$\mathbf{Z} = \bigoplus_{i=1}^I \bigoplus_{j=1}^{J_i} \mathbf{1}_{K_{ij}} = \text{diag} \left\{ \bigoplus_{j=1}^{J_1} \mathbf{1}_{K_{1j}}, \dots, \bigoplus_{j=1}^{J_I} \mathbf{1}_{K_{Ij}} \right\},$$

em que $\mathbf{1}_{(\cdot)}$ é um vetor de uns, com comprimento (\cdot) .

3.2.5 Modelo misto para dados provenientes de um experimento com dois fatores aninhados, ϕ variável e um componente de variância

Considerando um caso particular do modelo (64), em que existe apenas um componente de variância do efeito aleatório, o modelo será

$$\begin{aligned}
Y_{ijk} &\stackrel{ind.}{\sim} Beta(\mu_{ijk}, \phi_{ijk}), \quad i = 1, \dots, I, \quad j = 1, \dots, J_i, \quad k = 1, \dots, K_{ij} \\
\mu_{ijk} &= g_1^{-1}(\eta_{1ijk}), \\
\eta_{1ijk} &= \beta_i + b_{ij}, \\
\phi_{ijk} &= g_2^{-1}(\eta_{2ijk}), \\
\eta_{2ijk} &= \lambda_i, \\
b_{ij} &\stackrel{ind.}{\sim} N(0, \sigma_b^2),
\end{aligned} \tag{66}$$

em que: μ_{ijk} é o valor esperado de Y_{ijk} ; β_i é o efeito fixo associado ao i -ésimo grupo, $i = 1, \dots, I$; λ_i é o parâmetro associado a precisão; b_{ij} é o efeito aleatório associado ao i -ésimo grupo no j -ésimo tratamento, $j = 1, \dots, J_i$; $g_1(\cdot)$ e $g_2(\cdot)$ são funções de ligação e σ_b^2 é a variância da distribuição do efeito aleatório. O modelo (66) matricialmente será denotado por

$$\begin{aligned}
Y_{ijk} &\stackrel{ind.}{\sim} Beta(\mu_{ijk}, \phi_{ijk}), \quad i = 1, \dots, I, \quad j = 1, \dots, J_i, \quad k = 1, \dots, K_{ij} \\
\boldsymbol{\mu} &= g_1^{-1}(\boldsymbol{\eta}_1), \\
\boldsymbol{\eta}_1 &= \mathbf{X}_1 \boldsymbol{\beta} + \mathbf{Z} \mathbf{b}, \\
\boldsymbol{\phi} &= g_2^{-1}(\boldsymbol{\eta}_2), \\
\boldsymbol{\eta}_2 &= \mathbf{X}_2 \boldsymbol{\lambda}, \\
\mathbf{b} &\sim N_{(\sum_{i=1}^I J_i)} \left(\mathbf{0}, \mathbf{I}_{(\sum_{i=1}^I J_i)} \sigma_b^2 \right),
\end{aligned} \tag{67}$$

em que: $\boldsymbol{\beta}$ é o vetor de parâmetros associado aos efeitos dos grupos, com comprimento I ; $\boldsymbol{\lambda}$ é o vetor de parâmetros associado a precisão, com comprimento I ; \mathbf{b} é o vetor de variáveis aleatórias, com comprimento $\sum_{i=1}^I J_i$; $\mathbf{0}$ é um vetor de zeros, com comprimento $\sum_{i=1}^I J_i$; $\mathbf{I}_{(\cdot)}$ é uma matriz identidade com dimensão (\cdot) ; \mathbf{X}_1 e \mathbf{X}_2 são matrizes de delineamento associado aos efeitos dos grupos, com dimensões $\left(\sum_{i=1}^I \sum_{j=1}^{J_i} K_{ij} \right) \times I_1$ e $\left(\sum_{i=1}^I \sum_{j=1}^{J_i} K_{ij} \right) \times I_2$; \mathbf{Z} é a matriz de delineamento associado aos efeitos das cepas, com dimensões $\left(\sum_{i=1}^I \sum_{j=1}^{J_i} K_{ij} \right) \times \sum_{i=1}^I J_i$; $\boldsymbol{\eta}_1$ e $\boldsymbol{\eta}_2$ são vetores de preditores lineares, com comprimento $\sum_{i=1}^I \sum_{j=1}^{J_i} K_{ij}$; e $\boldsymbol{\mu}$ e $\boldsymbol{\phi}$ são vetores de parâmetros, com comprimento $\sum_{i=1}^I \sum_{j=1}^{J_i} K_{ij}$

3.2.6 Função de verossimilhança, distribuições a priori, a posteriori, marginais e condicionais completas para o modelo considerando dados de uma amostra aleatória simples de Y

Dado o modelo (52) apresentaremos, a seguir, o procedimento de estimação dos parâmetros do mesmo. Inicialmente, temos que a função de verossimilhança é dada por

$$\begin{aligned}
 f(\mathbf{y}|\beta, \phi) &= \prod_{u=1}^n f(y_u|\beta, \phi) = \prod_{u=1}^n \frac{y_u^{\mu_u \phi - 1} (1 - y_u)^{(1 - \mu_u) \phi - 1}}{B(\mu_u \phi, (1 - \mu_u) \phi)} \\
 &= \prod_{u=1}^n \frac{y_u^{g^{-1}(\eta_u) \phi - 1} (1 - y_u)^{(1 - g^{-1}(\eta_u)) \phi - 1}}{B(g^{-1}(\eta_u) \phi, (1 - g^{-1}(\eta_u)) \phi)} \\
 &= \prod_{u=1}^n \frac{y_u^{g^{-1}(\beta) \phi - 1} (1 - y_u)^{(1 - g^{-1}(\beta)) \phi - 1}}{B(g^{-1}(\beta) \phi, (1 - g^{-1}(\beta)) \phi)}. \tag{68}
 \end{aligned}$$

Como distribuição a priori para β , sugerimos a distribuição normal com média 0 e variância σ_β^2 , pois $\beta \in \mathbb{R}$ tem o mesmo suporte da distribuição e a escolha do valor de σ_β pode levar a prioris minimamente informativas (ou informativas, caso desejado). Temos, assim, que β tem uma função densidade de probabilidade dada por

$$f(\beta) = \frac{1}{\sqrt{2\pi\sigma_\beta^2}} \exp\left\{-\frac{\beta^2}{2\sigma_\beta^2}\right\} \propto \frac{1}{\sigma_\beta} \exp\left\{-\frac{\beta^2}{2\sigma_\beta^2}\right\}, \tag{69}$$

com $\beta \in \mathbb{R}$ e $\sigma_\beta > 0$. Por outro lado, para ϕ , sugerimos $\phi \sim \text{LogNormal}(0, \sigma_\phi^2)$, pois o suporte dessa distribuição é o mesmo do parâmetro, e a escolha do valor de σ_ϕ contribuirá para gerar prioris minimamente informativas (ou informativas, caso desejado), ou seja, ϕ terá função densidade de probabilidade dada por

$$f(\phi) = \frac{1}{\phi \sqrt{2\pi\sigma_\phi^2}} \exp\left\{-\frac{(\log \phi)^2}{2\sigma_\phi^2}\right\} \propto \frac{1}{\phi \sigma_\phi} \exp\left\{-\frac{(\log \phi)^2}{2\sigma_\phi^2}\right\}, \tag{70}$$

com $\phi > 0$ e $\sigma_\phi > 0$.

Supondo, agora, que β e ϕ sejam, a priori, independentes, então teremos que $f(\beta, \phi) = f(\beta)f(\phi)$ e dado o vetor \mathbf{y} de observações da variável \mathbf{Y} , a distribuição a posteriori conjunta de (β, ϕ) será, pelo teorema de Bayes,

$$\begin{aligned}
 f(\beta, \phi|\mathbf{y}) &= \frac{f(\mathbf{y}, \beta, \phi)}{f(\mathbf{y})} = \frac{f(\mathbf{y}|\beta, \phi)f(\beta, \phi)}{f(\mathbf{y})} = \frac{f(\mathbf{y}|\beta, \phi)f(\beta, \phi)}{\int \int f(\mathbf{y}|\beta, \phi)f(\beta, \phi) d\beta d\phi} \\
 &= \frac{f(\mathbf{y}|\beta, \phi)f(\beta)f(\phi)}{\int \int f(\mathbf{y}|\beta, \phi)f(\beta, \phi) d\beta d\phi} \tag{71}
 \end{aligned}$$

e como o divisor da equação (71) é uma constante, teremos que

$$f(\beta, \phi|\mathbf{y}) \propto f(\mathbf{y}|\beta, \phi)f(\beta)f(\phi). \tag{72}$$

As distribuições marginais a posteriori de β e ϕ , por sua vez, serão dadas por:

$$\begin{aligned} f(\beta|\mathbf{y}) &= \int f(\beta, \phi|\mathbf{y}) d\phi \propto \int f(\mathbf{y}|\beta, \phi)f(\beta)f(\phi) d\phi \\ &= f(\beta) \int f(\mathbf{y}|\beta, \phi)f(\phi) d\phi \end{aligned} \quad (73)$$

e

$$\begin{aligned} f(\phi|\mathbf{y}) &= \int f(\beta, \phi|\mathbf{y}) d\beta \propto \int f(\mathbf{y}|\beta, \phi)f(\beta)f(\phi) d\beta \\ &= f(\phi) \int f(\mathbf{y}|\beta, \phi)f(\beta) d\beta. \end{aligned} \quad (74)$$

No entanto, como, para o caso em questão, as densidades (73) e (74) não correspondem a distribuições conhecidas, devemos utilizar os métodos de Monte Carlo via Cadeias de Markov (*Markov Chain Monte Carlo - MCMC*) para obter amostras das distribuições a posteriori (LESAFFRE; LAWSON, 2012). Estes métodos, por sua vez, necessitam do conhecimento das distribuições condicionais a posteriori de cada um dos parâmetros, chamadas de distribuições condicionais completas. No caso em questão, as distribuições condicionais completas de β e ϕ , serão, respectivamente, dadas por:

$$f(\beta|\phi, \mathbf{y}) = \frac{f(\beta, \phi, \mathbf{y})}{f(\phi, \mathbf{y})} = \frac{f(\mathbf{y}|\beta, \phi)f(\beta, \phi)}{f(\phi, \mathbf{y})} = \frac{f(\mathbf{y}|\beta, \phi)f(\beta)f(\phi)}{f(\phi, \mathbf{y})} \propto f(\mathbf{y}|\beta, \phi)f(\beta) \quad (75)$$

e

$$f(\phi|\beta, \mathbf{y}) = \frac{f(\beta, \phi, \mathbf{y})}{f(\beta, \mathbf{y})} = \frac{f(\mathbf{y}|\beta, \phi)f(\beta, \phi)}{f(\beta, \mathbf{y})} = \frac{f(\mathbf{y}|\beta, \phi)f(\beta)f(\phi)}{f(\beta, \mathbf{y})} \propto f(\mathbf{y}|\beta, \phi)f(\phi). \quad (76)$$

3.2.7 Função de verossimilhança, distribuições a priori, a posteriori, marginais e condicionais completas para o modelo considerando dados provenientes de experimentos inteiramente ao acaso com um fator e ϕ não variável

Neste caso, a função de verossimilhança para o modelo correspondente (58) será

$$\begin{aligned} f(\mathbf{y}|\boldsymbol{\beta}, \phi) &= \prod_{i=1}^I \prod_{j=1}^J f(y_{ij}|\boldsymbol{\beta}, \phi) = \prod_{i=1}^I \prod_{j=1}^J \frac{y_{ij}^{g^{-1}(\eta_{ij})\phi-1} (1-y_{ij})^{(1-g^{-1}(\eta_{ij}))\phi-1}}{B(g^{-1}(\eta_{ij})\phi, (1-g^{-1}(\eta_{ij}))\phi)} \\ &= \prod_{i=1}^I \prod_{j=1}^J \frac{y_{ij}^{g^{-1}(\beta_i)\phi-1} (1-y_{ij})^{(1-g^{-1}(\beta_i))\phi-1}}{B(g^{-1}(\beta_i)\phi, (1-g^{-1}(\beta_i))\phi)}. \end{aligned} \quad (77)$$

Como distribuições a priori para ϕ e $\boldsymbol{\beta}$, consideremos, para ϕ , a distribuição descrita em (70), e para $\boldsymbol{\beta}$, a distribuição normal multivariada com vetor de valores esperados $\mathbf{0}$ e matriz de variâncias e covariâncias $\mathbf{I}_I \sigma_\beta^2$, ou seja, $\boldsymbol{\beta} \sim N_I(\mathbf{0}, \mathbf{I}_I \sigma_\beta^2)$, em que: $\mathbf{0}$ é o vetor de

zeros, com comprimento I ; \mathbf{I} é a matriz identidade com dimensões $I \times I$ e σ_β^2 é a variância de beta, ou seja, $\beta_i \stackrel{ind.}{\sim} N(0, \sigma_\beta^2)$, $i = 1, \dots, I$, com $\beta_i \in \mathbb{R}$ e $\sigma_\beta > 0$. Assim, teremos que

$$f(\beta_i) = \frac{1}{\sqrt{2\pi\sigma_\beta^2}} \exp\left\{-\frac{\beta_i^2}{2\sigma_\beta^2}\right\}$$

e

$$\begin{aligned} f(\boldsymbol{\beta}) &= \prod_{i=1}^I f(\beta_i) = \det(2\pi\mathbf{I}_I\sigma_\beta^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\boldsymbol{\beta}^T(\mathbf{I}_I\sigma_\beta^2)^{-1}\boldsymbol{\beta}\right\} = (2\pi\sigma_\beta^2)^{-\frac{I}{2}} \exp\left\{-\frac{1}{2\sigma_\beta^2}\boldsymbol{\beta}^T\boldsymbol{\beta}\right\} \\ &\propto \sigma_\beta^{-I} \exp\left\{-\frac{1}{2\sigma_\beta^2}\boldsymbol{\beta}^T\boldsymbol{\beta}\right\} = \sigma_\beta^{-I} \exp\left\{-\frac{1}{2\sigma_\beta^2}\sum_{i=1}^I\beta_i^2\right\}. \end{aligned} \quad (78)$$

Supondo, agora, que $\boldsymbol{\beta}$ e ϕ sejam, a priori, independentes, então teremos que $f(\boldsymbol{\beta}, \phi) = f(\boldsymbol{\beta})f(\phi)$ e dessa maneira, a distribuição à posteriori conjunta de $(\boldsymbol{\beta}^T, \phi)$ será, removendo a constante normalizadora, dada por:

$$f(\boldsymbol{\beta}, \phi|\mathbf{y}) = \frac{f(\boldsymbol{\beta}, \phi, \mathbf{y})}{f(\mathbf{y})} = \frac{f(\mathbf{y}|\boldsymbol{\beta}, \phi)f(\boldsymbol{\beta}, \phi)}{f(\mathbf{y})} \propto f(\mathbf{y}|\boldsymbol{\beta}, \phi)f(\boldsymbol{\beta}, \phi) = f(\mathbf{y}|\boldsymbol{\beta}, \phi)f(\boldsymbol{\beta})f(\phi),$$

em que $f(\mathbf{y}|\boldsymbol{\beta}, \phi)$, $f(\boldsymbol{\beta})$ e $f(\phi)$ são dadas, respectivamente, por (77), (78) e (70).

As distribuições marginais a posteriori de $\boldsymbol{\beta}$ e ϕ serão, deste modo, dadas. Respetivamente, por:

$$f(\boldsymbol{\beta}|\mathbf{y}) = \int f(\boldsymbol{\beta}, \phi|\mathbf{y}) d\phi \propto \int f(\mathbf{y}|\boldsymbol{\beta}, \phi)f(\boldsymbol{\beta})f(\phi) d\phi = f(\boldsymbol{\beta}) \int f(\mathbf{y}|\boldsymbol{\beta}, \phi)f(\phi) d\phi \quad (79)$$

e

$$\begin{aligned} f(\phi|\mathbf{y}) &= \int \dots \int f(\boldsymbol{\beta}, \phi|\mathbf{y}) d\boldsymbol{\beta} \propto \int \dots \int f(\mathbf{y}|\boldsymbol{\beta}, \phi)f(\boldsymbol{\beta})f(\phi) d\boldsymbol{\beta} \\ &= f(\phi) \int \dots \int f(\mathbf{y}|\boldsymbol{\beta}, \phi)f(\boldsymbol{\beta}) d\boldsymbol{\beta}. \end{aligned} \quad (80)$$

As distribuições condicionais completas de $\boldsymbol{\beta}$ e ϕ serão, por sua vez, dadas, respectivamente, por:

$$f(\boldsymbol{\beta}|\phi, \mathbf{y}) = \frac{f(\boldsymbol{\beta}, \phi, \mathbf{y})}{f(\phi, \mathbf{y})} = \frac{f(\mathbf{y}|\boldsymbol{\beta}, \phi)f(\boldsymbol{\beta}, \phi)}{f(\phi, \mathbf{y})} = \frac{f(\mathbf{y}|\boldsymbol{\beta}, \phi)f(\boldsymbol{\beta})f(\phi)}{f(\phi, \mathbf{y})} \propto f(\mathbf{y}|\boldsymbol{\beta}, \phi)f(\boldsymbol{\beta}) \quad (81)$$

e

$$f(\phi|\boldsymbol{\beta}, \mathbf{y}) = \frac{f(\boldsymbol{\beta}, \phi, \mathbf{y})}{f(\boldsymbol{\beta}, \mathbf{y})} = \frac{f(\mathbf{y}|\boldsymbol{\beta}, \phi)f(\boldsymbol{\beta}, \phi)}{f(\boldsymbol{\beta}, \mathbf{y})} = \frac{f(\mathbf{y}|\boldsymbol{\beta}, \phi)f(\boldsymbol{\beta})f(\phi)}{f(\boldsymbol{\beta}, \mathbf{y})} \propto f(\mathbf{y}|\boldsymbol{\beta}, \phi)f(\phi). \quad (82)$$

3.2.8 Função de verossimilhança, distribuições a priori, a posteriori, marginais e condicionais completas para o modelo considerando dados provenientes de experimentos inteiramente ao acaso com um fator e ϕ variável

Neste caso, temos que a função de verossimilhança para o modelo (62) é:

$$\begin{aligned}
 f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda}) &= \prod_{i=1}^I \prod_{j=1}^{J_i} f(y_{ij}|\boldsymbol{\beta}, \boldsymbol{\lambda}) = \prod_{i=1}^I \prod_{j=1}^{J_i} \frac{y_{ij}^{g_1^{-1}(\eta_{1ij})g_2^{-1}(\eta_{2ij})-1} (1-y_{ij})^{(1-g_1^{-1}(\eta_{1ij})g_2^{-1}(\eta_{2ij}))-1}}{B(g_1^{-1}(\eta_{1ij})g_2^{-1}(\eta_{2ij}), (1-g_1^{-1}(\eta_{1ij})g_2^{-1}(\eta_{2ij})))} \\
 &= \prod_{i=1}^I \prod_{j=1}^{J_i} \frac{y_{ij}^{g_1^{-1}(\beta_i)g_2^{-1}(\lambda_i)-1} (1-y_{ij})^{(1-g_1^{-1}(\beta_i)g_2^{-1}(\lambda_i))-1}}{B(g_1^{-1}(\beta_i)g_2^{-1}(\lambda_i), (1-g_1^{-1}(\beta_i)g_2^{-1}(\lambda_i)))}. \tag{83}
 \end{aligned}$$

As distribuições a priori que consideraremos para $\boldsymbol{\beta}$ e $\boldsymbol{\lambda}$ serão, por sua vez: $\boldsymbol{\beta} \sim N_I(\mathbf{0}, I_I \sigma_\beta^2)$ e $\boldsymbol{\lambda} \sim N_I(\mathbf{0}, I_I \sigma_\lambda^2)$, em que: $\mathbf{0}$ é o vetor de zeros, com comprimento I ; I_I é a matriz identidade com dimensões $I \times I$, e $I_I \sigma_\lambda^2$ é a matriz de variâncias e covariâncias de $\boldsymbol{\lambda}$, ou seja, $\lambda_i \stackrel{ind.}{\sim} N(0, \sigma_\lambda^2)$, com $\lambda_i \in \mathbb{R}$ e $\sigma_\lambda > 0$. Logo, teremos que

$$f(\lambda_i) = \frac{1}{\sqrt{2\pi\sigma_\lambda^2}} \exp\left\{-\frac{\lambda_i^2}{2\sigma_\lambda^2}\right\},$$

e

$$\begin{aligned}
 f(\boldsymbol{\lambda}) &= \prod_{i=1}^I f(\lambda_i) = \det(2\pi I_I \sigma_\lambda^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \boldsymbol{\lambda}^T (I_I \sigma_\lambda^2)^{-1} \boldsymbol{\beta}\right\} = (2\pi\sigma_\lambda^2)^{-\frac{I}{2}} \exp\left\{-\frac{1}{2\sigma_\lambda^2} \boldsymbol{\lambda}^T \boldsymbol{\lambda}\right\} \\
 &\propto \sigma_\lambda^{-I} \exp\left\{-\frac{1}{2\sigma_\lambda^2} \boldsymbol{\lambda}^T \boldsymbol{\lambda}\right\} = \sigma_\lambda^{-I} \exp\left\{-\frac{1}{2\sigma_\lambda^2} \sum_{i=1}^I \lambda_i^2\right\}. \tag{84}
 \end{aligned}$$

Assumindo, agora, que $\boldsymbol{\beta}$ e $\boldsymbol{\lambda}$ sejam, a priori, independentes, então teremos que $f(\boldsymbol{\beta}, \boldsymbol{\lambda}) = f(\boldsymbol{\beta})f(\boldsymbol{\lambda})$ e dado o vetor \mathbf{y} de observações da variável \mathbf{Y} , a distribuição a posteriori conjunta de $(\boldsymbol{\beta}^T, \boldsymbol{\lambda}^T)$ será:

$$f(\boldsymbol{\beta}, \boldsymbol{\lambda}|\mathbf{y}) \propto f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda})f(\boldsymbol{\beta}, \boldsymbol{\lambda}) = f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda})f(\boldsymbol{\beta})f(\boldsymbol{\lambda}), \tag{85}$$

em que $f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda})$, $f(\boldsymbol{\beta})$ e $f(\boldsymbol{\lambda})$ são dados por (83), (78) e (84).

Por outro lado as distribuições marginais a posteriori, para $\boldsymbol{\beta}$ e $\boldsymbol{\lambda}$ serão dadas, respectivamente, por:

$$\begin{aligned}
 f(\boldsymbol{\beta}|\mathbf{y}) &= \int \cdots \int f(\boldsymbol{\beta}, \boldsymbol{\lambda}|\mathbf{y}) d\boldsymbol{\lambda} \propto \int \cdots \int f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda})f(\boldsymbol{\beta}, \boldsymbol{\lambda}) d\boldsymbol{\lambda} \\
 &= \int \cdots \int f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda})f(\boldsymbol{\beta})f(\boldsymbol{\lambda}) d\boldsymbol{\lambda} \\
 &= f(\boldsymbol{\beta}) \int \cdots \int f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda})f(\boldsymbol{\lambda}) d\boldsymbol{\lambda} \tag{86}
 \end{aligned}$$

e

$$\begin{aligned}
f(\boldsymbol{\lambda}|\mathbf{y}) &= \int \cdots \int f(\boldsymbol{\beta}, \boldsymbol{\lambda}|\mathbf{y}) d\boldsymbol{\beta} \propto \int \cdots \int f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda})f(\boldsymbol{\beta}, \boldsymbol{\lambda}) d\boldsymbol{\beta} \\
&= \int \cdots \int f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda})f(\boldsymbol{\beta})f(\boldsymbol{\lambda}) d\boldsymbol{\beta} \\
&= f(\boldsymbol{\lambda}) \int \cdots \int f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda})f(\boldsymbol{\beta}) d\boldsymbol{\beta}
\end{aligned} \tag{87}$$

Finalmente, as distribuições condicionais completas para $\boldsymbol{\beta}$ e $\boldsymbol{\lambda}$ serão dadas, respectivamente, por:

$$\begin{aligned}
f(\boldsymbol{\beta}|\boldsymbol{\lambda}, \mathbf{y}) &= \frac{f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{y})}{f(\boldsymbol{\lambda}, \mathbf{y})} = \frac{f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda})f(\boldsymbol{\beta}, \boldsymbol{\lambda})}{f(\boldsymbol{\lambda}, \mathbf{y})} = \frac{f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda})f(\boldsymbol{\beta})f(\boldsymbol{\lambda})}{f(\boldsymbol{\lambda}, \mathbf{y})} \\
&\propto f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda})f(\boldsymbol{\beta})
\end{aligned} \tag{88}$$

e

$$\begin{aligned}
f(\boldsymbol{\lambda}|\boldsymbol{\beta}, \mathbf{y}) &= \frac{f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{y})}{f(\boldsymbol{\beta}, \mathbf{y})} = \frac{f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda})f(\boldsymbol{\beta}, \boldsymbol{\lambda})}{f(\boldsymbol{\beta}, \mathbf{y})} = \frac{f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda})f(\boldsymbol{\beta})f(\boldsymbol{\lambda})}{f(\boldsymbol{\beta}, \mathbf{y})} \\
&\propto f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\lambda})f(\boldsymbol{\lambda})
\end{aligned} \tag{89}$$

3.2.9 Função de verossimilhança, distribuições a priori, a posteriori, marginais e condicionais completas para o modelo misto considerando dados provenientes de um experimento com dois fatores aninhados, com ϕ variável e mais de um componente de variância

Para o modelo (64) a função de verossimilhança será:

$$\begin{aligned}
f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) &= \prod_{i=1}^I \prod_{j=1}^{J_i} \prod_{k=1}^{K_{ij}} f(y_{ijk}|\boldsymbol{\beta}, \boldsymbol{\lambda}, b_{ij}, \boldsymbol{\sigma})f(b_{ij}) \\
&= \prod_{i=1}^I \prod_{j=1}^{J_i} \prod_{k=1}^{K_{ij}} \frac{y_{ijk}^{g_1^{-1}(\eta_{1ijk})g_2^{-1}(\eta_{2ijk})-1} (1-y_{ijk})^{(1-g_1^{-1}(\eta_{1ijk}))g_2^{-1}(\eta_{2ijk})-1}}{B(g_1^{-1}(\eta_{1ijk})g_2^{-1}(\eta_{2ijk}), (1-g_1^{-1}(\eta_{1ijk}))g_2^{-1}(\eta_{2ijk}))} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left\{-\frac{b_{ij}^2}{2\sigma_i^2}\right\} \\
&= \prod_{i=1}^I \prod_{j=1}^{J_i} \prod_{k=1}^{K_{ij}} \frac{y_{ijk}^{g_1^{-1}(\beta_i+b_{ij})g_2^{-1}(\lambda_i)-1} (1-y_{ijk})^{(1-g_1^{-1}(\beta_i+b_{ij}))g_2^{-1}(\lambda_i)-1}}{B(g_1^{-1}(\beta_i+b_{ij})g_2^{-1}(\lambda_i), (1-g_1^{-1}(\beta_i+b_{ij}))g_2^{-1}(\lambda_i))} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left\{-\frac{b_{ij}^2}{2\sigma_i^2}\right\}
\end{aligned} \tag{90}$$

As distribuições a priori que consideraremos para $\boldsymbol{\beta}$ e $\boldsymbol{\lambda}$ serão, por sua vez: $\boldsymbol{\beta} \sim N_I(\mathbf{0}, \mathbf{I}_I\sigma_\beta^2)$ e $\boldsymbol{\lambda} \sim N_I(\mathbf{0}, \mathbf{I}_I\sigma_\lambda^2)$, cujas funções densidades de probabilidade são dadas, respectivamente, por (78) e (84). Para o parâmetro σ_i recomendamos distribuição LogNormal com média 0 e variância δ^2 , porque o suporte dessa distribuição é o mesmo do parâmetro, e a escolha do valor de δ contribuirá para gerar prioris minimamente informativas (ou informativas,

caso desejado). Assim teremos que

$$f(\boldsymbol{\sigma}) = \prod_{i=1}^I f(\sigma_i) = \prod_{i=1}^I \frac{1}{\sigma_i^2 \delta \sqrt{2\pi}} \exp \left\{ -\frac{(\log \sigma_i^2)^2}{2\delta^2} \right\} \propto \prod_{i=1}^I \frac{1}{\sigma_i^2 \delta} \exp \left\{ -\frac{(\log \sigma_i^2)^2}{2\delta^2} \right\}, \quad (91)$$

com $\sigma_i > 0$ e $\delta > 0$.

Supondo, agora, que as distribuições a priori de $\boldsymbol{\beta}$, $\boldsymbol{\lambda}$ e σ_i sejam independentes, ou seja, $f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_i) = f(\boldsymbol{\beta})f(\boldsymbol{\lambda})f(\sigma_i)$ teremos que a distribuição à posteriori conjunta de $(\boldsymbol{\beta}^T, \boldsymbol{\lambda}^T, \sigma_i)$ será

$$f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_i | \mathbf{y}, \mathbf{b}) \propto f(\mathbf{y}, \mathbf{b} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_i) f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_i) = f(\mathbf{y}, \mathbf{b} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_i) f(\boldsymbol{\beta}) f(\boldsymbol{\lambda}) f(\sigma_i), \quad (92)$$

em que $f(\mathbf{y}, \mathbf{b} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_i)$, $f(\boldsymbol{\beta})$, $f(\boldsymbol{\lambda})$ e $f(\sigma_i)$ são dadas por (90), (78), (84) e (91).

Consequentemente as distribuições marginais a posteriori de $\boldsymbol{\beta}$, $\boldsymbol{\lambda}$ e $\boldsymbol{\sigma}$ serão dadas, respectivamente, por:

$$\begin{aligned} f(\boldsymbol{\beta} | \mathbf{y}, \mathbf{b}) &= \int \cdots \int f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma} | \mathbf{y}, \mathbf{b}) d\boldsymbol{\lambda} d\boldsymbol{\sigma} \propto \int \cdots \int f(\mathbf{y}, \mathbf{b} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) d\boldsymbol{\lambda} d\boldsymbol{\sigma} \\ &= \int \cdots \int f(\mathbf{y}, \mathbf{b} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) f(\boldsymbol{\beta}) f(\boldsymbol{\lambda}) f(\boldsymbol{\sigma}) d\boldsymbol{\lambda} d\boldsymbol{\sigma} \\ &= f(\boldsymbol{\beta}) \int \cdots \int f(\mathbf{y}, \mathbf{b} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) f(\boldsymbol{\lambda}) f(\boldsymbol{\sigma}) d\boldsymbol{\lambda} d\boldsymbol{\sigma}, \end{aligned} \quad (93)$$

$$\begin{aligned} f(\boldsymbol{\lambda} | \mathbf{y}, \mathbf{b}) &= \int \cdots \int f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma} | \mathbf{y}, \mathbf{b}) d\boldsymbol{\beta} d\boldsymbol{\sigma} \propto \int \cdots \int f(\mathbf{y}, \mathbf{b} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) d\boldsymbol{\beta} d\boldsymbol{\sigma} \\ &= \int \cdots \int f(\mathbf{y}, \mathbf{b} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) f(\boldsymbol{\beta}) f(\boldsymbol{\lambda}) f(\boldsymbol{\sigma}) d\boldsymbol{\beta} d\boldsymbol{\sigma} \\ &= f(\boldsymbol{\lambda}) \int \cdots \int f(\mathbf{y}, \mathbf{b} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) f(\boldsymbol{\beta}) f(\boldsymbol{\sigma}) d\boldsymbol{\beta} d\boldsymbol{\sigma} \end{aligned} \quad (94)$$

e

$$\begin{aligned} f(\boldsymbol{\sigma} | \mathbf{y}, \mathbf{b}) &= \int \cdots \int f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma} | \mathbf{y}, \mathbf{b}) d\boldsymbol{\beta} d\boldsymbol{\lambda} \propto \int \cdots \int f(\mathbf{y}, \mathbf{b} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) d\boldsymbol{\beta} d\boldsymbol{\lambda} \\ &= \int \cdots \int f(\mathbf{y}, \mathbf{b} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) f(\boldsymbol{\beta}) f(\boldsymbol{\lambda}) f(\boldsymbol{\sigma}) d\boldsymbol{\beta} d\boldsymbol{\lambda} \\ &= f(\boldsymbol{\sigma}) \int \cdots \int f(\mathbf{y}, \mathbf{b} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) f(\boldsymbol{\beta}) f(\boldsymbol{\lambda}) d\boldsymbol{\beta} d\boldsymbol{\lambda}. \end{aligned} \quad (95)$$

As distribuições condicionais completas do vetor de parâmetros $\boldsymbol{\beta}$ e dos parâmetros $\boldsymbol{\lambda}$ e $\boldsymbol{\sigma}$ serão, por sua vez, dadas, respectivamente, por:

$$\begin{aligned} f(\boldsymbol{\beta} | \boldsymbol{\lambda}, \sigma_i, \mathbf{y}, \mathbf{b}) &= \frac{f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}, \mathbf{y}, \mathbf{b})}{f(\boldsymbol{\lambda}, \boldsymbol{\sigma}, \mathbf{y}, \mathbf{b})} = \frac{f(\mathbf{y}, \mathbf{b} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_i)}{f(\boldsymbol{\lambda}, \boldsymbol{\sigma}, \mathbf{y}, \mathbf{b})} \\ &= \frac{f(\mathbf{y}, \mathbf{b} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) f(\boldsymbol{\beta}) f(\boldsymbol{\lambda}) f(\sigma_i)}{f(\boldsymbol{\lambda}, \boldsymbol{\sigma}, \mathbf{y}, \mathbf{b})} \propto f(\mathbf{y}, \mathbf{b} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) f(\boldsymbol{\beta}), \end{aligned} \quad (96)$$

$$\begin{aligned}
f(\boldsymbol{\lambda}|\boldsymbol{\beta}, \boldsymbol{\sigma}, \mathbf{y}, \mathbf{b}) &= \frac{f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}, \mathbf{y}, \mathbf{b})}{f(\boldsymbol{\beta}, \boldsymbol{\sigma}, \mathbf{y}, \mathbf{b})} = \frac{f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma})f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}_i)}{f(\boldsymbol{\beta}, \boldsymbol{\sigma}_i, \mathbf{y}, \mathbf{b})} \\
&= \frac{f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma})f(\boldsymbol{\beta})f(\boldsymbol{\lambda})f(\boldsymbol{\sigma})}{f(\boldsymbol{\beta}, \boldsymbol{\sigma}, \mathbf{y}, \mathbf{b})} \propto f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma})f(\boldsymbol{\lambda}) \quad (97)
\end{aligned}$$

e

$$\begin{aligned}
f(\boldsymbol{\sigma}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{y}, \mathbf{b}) &= \frac{f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma}, \mathbf{y}, \mathbf{b})}{f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{y}, \mathbf{b})} = \frac{f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma})f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma})}{f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{y}, \mathbf{b})} \\
&= \frac{f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma})f(\boldsymbol{\beta})f(\boldsymbol{\lambda})f(\boldsymbol{\sigma})}{f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{y}, \mathbf{b})} \propto f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\sigma})f(\boldsymbol{\sigma}). \quad (98)
\end{aligned}$$

3.2.10 Função de verossimilhança, distribuições a priori, a posteriori, marginais e condicionais completas para o modelo misto considerando dados provenientes de um experimento com dois fatores aninhados, com ϕ variável e um componente de variância

Para o modelo (65) a função de verossimilhança será:

$$\begin{aligned}
f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) &= \prod_{i=1}^I \prod_{j=1}^{J_i} \prod_{k=1}^{K_{ij}} f(y_{ijk}|\boldsymbol{\beta}, \boldsymbol{\lambda}, b_{ij}, \sigma_b^2) f(b_{ij}) \quad (99) \\
&= \prod_{i=1}^I \prod_{j=1}^{J_i} \prod_{k=1}^{K_{ij}} \frac{y_{ijk}^{g_1^{-1}(\eta_{1ijk})g_2^{-1}(\eta_{2ijk})-1} (1-y_{ijk})^{(1-g_1^{-1}(\eta_{1ijk}))g_2^{-1}(\eta_{2ijk})-1}}{B(g_1^{-1}(\eta_{1ijk})g_2^{-1}(\eta_{2ijk}), (1-g_1^{-1}(\eta_{1ijk}))g_2^{-1}(\eta_{2ijk}))} \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp\left\{-\frac{b_{ij}^2}{2\sigma_b^2}\right\} \\
&= \prod_{i=1}^I \prod_{j=1}^{J_i} \prod_{k=1}^{K_{ij}} \frac{y_{ijk}^{g_1^{-1}(\beta_i+b_{ij})g_2^{-1}(\lambda_i)-1} (1-y_{ijk})^{(1-g_1^{-1}(\beta_i+b_{ij}))g_2^{-1}(\lambda_i)-1}}{B(g_1^{-1}(\beta_i+b_{ij})g_2^{-1}(\lambda_i), (1-g_1^{-1}(\beta_i+b_{ij}))g_2^{-1}(\lambda_i))} \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp\left\{-\frac{b_{ij}^2}{2\sigma_b^2}\right\}
\end{aligned}$$

As distribuições a priori que consideraremos para $\boldsymbol{\beta}$ e $\boldsymbol{\lambda}$ serão, por sua vez: $\boldsymbol{\beta} \sim N_I(\mathbf{0}, I_I\sigma_\beta^2)$ e $\boldsymbol{\lambda} \sim N_I(\mathbf{0}, I_I\sigma_\lambda^2)$, cujas funções densidades de probabilidade são dadas, respectivamente, por (78) e (84). Para o parâmetro σ_b recomendamos distribuição LogNormal com média 0 e variância δ^2 , porque o suporte dessa distribuição é o mesmo do parâmetro, e a escolha do valor de δ contribuirá para gerar prioris minimamente informativas (ou informativas, caso desejado). Assim teremos que

$$f(\sigma_b) = \frac{1}{\sigma_b^2 \delta \sqrt{2\pi}} \exp\left\{-\frac{(\log \sigma_b^2)^2}{2\delta^2}\right\} \propto \frac{1}{\sigma_b^2 \delta} \exp\left\{-\frac{(\log \sigma_b^2)^2}{2\delta^2}\right\}, \quad (100)$$

com $\sigma_b > 0$ e $\delta > 0$.

Supondo, agora, que as distribuições a priori de $\boldsymbol{\beta}$, $\boldsymbol{\lambda}$ e σ_b sejam independentes, ou seja, $f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) = f(\boldsymbol{\beta})f(\boldsymbol{\lambda})f(\sigma_b)$ teremos que a distribuição à posteriori conjunta de $(\boldsymbol{\beta}^T, \boldsymbol{\lambda}^T, \sigma_b)$ será

$$f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b|\mathbf{y}, \mathbf{b}) \propto f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b)f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) = f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b)f(\boldsymbol{\beta})f(\boldsymbol{\lambda})f(\sigma_b), \quad (101)$$

em que $f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b)$, $f(\boldsymbol{\beta})$, $f(\boldsymbol{\lambda})$ e $f(\sigma_b)$ são dadas por (99), (78), (84) e (100).

Consequentemente as distribuições marginais a posteriori de $\boldsymbol{\beta}$, $\boldsymbol{\lambda}$ e σ_b serão dadas, respectivamente, por:

$$\begin{aligned} f(\boldsymbol{\beta}|\mathbf{y}, \mathbf{b}) &= \int \int f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b|\mathbf{y}, \mathbf{b}) d\boldsymbol{\lambda} d\sigma_b \propto \int \int f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) d\boldsymbol{\lambda} d\sigma_b \\ &= \int \int f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\beta}) f(\boldsymbol{\lambda}) f(\sigma_b) d\boldsymbol{\lambda} d\sigma_b \\ &= f(\boldsymbol{\beta}) \int \int f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\lambda}) f(\sigma_b) d\boldsymbol{\lambda} d\sigma_b, \end{aligned} \quad (102)$$

$$\begin{aligned} f(\boldsymbol{\lambda}|\mathbf{y}, \mathbf{b}) &= \int \cdots \int f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b|\mathbf{y}, \mathbf{b}) d\boldsymbol{\beta} d\sigma_b \propto \int \cdots \int f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) d\boldsymbol{\beta} d\sigma_b \\ &= \int \cdots \int f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\beta}) f(\boldsymbol{\lambda}) f(\sigma_b) d\boldsymbol{\beta} d\sigma_b \\ &= f(\boldsymbol{\lambda}) \int \cdots \int f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\beta}) f(\sigma_b) d\boldsymbol{\beta} d\sigma_b \end{aligned} \quad (103)$$

e

$$\begin{aligned} f(\sigma_b|\mathbf{y}, \mathbf{b}) &= \int \cdots \int f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b|\mathbf{y}, \mathbf{b}) d\boldsymbol{\beta} d\boldsymbol{\lambda} \propto \int \cdots \int f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) d\boldsymbol{\beta} d\boldsymbol{\lambda} \\ &= \int \cdots \int f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\beta}) f(\boldsymbol{\lambda}) f(\sigma_b) d\boldsymbol{\beta} d\boldsymbol{\lambda} \\ &= f(\sigma_b) \int \cdots \int f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\beta}) f(\boldsymbol{\lambda}) d\boldsymbol{\beta} d\boldsymbol{\lambda}. \end{aligned} \quad (104)$$

As distribuições condicionais completas do vetor de parâmetros $\boldsymbol{\beta}$ e dos parâmetros $\boldsymbol{\lambda}$ e σ_b serão, por sua vez, dadas, respectivamente, por:

$$\begin{aligned} f(\boldsymbol{\beta}|\boldsymbol{\lambda}, \sigma_b, \mathbf{y}, \mathbf{b}) &= \frac{f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b, \mathbf{y}, \mathbf{b})}{f(\boldsymbol{\lambda}, \sigma_b, \mathbf{y}, \mathbf{b})} = \frac{f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b)}{f(\boldsymbol{\lambda}, \sigma_b, \mathbf{y}, \mathbf{b})} \\ &= \frac{f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\beta}) f(\boldsymbol{\lambda}) f(\sigma_b)}{f(\boldsymbol{\lambda}, \sigma_b, \mathbf{y}, \mathbf{b})} \propto f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\beta}), \end{aligned} \quad (105)$$

$$\begin{aligned} f(\boldsymbol{\lambda}|\boldsymbol{\beta}, \sigma_b, \mathbf{y}, \mathbf{b}) &= \frac{f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b, \mathbf{y}, \mathbf{b})}{f(\boldsymbol{\beta}, \sigma_b, \mathbf{y}, \mathbf{b})} = \frac{f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b)}{f(\boldsymbol{\beta}, \sigma_b, \mathbf{y}, \mathbf{b})} \\ &= \frac{f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\beta}) f(\boldsymbol{\lambda}) f(\sigma_b)}{f(\boldsymbol{\beta}, \sigma_b, \mathbf{y}, \mathbf{b})} \propto f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\lambda}) \end{aligned} \quad (106)$$

e

$$\begin{aligned} f(\sigma_b|\boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{y}, \mathbf{b}) &= \frac{f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b, \mathbf{y}, \mathbf{b})}{f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{y}, \mathbf{b})} = \frac{f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b)}{f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{y}, \mathbf{b})} \\ &= \frac{f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\boldsymbol{\beta}) f(\boldsymbol{\lambda}) f(\sigma_b)}{f(\boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{y}, \mathbf{b})} \propto f(\mathbf{y}, \mathbf{b}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \sigma_b) f(\sigma_b). \end{aligned} \quad (107)$$

3.2.11 Convergência e ajuste dos modelos

Para obtenção das estimativas a posteriori dos modelos propostos com algoritmo *HMC*, foi utilizado o pacote **rstan** da linguagem R (R Core Team, 2021; Stan Development Team, 2020).

As cadeias foram simuladas a partir dos modelos supracitados e, após, verificou-se o início de convergência dos algoritmos de *MCMC*. Para essa verificação entre as simulações das cadeias, foi utilizado a medida do fator de redução de escala potencial, que é descrito na equação (31). Além dessa medida, foi utilizados gráficos para análise da convergência e autocorrelação das estimativas dos parâmetros a posteriori de cada modelo.

Depois, foi calculada para cada modelo a medida de verificação do indícios de ajuste p_d , com sua estatística de teste $T(\cdot)$ sendo a média, que está descrito na equação (37). Com base nos valores p_d foram evidenciados quais modelos se ajustaram melhor aos dados.

Os modelos que tiveram indícios de bom ajuste aos dados foram comparados pelos critérios de informação *WAIC* e *LOOIC*, ou seja, o modelo que teve os menores valores nos critérios, é o que foi selecionado. Com o modelo escolhido, tabelas resumos foram elaboradas, contendo as estimativas a posteriori dos parâmetros, como por exemplo: média, mediana, variância, desvio padrão, intervalos de credibilidade, dentre outras.

4 RESULTADOS E DISCUSSÃO

Na Figura 4 podemos observar que o grupos de cepas 2 e 3, possuem maior variabilidade dentro do grupo, tanto entre as observações, quanto entre as cepas. No grupo 2 as plantas inoculadas com a cepa GO5_14 tem menor valor de severidade, essa cepa em geral causa menor severidade se comparado as demais de seu grupo, gerando assim uma grande variabilidade entre as observações presente grupo.

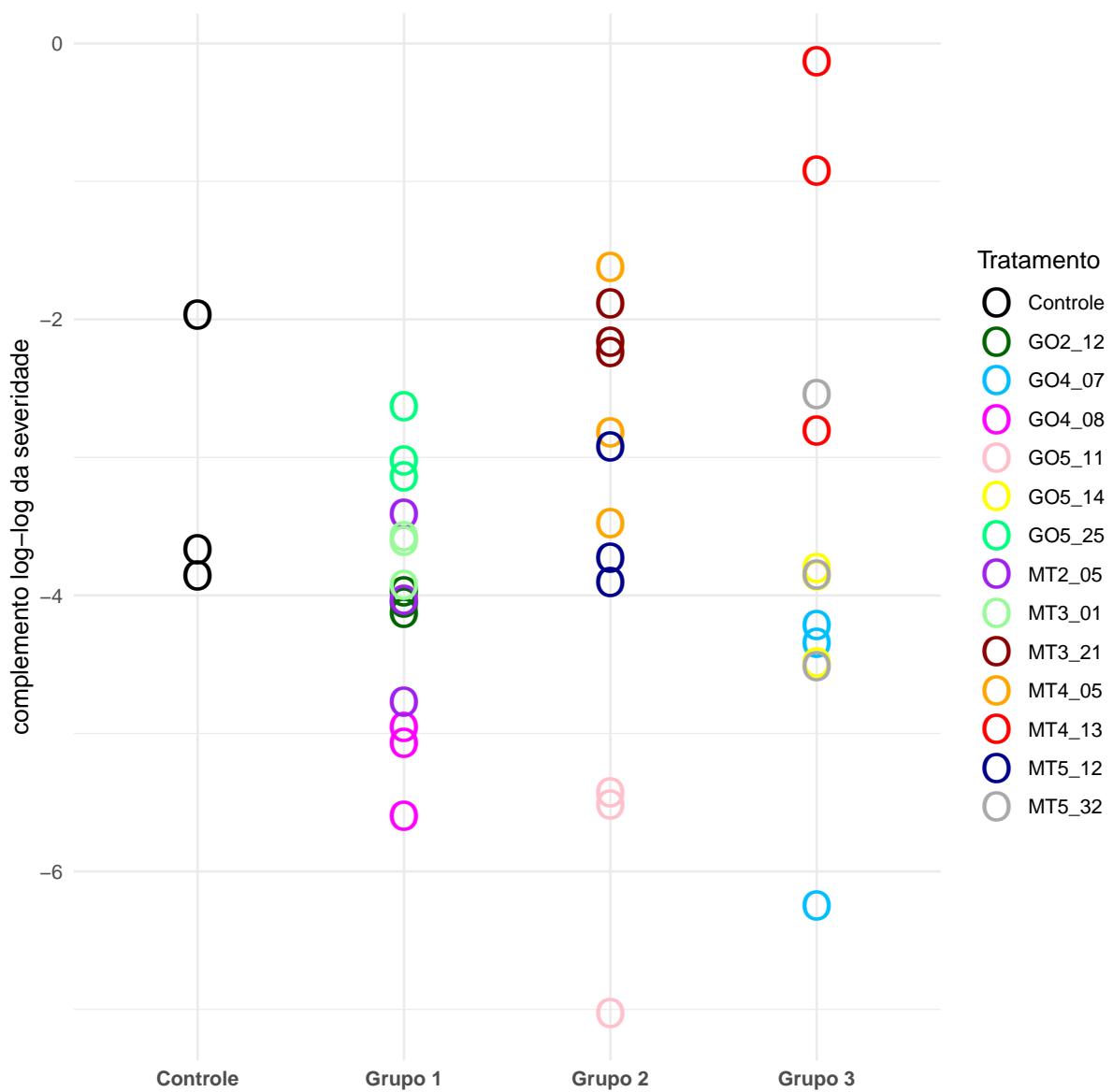


Figura 4. Gráfico de dispersão da severidade na escala da função de ligação complemento log-log em função dos grupos de cepas.

Inicialmente foi ajustado o modelo 1, utilizando a equação (60) e considerando os dados da Tabela 2, a severidade sendo a variável resposta y e os tratamentos sendo as cepas do fungo. O modelo 1 ficou descrito da seguinte forma:

$$\begin{aligned} Y_{ij} &\stackrel{ind.}{\sim} Beta(\mu_{ij}, \phi), \\ \boldsymbol{\mu} &= \mathbf{1} - \exp\{-\exp(\boldsymbol{\eta})\}, \\ \boldsymbol{\eta} &= \mathbf{X}\boldsymbol{\beta}, \\ \phi &= \exp(\lambda), \end{aligned} \tag{108}$$

em que: $i = 1, \dots, 14$; $j = 1, \dots, 3$; $\boldsymbol{\beta}$ é o vetor de parâmetros desconhecidos, com comprimento 14; $\mathbf{1}$ é o vetor de uns, com comprimento 14; $\mathbf{X} = \mathbf{I}_{14} \otimes \mathbf{1}_3$ é a matriz conhecida do delineamento, com dimensões 42×14 e $\boldsymbol{\mu}$ é o vetor de valores esperados, com comprimento 42.

O segundo modelo foi ajustado, agora considerando que o parâmetro ϕ é variável, utilizando a equação (62). Logo o modelo 2 é definido por:

$$\begin{aligned} Y_{ij} &\stackrel{ind.}{\sim} Beta(\mu_{ij}, \phi_{ij}), \\ \boldsymbol{\mu} &= \mathbf{1} - \exp\{-\exp(\boldsymbol{\eta}_1)\}, \\ \boldsymbol{\eta}_1 &= \mathbf{X}_1\boldsymbol{\beta}, \\ \phi &= \exp(\boldsymbol{\eta}_2), \\ \boldsymbol{\eta}_2 &= \mathbf{X}_2\boldsymbol{\lambda}, \end{aligned} \tag{109}$$

em que: $i = 1, \dots, 14$; $j = 1, \dots, 3$; $\boldsymbol{\beta}$ e $\boldsymbol{\lambda}$ são vetores de parâmetros, com comprimento 14; $\mathbf{1}$ é o vetor de uns, com comprimento 14; $\mathbf{X}_1 = \mathbf{X}_2 = \mathbf{I}_{14} \otimes \mathbf{1}_3$ são matrizes com dimensões 42×14 ; $\boldsymbol{\mu}$ e ϕ são vetores de parâmetros, com comprimento 42.

O modelo 1 e 2 foram ajustados sob uma abordagem da inferência bayesiana, para obtenção das estimativas a posteriori dos parâmetros de cada modelo foi fixado os seguintes parâmetros no algoritmo *HMC*: três cadeias, 3 mil iterações para cada cadeia, os 100 primeiros valores foram descartados e o espaçamento foi 5. No total, foram gerados um total de 1740 simulações para cada respectivo parâmetro a posteriori.

O indicativo de convergência entre as cadeias, foi verificado pela medida \hat{R} , que foi igual a 1 para todos os respectivos parâmetros de cada modelo. Note que as distribuições a priori para cada parâmetro dos modelos estão definidos na Tabela 6.

Tabela 6. Modelos e distribuições a priori dos parâmetros dos modelos considerando dados provenientes de experimentos inteiramente ao acaso com um fator

	Priori do $\boldsymbol{\beta}$	Priori do
Modelo 1	$N_{14}(\mathbf{0}, 100\mathbf{I}_{14})$	$\lambda \sim N(0, 100)$
Modelo 2	$N_{14}(\mathbf{0}, 100\mathbf{I}_{14})$	$\boldsymbol{\lambda} \sim N_{14}(\mathbf{0}, 100\mathbf{I}_{14})$

Os valores p_d para cada respectivo modelo são apresentados na Tabela 7. Podemos observar que o valor p_d mais próximo de 0,5, é o modelo 2, ou seja, é o modelo que melhor se ajusta aos dados. Os valores dos critérios de comparação $WAIC$ e $LOOIC$ para cada modelo, presentes na Tabela 7, também indica que o melhor modelo é o 2, em que existe um parâmetro de escala e precisão para cepa.

Tabela 7. Critério de informação amplamente aplicável ($WAIC$), Critério de informação leave-one-out ($LOOIC$) e o p-valor de verificação preditiva a posteriori (p_d). Menores valores dos critérios estão destacados em vermelho.

Modelo	p_d	$WAIC$	$LOOIC$
Modelo 1	0,69	-154,58	-151,96
Modelo 2	0,64	-210,32	-203,16

Por meio de ferramentas gráficas, fica evidenciado, que o modelo 2 tem uma boa capacidade preditiva se comparado ao modelo 1. A Figura 5 mostra em linhas azuis clara, a densidade das simulações da distribuição preditiva a posteriori para cada observação (y_{rep}) e a linha azul representando a densidade para os dados observados (y), onde pode-se notar na Figura 5 B) que a maioria dos valores y_{rep} estão se concentrando em torno da densidade das observações reais, indiciando um bom ajuste do modelo 2 aos dados.

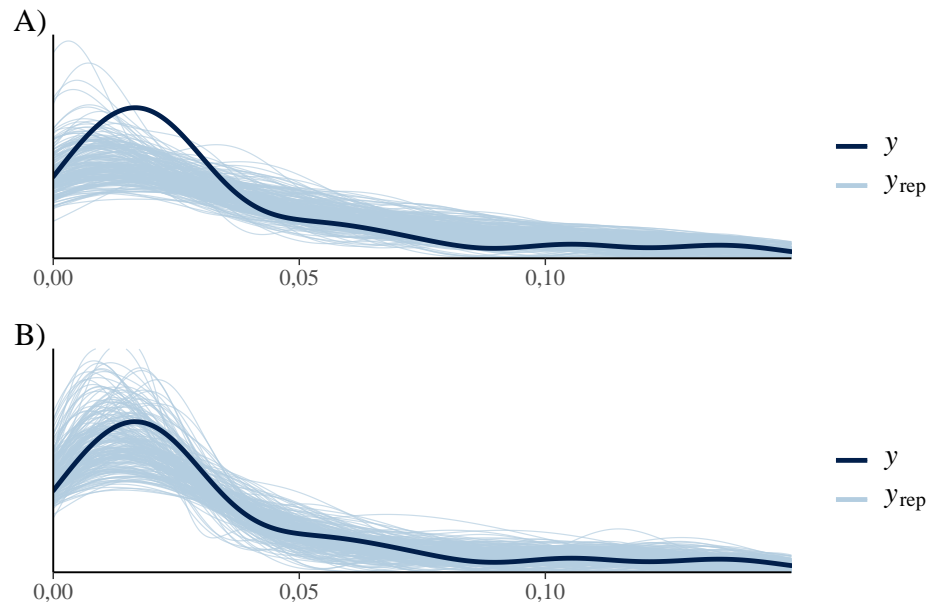


Figura 5. Gráfico das densidade estimada das simulações dos valores da distribuição preditiva a posteriori (y_{rep}) para o A) modelo 1, B) modelo 2 e y valores observados

Por meio da Figura 6 podemos concluir, que nos intervalos de credibilidade ao nível

de 95% o modelo 2 teve a maioria dos intervalos de credibilidade menores para cada respectiva observação se comparado ao modelo 1. Também nota-se que nos dois modelos as observações 31 a 33 não foram previstas bem pelos modelos, pois os valores observados ficaram quase ou fora do intervalo.

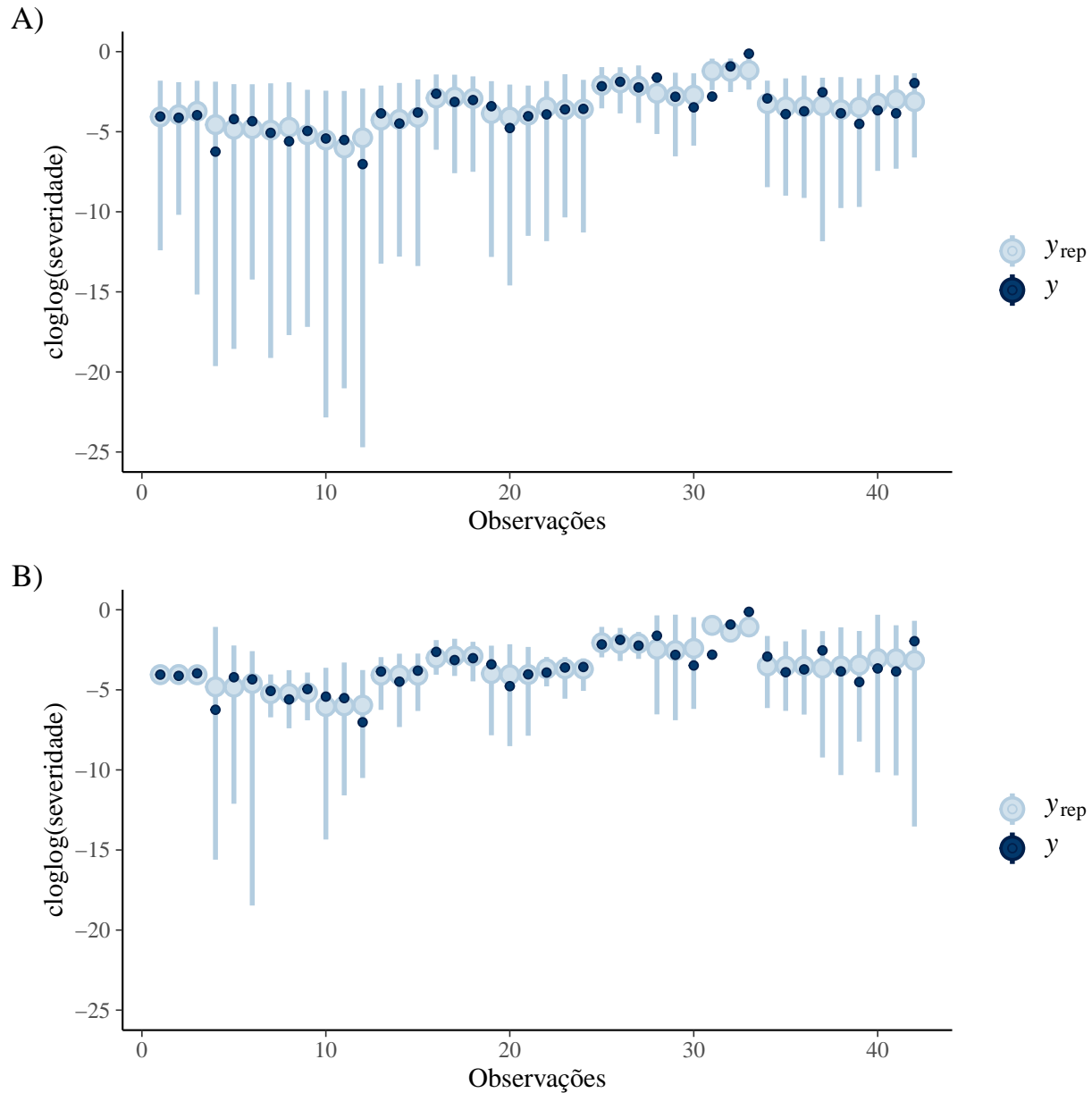


Figura 6. Gráfico dos intervalos de 95% das simulações dos valores da distribuição preditiva a posteriori (y_{rep}) para o A) modelo 1, B) modelo 2 e y valores observados, os valores estão na escala complemento log-log.

Na Tabela 8 são apresentados as estimativas a posteriori dos parâmetros do modelo

2, que foi o escolhido. Pode-se observar nas médias a posteriori que os tratamentos MT4_13, MT3_21 e MT4_05 causam o maior índice de severidade esperado. Já os tratamentos que tem menor variabilidade são GO2_12, GO4_08 e MT3_01, pois seu parâmetros associados a precisão foram os maiores.

Tabela 8. Estimativas a posteriori dos parâmetros do modelo 2. médias, desvios padrões, medianas e intervalos de 95% de credibilidade (IC 95%).

Tratamento	Parâmetro	Média	Desvio padrão	Mediana	IC 95%
Controle	β_1	-2,64	0,65	-2,68	(-3,89;-1,22)
GO2_12	β_2	-4,04	0,12	-4,05	(-4,25;-3,80)
GO4_08	β_3	-5,10	0,38	-5,14	(-5,73;-4,19)
GO5_25	β_4	-2,87	0,27	-2,89	(-3,33;-2,28)
MT2_05	β_5	-3,79	0,55	-3,87	(-4,67;-2,39)
MT3_01	β_6	-3,66	0,21	-3,68	(-4,01;-3,19)
GO5_11	β_7	-5,53	0,78	-5,69	(-6,67;-3,48)
MT3_21	β_8	-2,07	0,20	-2,08	(-2,45;-1,63)
MT4_05	β_9	-2,29	0,60	-2,33	(-3,41;-0,98)
MT5_12	β_{10}	-3,32	0,49	-3,39	(-4,16;-1,96)
GO4_07	β_{11}	-4,35	0,78	-4,46	(-5,60;-2,34)
GO5_14	β_{12}	-3,93	0,42	-3,98	(-4,63;-2,90)
MT4_13	β_{13}	-0,99	0,57	-0,97	(-2,20; 0,13)
MT5_32	β_{14}	-3,08	0,74	-3,17	(-4,37;-1,27)
Controle	λ_1	2,55	1,04	2,69	(0,15; 4,20)
GO2_12	λ_2	8,40	1,30	8,63	(5,36;10,37)
GO4_08	λ_3	6,88	1,26	7,06	(3,78; 8,68)
GO5_25	λ_4	5,07	1,13	5,23	(2,57; 6,77)
MT2_05	λ_5	4,39	1,15	4,60	(1,57; 6,21)
MT3_01	λ_6	6,51	1,22	6,69	(3,70; 8,38)
GO5_11	λ_7	5,74	1,38	6,01	(2,10; 7,62)
MT3_21	λ_8	4,98	1,13	5,12	(2,37; 6,80)
MT4_05	λ_9	2,40	1,03	2,56	(0,01; 4,07)
MT5_12	λ_{10}	4,24	1,19	4,42	(1,33; 6,00)
GO4_07	λ_{11}	4,25	1,28	4,49	(1,21; 6,18)
GO5_14	λ_{12}	5,51	1,22	5,67	(2,65; 7,40)
MT4_13	λ_{13}	0,88	0,88	0,98	(-1,05; 2,24)
MT5_32	λ_{14}	3,03	1,19	3,20	(0,10; 4,73)

Nas Figuras 7 e 8, podemos observar graficamente o que foi mostrado na Tabela 8 sobre os parâmetros do modelo 2. Fica evidenciado que alguns tratamentos possuem menores intervalos de credibilidade, tais como: β_2 , β_6 e β_8 .

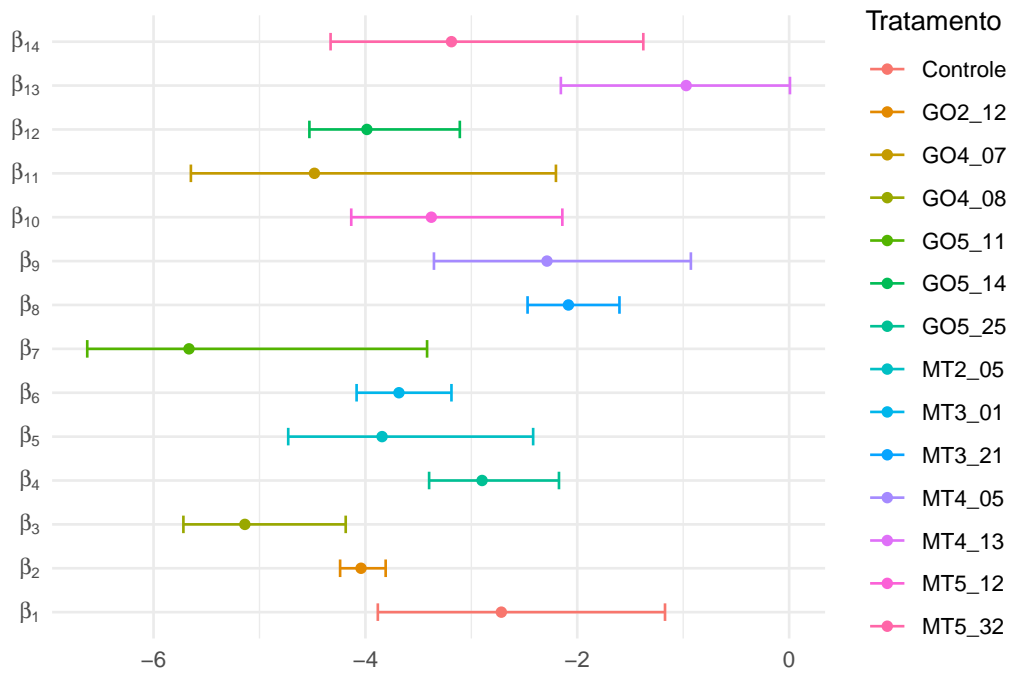


Figura 7. Intervalos de 95% de credibilidade com a mediana dos parâmetros $\beta_1, \dots, \beta_{14}$ do modelo 2

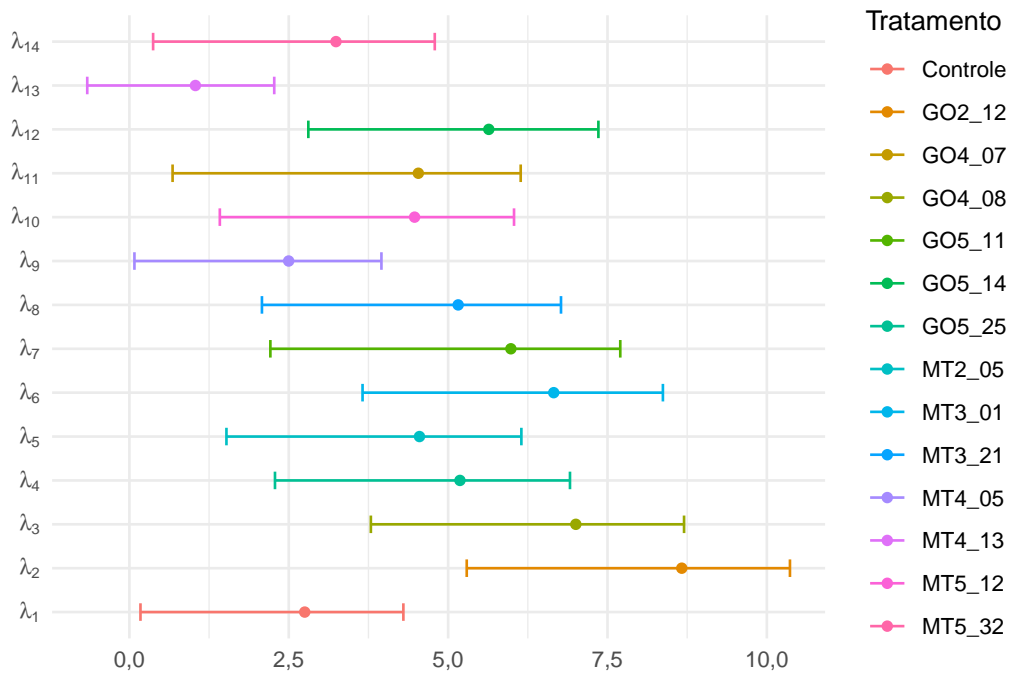


Figura 8. Intervalos de 95% de credibilidade com a mediana dos parâmetros $\lambda_1, \dots, \lambda_{14}$ do modelo 2

Podemos observa na Tabela 9 que os valores esperados das cepas do fungo que causam maior índice de severidade, são as cepas MT4_13 e a MT3_21. Já o tratamentos que obtiveram os menores índices de severidade esperado foram as cepas GO4-07, GO4-08, GO5-11 e GO5-14.

Tabela 9. Estimativas a posteriori dos valores esperados do modelo 2

Tratamento	Parâmetro	Média	IC 95%
Controle	μ_1	0,067	(0,020; 0,202)
GO2_12	μ_2	0,017	(0,014; 0,020)
GO4_08	μ_3	0,006	(0,003; 0,011)
GO5_25	μ_4	0,055	(0,033; 0,086)
MT2_05	μ_5	0,023	(0,009; 0,061)
MT3_01	μ_6	0,025	(0,017; 0,034)
GO5_11	μ_7	0,004	(0,001; 0,016)
MT3_21	μ_8	0,118	(0,081; 0,161)
MT4_05	μ_9	0,100	(0,034; 0,246)
MT5_12	μ_{10}	0,036	(0,016; 0,080)
GO4_07	μ_{11}	0,013	(0,004; 0,056)
GO5_14	μ_{12}	0,019	(0,011; 0,036)
MT4_13	μ_{13}	0,309	(0,109; 0,568)
MT5_32	μ_{14}	0,044	(0,013; 0,156)

As hipóteses que a pesquisadora tinha interesse, que foram os seguintes:

$$\begin{cases} H_0 : C = (G_1 + G_2 + G_3)/3 \\ H_1 : C \neq (G_1 + G_2 + G_3)/3, \end{cases} \quad (110)$$

$$\begin{cases} H_0 : G_1 = G_2 \\ H_1 : G_1 \neq G_2, \end{cases} \quad (111)$$

$$\begin{cases} H_0 : G_1 = G_3 \\ H_1 : G_1 \neq G_3, \end{cases} \quad (112)$$

$$\begin{cases} H_0 : G_1 = C \\ H_1 : G_1 \neq C, \end{cases} \quad (113)$$

$$\begin{cases} H_0 : G_2 = G_3 \\ H_1 : G_2 \neq G_3, \end{cases} \quad (114)$$

$$\begin{cases} H_0 : G_2 = C \\ H_1 : G_2 \neq C, \end{cases} \quad (115)$$

$$\begin{cases} H_0 : G_3 = C \\ H_1 : G_3 \neq C. \end{cases} \quad (116)$$

O próximo passo é testar as hipóteses mostradas anteriormente, que foram formuladas por contrastes dos parâmetros do modelo, mas primeiro foi formulado quais β compõem o controle e os grupos genéticos de cepas 1, 2 e 3, definidos por:

- $C = \beta_1$,
- $G_1 = (\beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6)/5$,
- $G_2 = (\beta_7 + \beta_8 + \beta_9 + \beta_{10})/4$,
- $G_3 = (\beta_{11} + \beta_{12} + \beta_{13} + \beta_{14})/4$.

e os contrastes de interesse definidos pela pesquisadora foram as seguintes:

- $C - (G_1 + G_2 + G_3)/3$;
- $G_1 - G_2$;
- $G_1 - G_3$;
- $G_1 - C$;
- $G_2 - G_3$;
- $G_2 - C$;
- $G_3 - C$.

Após testar os contrastes mostrados anteriormente, fica constatado na Tabela 10 que no intervalo de credibilidade ao nível de 95%, que G_1 difere do G_2 e G_3 , como as diferenças dos contraste 1 e 2 foram negativas, o grupo de cepa 1 causa menor índice de severidade na cultura da soja se comparado ao grupo 2 e 3. Já os demais contrastes não diferem entre si, pois incluem o valor zero no intervalo de credibilidade.

Tabela 10. Resultados dos contrastes dos parâmetros do modelo 2, mediana e intervalo de 95% de credibilidade (IC 95%).

Contraste	Média	Mediana	IC 95%
$C - (G_1 + G_2 + G_3)/3$	0,79	0,75	(-0,51; 2,34)
$G_1 - G_2$	-0,60	-0,58	(-1,31;-0,01)
$G_1 - G_3$	-0,79	-0,78	(-1,52;-0,15)
$G_1 - C$	-1,22	-1,18	(-2,79; 0,11)
$G_2 - G_3$	-0,19	-0,19	(-0,99; 0,65)
$G_2 - C$	-0,62	-0,58	(-2,24; 0,81)
$G_3 - C$	-0,43	-0,40	(-2,07; 0,99)

Visualmente na Figura 9 se observa na linha vermelha tracejada, que nos intervalos de credibilidade do grupo 1 versus o grupo 2 e 3, não incluem o valor zero, como foi descrito na Tabela 10.

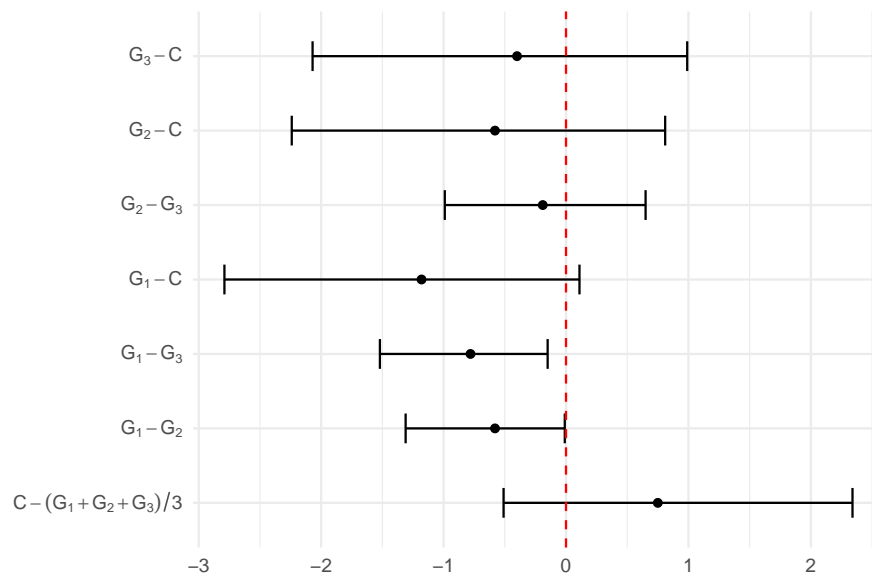


Figura 9. Intervalos de 95% de credibilidade com a mediana dos contrastes do modelo 2

A partir daqui foram ajustados 5 modelos, alguns com apenas efeitos fixos e outros com dição do afeito aleatório, os modelos utilizaram os dados da Tabela 2 considerando os dois fatores, grupos e cepas são covariáveis, onde existe um aninhamento de cepas aos grupos e a variável resposta é a severidade. Note que a partir daqui o parâmetro β_1 ainda se refere ao controle, mas os demais não, estão agora associado aos grupos de cepas, por exemplo o β_2 é referente ao grupo de cepas 1, o β_3 é referente ao grupo de cepas 2 e o β_4 referente ao grupo de cepas 3.

Os modelos nessa subseção, serão nomeados de M1, M2, M3, M4 e M5. Para criação deles foram utilizadas as equações (65) e (67), os mesmo são definidos por:

- Modelo M1:

$$\begin{aligned}
 Y_{ijk} &\stackrel{ind.}{\sim} Beta(\mu_{ijk}, \phi_{ijk}), \\
 \boldsymbol{\mu} &= \mathbf{1} - \exp\{-\exp(\boldsymbol{\eta}_1)\}, \\
 \boldsymbol{\eta}_1 &= \mathbf{X}_1\boldsymbol{\beta}, \\
 \phi &= \exp(\boldsymbol{\eta}_2), \\
 \boldsymbol{\eta}_2 &= \mathbf{X}_2\boldsymbol{\lambda},
 \end{aligned} \tag{117}$$

em que: $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)^T$ é o vetor de parâmetros associado aos efeitos dos grupos, com comprimento 4; $\mathbf{1}$ e um vetor de uns, com comprimento 14;

$\boldsymbol{\lambda} = (\lambda_{11}, \lambda_{21}, \lambda_{22}, \lambda_{23}, \lambda_{24}, \lambda_{25}, \lambda_{31}, \lambda_{32}, \lambda_{33}, \lambda_{34}, \lambda_{41}, \lambda_{42}, \lambda_{43}, \lambda_{44})^T$ é o vetor de parâmetros associado a precisão, com comprimento 14; $\mathbf{X}_1 = \text{diag}\{\mathbf{1}_3, \mathbf{1}_{15}, \mathbf{1}_{12}, \mathbf{1}_{12}\}$ é a matriz de delineamento associado aos efeitos dos grupos, com dimensões 42×4 e $\mathbf{X}_2 = \text{diag}\left\{\bigoplus_{j=1}^1 \mathbf{1}_3, \bigoplus_{j=1}^5 \mathbf{1}_3, \bigoplus_{j=1}^4 \mathbf{1}_3, \bigoplus_{j=1}^4 \mathbf{1}_3\right\}$ é a matriz de delineamento associado aos efeitos das cepas dentro dos grupos, com dimensão 42×14 .

- Modelo M2:

$$\begin{aligned}
 Y_{ijk} &\stackrel{ind.}{\sim} Beta(\mu_{ijk}, \phi), \\
 \boldsymbol{\mu} &= \mathbf{1} - \exp\{-\exp(\boldsymbol{\eta}_1)\}, \\
 \boldsymbol{\eta}_1 &= \mathbf{X}_1\boldsymbol{\beta} + \mathbf{Z}\mathbf{b}, \\
 \phi &= \exp(\boldsymbol{\eta}_2), \\
 \boldsymbol{\eta}_2 &= \lambda, \\
 \mathbf{b} &\sim N_{13}(\mathbf{0}, \boldsymbol{\Sigma}), \quad \boldsymbol{\Sigma} = \text{diag}\{I_5\sigma_2^2, I_4\sigma_3^2, I_4\sigma_4^2\},
 \end{aligned} \tag{118}$$

em que: $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)^T$ é o vetor de parâmetros associado aos efeitos dos grupos, com comprimento 4; λ é o parâmetro associado a precisão; $\mathbf{1}$ e um vetor de uns, com

comprimento 14; $\mathbf{0}$ e um vetor de zeros, com comprimento 13;

$\mathbf{b} = (b_{21}, b_{22}, b_{23}, b_{24}, b_{25}, b_{31}, b_{32}, b_{33}, b_{34}, b_{41}, b_{42}, b_{43}, b_{44})^T$ é o vetor de variáveis aleatórias, com comprimento 13; $\mathbf{X}_1 = \text{diag}\{\mathbf{1}_3, \mathbf{1}_{15}, \mathbf{1}_{12}, \mathbf{1}_{12}\}$ é a matriz de delineamento associado aos efeitos dos grupos, com dimensões 42×4 e

$\mathbf{Z} = \text{diag}\left\{\bigoplus_{j=1}^5 \mathbf{1}_3, \bigoplus_{j=1}^4 \mathbf{1}_3, \bigoplus_{j=1}^4 \mathbf{1}_3\right\}$ é a matriz de delineamento associado aos efeitos das cepas dentro dos grupos sem o efeito do controle, com dimensão 42×13 ; $\mathbf{\Sigma}$ é a matriz de variâncias e covariâncias do efeito aleatório, com dimensões 13×13 ; σ_2^2, σ_3^2 e σ_4^2 são os componentes de variâncias associado aos grupos de cepas, sem o efeito do controle.

- Modelo M3:

$$\begin{aligned}
 Y_{ijk} &\stackrel{ind.}{\sim} \text{Beta}(\mu_{ijk}, \phi_{ijk}), \\
 \mu &= \mathbf{1} - \exp\{-\exp(\boldsymbol{\eta}_1)\}, \\
 \boldsymbol{\eta}_1 &= \mathbf{X}_1\boldsymbol{\beta} + \mathbf{Z}\mathbf{b}, \\
 \phi &= \exp(\boldsymbol{\eta}_2), \\
 \boldsymbol{\eta}_2 &= \mathbf{X}_2\boldsymbol{\lambda}, \\
 \mathbf{b} &\sim N_{13}(\mathbf{0}, \mathbf{\Sigma}), \quad \mathbf{\Sigma} = \text{diag}\{I_5\sigma_2^2, I_4\sigma_3^2, I_4\sigma_4^2\}, \tag{119}
 \end{aligned}$$

em que: $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)^T$ é o vetor de parâmetros associado aos efeitos dos grupos, com comprimento 4; $\boldsymbol{\lambda} = (\lambda_{11}, \lambda_{21}, \lambda_{22}, \lambda_{23}, \lambda_{24}, \lambda_{25}, \lambda_{31}, \lambda_{32}, \lambda_{33}, \lambda_{34}, \lambda_{41}, \lambda_{42}, \lambda_{43}, \lambda_{44})^T$ é o vetor de parâmetros associado a precisão, com comprimento 14; $\mathbf{1}$ e um vetor de uns, com comprimento 14; $\mathbf{0}$ e um vetor de zeros, com comprimento 13; $\mathbf{b} = (b_{21}, b_{22}, b_{23}, b_{24}, b_{25}, b_{31}, b_{32}, b_{33}, b_{34}, b_{41}, b_{42}, b_{43}, b_{44})^T$ é o vetor de variáveis aleatórias, com comprimento 13; $\mathbf{X}_1 = \text{diag}\{\mathbf{1}_3, \mathbf{1}_{15}, \mathbf{1}_{12}, \mathbf{1}_{12}\}$ é a matriz de delineamento associado aos efeitos dos grupos, com dimensões 42×4 ;

$\mathbf{X}_2 = \text{diag}\left\{\bigoplus_{j=1}^1 \mathbf{1}_3, \bigoplus_{j=1}^5 \mathbf{1}_3, \bigoplus_{j=1}^4 \mathbf{1}_3, \bigoplus_{j=1}^4 \mathbf{1}_3\right\}$ é a matriz associado aos efeitos das cepas dentro dos grupos, com dimensão 42×14 ;

$\mathbf{Z} = \text{diag}\left\{\bigoplus_{j=1}^5 \mathbf{1}_3, \bigoplus_{j=1}^4 \mathbf{1}_3, \bigoplus_{j=1}^4 \mathbf{1}_3\right\}$ é a matriz de delineamento associado aos efeitos das cepas dentro dos grupos sem o efeito do controle, com dimensão 42×13 ; $\mathbf{\Sigma}$ é a matriz de variâncias e covariâncias do efeito aleatório, com dimensões 13×13 ; σ_2^2, σ_3^2 e σ_4^2 são os componentes de variâncias associado aos grupos de cepas, sem o efeito do controle.

- Modelo M4:

$$\begin{aligned}
Y_{ijk} &\stackrel{ind.}{\sim} \text{Beta}(\mu_{ijk}, \phi), \\
\boldsymbol{\mu} &= \mathbf{1} - \exp\{-\exp(\boldsymbol{\eta}_1)\}, \\
\boldsymbol{\eta}_1 &= \mathbf{X}_1\boldsymbol{\beta} + \mathbf{Z}\mathbf{b}, \\
\phi &= \exp(\eta_2), \\
\eta_2 &= \lambda, \\
\mathbf{b} &\sim N_{14}(\mathbf{0}, \mathbf{I}_{14}\sigma_b^2),
\end{aligned} \tag{120}$$

em que: $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)^T$ é o vetor de parâmetros associado aos efeitos dos grupos, com comprimento 4; λ é o parâmetro associado a precisão; $\mathbf{1}$ e um vetor de uns, com comprimento 14; $\mathbf{0}$ e um vetor de zeros, com comprimento 14;

$\mathbf{b} = (b_{11}, b_{21}, b_{22}, b_{23}, b_{24}, b_{25}, b_{31}, b_{32}, b_{33}, b_{34}, b_{41}, b_{42}, b_{43}, b_{44})^T$ é o vetor de variáveis aleatórias, com comprimento 14; $\mathbf{X}_1 = \text{diag}\{\mathbf{1}_3, \mathbf{1}_{15}, \mathbf{1}_{12}, \mathbf{1}_{12}\}$ é a matriz de delineamento associado aos efeitos dos grupos, com dimensões 42×4 ;

$\mathbf{Z} = \text{diag}\left\{\bigoplus_{j=1}^1 \mathbf{1}_3, \bigoplus_{j=1}^5 \mathbf{1}_3, \bigoplus_{j=1}^4 \mathbf{1}_3, \bigoplus_{j=1}^4 \mathbf{1}_3\right\}$ é a matriz de delineamento associado aos efeitos das cepas dentro dos grupos sem o efeito do controle, com dimensão 42×14 ; \mathbf{I}_{14} é uma matriz identidade, com dimensão 14; σ_b^2 é o componente de variância associado as cepas dentro de grupos.

- Modelo M5:

$$\begin{aligned}
Y_{ijk} &\stackrel{ind.}{\sim} \text{Beta}(\mu_{ijk}, \phi_{ijk}), \\
\boldsymbol{\mu} &= \mathbf{1} - \exp\{-\exp(\boldsymbol{\eta}_1)\}, \\
\boldsymbol{\eta}_1 &= \mathbf{X}_1\boldsymbol{\beta} + \mathbf{Z}\mathbf{b}, \\
\phi &= \exp(\boldsymbol{\eta}_2), \\
\boldsymbol{\eta}_2 &= \mathbf{X}_2\boldsymbol{\lambda}, \\
\mathbf{b} &\sim N_{14}(\mathbf{0}, \mathbf{I}_{14}\sigma_b^2)
\end{aligned} \tag{121}$$

em que: $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)^T$ é o vetor de parâmetros associado aos efeitos dos grupos, com comprimento 4; $\boldsymbol{\lambda} = (\lambda_{11}, \lambda_{21}, \lambda_{22}, \lambda_{23}, \lambda_{24}, \lambda_{25}, \lambda_{31}, \lambda_{32}, \lambda_{33}, \lambda_{34}, \lambda_{41}, \lambda_{42}, \lambda_{43}, \lambda_{44})^T$ é o vetor de parâmetros associado a precisão, com comprimento 14; $\mathbf{1}$ e um vetor de uns, com comprimento 14; $\mathbf{0}$ e um vetor de zeros, com comprimento 14; $\mathbf{b} = (b_{11}, b_{21}, b_{22}, b_{23}, b_{24}, b_{25}, b_{31}, b_{32}, b_{33}, b_{34}, b_{41}, b_{42}, b_{43}, b_{44})^T$ é o vetor de variáveis aleatórias, com comprimento 14; $\mathbf{X}_1 = \text{diag}\{\mathbf{1}_3, \mathbf{1}_{15}, \mathbf{1}_{12}, \mathbf{1}_{12}\}$ é a matriz de delineamento associado aos efeitos dos grupos, com dimensões 42×4 ;

$\mathbf{X}_2 = \text{diag}\left\{\bigoplus_{j=1}^1 \mathbf{1}_3, \bigoplus_{j=1}^5 \mathbf{1}_3, \bigoplus_{j=1}^4 \mathbf{1}_3, \bigoplus_{j=1}^4 \mathbf{1}_3\right\}$ é a matriz de delineamento associado aos efeitos das cepas dentro dos grupos, com dimensão 42×14 ;

$\mathbf{Z} = \text{diag} \left\{ \bigoplus_{j=1}^1 \mathbf{1}_3, \bigoplus_{j=1}^5 \mathbf{1}_3, \bigoplus_{j=1}^4 \mathbf{1}_3, \bigoplus_{j=1}^4 \mathbf{1}_3 \right\}$ é a matriz de delineamento associado aos efeitos das cepas dentro dos grupos sem o efeito do controle, com dimensão 42×14 ; \mathbf{I}_{14} é a matriz identidade, com dimensão 14; σ_b^2 é o componente de variância associado as cepas dentro de grupos.

Para obtenção das estimativas a posteriori dos parâmetros dos modelos apresentados, foi fixado os seguintes parâmetros no algoritmo *HMC*: para o modelo M1 foram geradas três cadeias, 3 mil iterações para cada cadeia, os 100 primeiros valores foram descartados e o espaçamento foi 5. No total, foram gerados um total de 1740 simulações para cada respectivo parâmetro a posteriori.

Já para o restante dos modelos foram geradas três cadeias, 15 mil iterações para cada cadeia, os 3000 primeiros valores foram descartados e o espaçamento foi 20. No total, foram gerados um total de 1050 simulações para cada respectivo parâmetro a posteriori.

O indicativo de convergência entre as cadeias, foi verificado pela medida \hat{R} , que foi igual a 1 para todos os respectivos parâmetros de cada modelo. Note que as distribuições a priori para cada parâmetro dos modelos estão definidos na Tabela 11.

Tabela 11. Distribuições a priori dos parâmetros dos modelos considerando dados provenientes de um experimento com dois fatores aninhados.

Modelo	Priori do β	Priori do	Priori do σ_b	Priori do σ_i
M1	$N_4(\mathbf{0}, 100\mathbf{I}_4)$	$\lambda \sim N_{14}(\mathbf{0}, 100\mathbf{I}_{14})$		
M2	$N_4(\mathbf{0}, 100\mathbf{I}_4)$	$\lambda \sim N(0, 100)$		$LN(0, 100)$
M3	$N_4(\mathbf{0}, 100\mathbf{I}_4)$	$\lambda \sim N_{14}(\mathbf{0}, 100\mathbf{I}_{14})$		$LN(0, 100)$
M4	$N_4(\mathbf{0}, 100\mathbf{I}_4)$	$\lambda \sim N(0, 100)$	$LN(0, 100)$	
M5	$N_4(\mathbf{0}, 100\mathbf{I}_4)$	$\lambda \sim N_{14}(\mathbf{0}, 100\mathbf{I}_{14})$	$LN(0, 100)$	

Por meio dos valores p_d presentes na Tabela 12, podemos identificar que há indícios que os modelos M3 e M5 se ajustaram bem aos dados, pois seus respectivos p_d são os mais próximos do valor 0,5. Além dessa medida de ajuste, temos medidas de comparações de modelos que são critérios de informações *WAIC* e *LOOIC*, que indicam que o modelo M5.

Tabela 12. Critério de informação amplamente aplicável (WAIC), Critério de informação leave-one-out (LOOIC) e o p-valor de verificação preditiva a posterior (p_d).

Modelo	p_d	WAIC	LOOIC
M1	0,33	-172,26	-169,36
M2	0,65	-151,41	-148,36
M3	0,46	-192,52	-189,12
M4	0,68	-155,01	-152,78
M5	0,42	-207,36	-203,47

Na Figura 10, pode-se observar que as densidades dos valores simulados preditos a posteriori (y_{rep}), estão em torno da densidade dos valores observados (y). Isso mostra graficamente indícios que o modelo está bem ajustado aos dados.

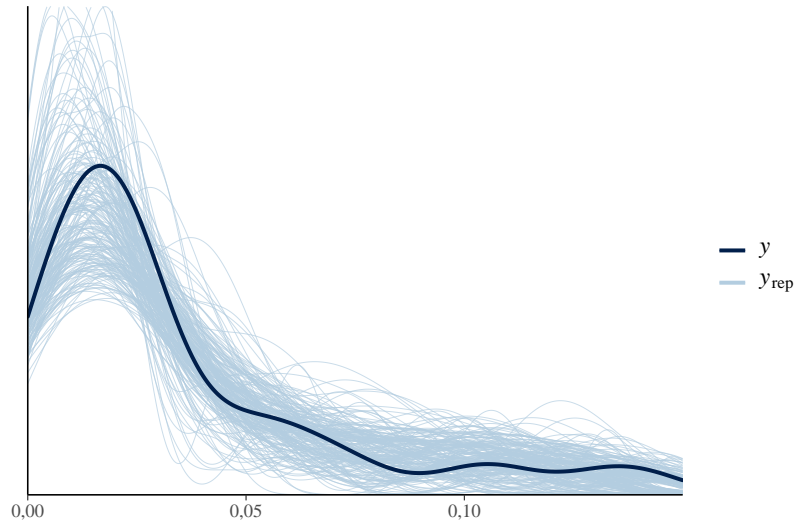


Figura 10. Gráfico das densidade estimadas das simulações dos valores da distribuição preditiva a posteriori do modelo M5.

Na Tabela 13 estão descritos as estimavas a posteriori de cada parâmetro do modelo M5, podemos observar que o β_2 associado ao grupo genético 1 obteve a menor média a posteriori, ou seja, podemos concluir que esse grupo genético de cepas causa o menor índice de severidade esperado na sojicultura. Já o intervalo 95% de credibilidade do componente aleatório σ_b , indica uma baixa variação entre as cepas.

Tabela 13. Estimativas a posteriori dos parâmetros do modelo M5 e respectivos intervalos de 95% de credibilidade (IC 95%).

Parâmetro	Média	Desvio padrão	Mediana	IC 95%
β_1	-2,59	1,37	-2,60	(-5,30; 0,23)
β_2	-3,88	0,57	-3,88	(-5,04;-2,74)
β_3	-3,03	0,73	-3,00	(-4,52;-1,66)
β_4	-3,14	0,65	-3,14	(-4,43;-1,80)
λ_{11}	2,56	1,06	2,74	(-0,09; 4,12)
λ_{21}	8,46	1,27	8,65	(5,39;10,37)
λ_{22}	6,51	1,44	6,81	(3,16; 8,56)
λ_{23}	5,09	1,11	5,22	(2,56; 6,85)
λ_{24}	4,53	1,02	4,65	(2,18; 6,13)
λ_{25}	6,58	1,13	6,74	(3,97; 8,42)
λ_{31}	4,01	2,07	4,12	(-0,02; 7,25)
λ_{32}	4,95	1,14	5,13	(2,27; 6,72)
λ_{33}	2,62	0,91	2,73	(0,51; 4,10)
λ_{34}	4,31	1,16	4,49	(1,69; 6,10)
λ_{41}	3,89	1,39	4,06	(0,73; 6,09)
λ_{42}	5,44	1,27	5,62	(2,16; 7,30)
λ_{43}	1,02	0,72	1,10	(-0,57; 2,24)
λ_{44}	3,12	1,08	3,28	(0,64; 4,80)
σ_b	1,16	0,41	1,11	(0,52; 2,12)

Na Figura 11, fica visível no intervalo de credibilidade ao nível de 95% do parâmetro β_1 referente ao controle, que o zero está incluso, ou seja, podemos considerar que controle causa um índice de severidade médio quase zero ao nível de 95% de credibilidade. Já os demais não incluem o valor zero em seus respectivos intervalos de credibilidade, podem ser considerados diferentes de zero.

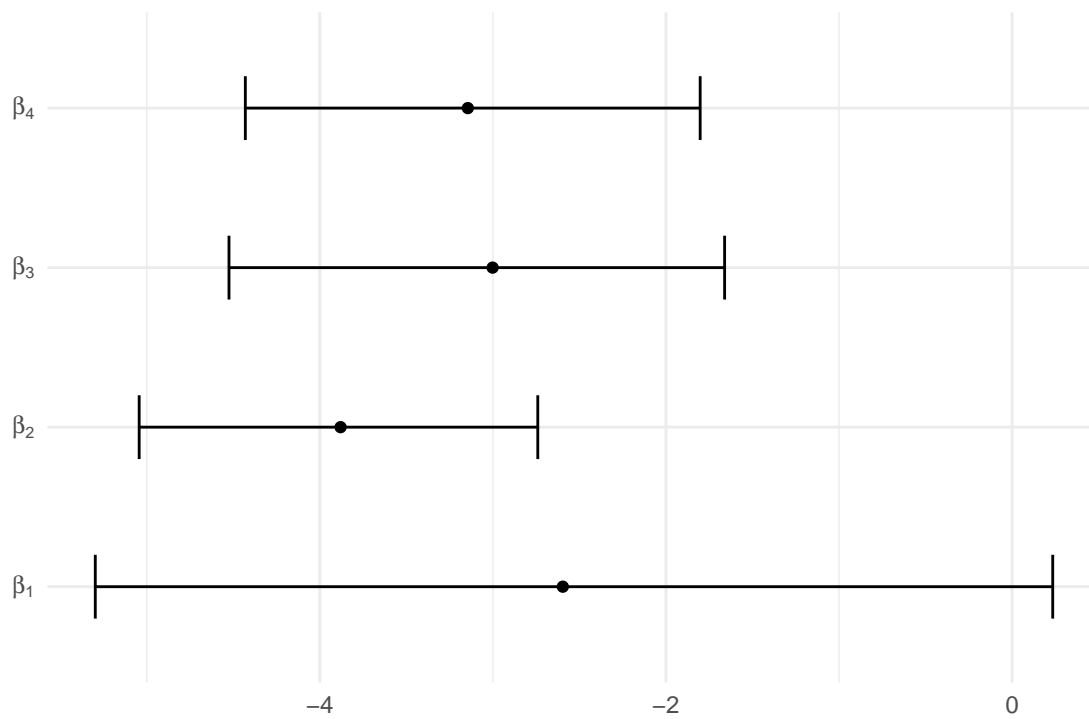


Figura 11. Intervalos de 95% de credibilidade com a mediana dos parâmetros β_1, \dots, β_4 do modelo M5

Nesta subseção foram formuladas as hipóteses que a pesquisadora tinha interesse,

que foram os seguintes:

$$\begin{cases} H_0 : \beta_1 = \beta_2 + \beta_3 + \beta_4 \\ H_1 : \beta_1 \neq \beta_2 + \beta_3 + \beta_4, \end{cases}$$

$$\begin{cases} H_0 : \beta_2 = \beta_3 \\ H_1 : \beta_2 \neq \beta_3, \end{cases}$$

$$\begin{cases} H_0 : \beta_2 = \beta_4 \\ H_1 : \beta_2 \neq \beta_4, \end{cases}$$

$$\begin{cases} H_0 : \beta_2 = \beta_1 \\ H_1 : \beta_2 \neq \beta_1, \end{cases}$$

$$\begin{cases} H_0 : \beta_3 = \beta_4 \\ H_1 : \beta_3 \neq \beta_4, \end{cases}$$

$$\begin{cases} H_0 : \beta_3 = \beta_1 \\ H_1 : \beta_3 \neq \beta_1, \end{cases}$$

$$\begin{cases} H_0 : \beta_4 = \beta_1 \\ H_1 : \beta_4 \neq \beta_1. \end{cases}$$

a partir das hipóteses mostradas anteriormente foram formulados os seguintes contrastes:

- $\beta_1 - (\beta_2 + \beta_3 + \beta_4)/3$;
- $\beta_2 - \beta_3$;
- $\beta_2 - \beta_4$;
- $\beta_2 - \beta_1$;
- $\beta_3 - \beta_4$;
- $\beta_3 - \beta_1$;
- $\beta_4 - \beta_1$.

Após testar os contrastes de interesse mostradas anteriormente, fica constatado na Tabela 14 que no intervalo de credibilidade ao nível de 95%, que todos intervalos incluem o valor zero, ou seja, todas as diferenças podem ser consideradas iguais ao nível de 95% de credibilidade. Isso ocorre parecido na cultura do espinafre, no trabalho de Liu et al. (2021), em que três cepas do *Colletotrichum truncatum* não diferem no índice de severidade, mas nesse trabalho é utilizado a análise de variância, em que um dos pré supostos é que os resíduos do modelo siga normalidade.

Na Figura 12 se observa que cada contraste não se diferencia, pois todos os intervalos das diferenças testadas incluem o valor zero, como a linha vermelha tracejada ressalta no gráfico.

Tabela 14. Resultados dos contrastes do modelo M5. Média, mediana e intervalo de 95% de credibilidade (IC 95%).

Contraste	Média	Mediana	IC 95%
$\beta_1 - (\beta_2 + \beta_3 + \beta_4)/3$	0,76	0,74	(-2,03; 3,83)
$\beta_2 - \beta_3$	-0,84	-0,86	(-2,64; 1,00)
$\beta_2 - \beta_4$	-0,74	-0,74	(-2,52; 0,97)
$\beta_2 - \beta_1$	-1,29	-1,30	(-4,44; 1,69)
$\beta_3 - \beta_4$	0,10	0,13	(-1,86; 1,87)
$\beta_3 - \beta_1$	-0,45	-0,42	(-3,69; 2,57)
$\beta_4 - \beta_1$	-0,55	-0,50	(-3,72; 2,67)

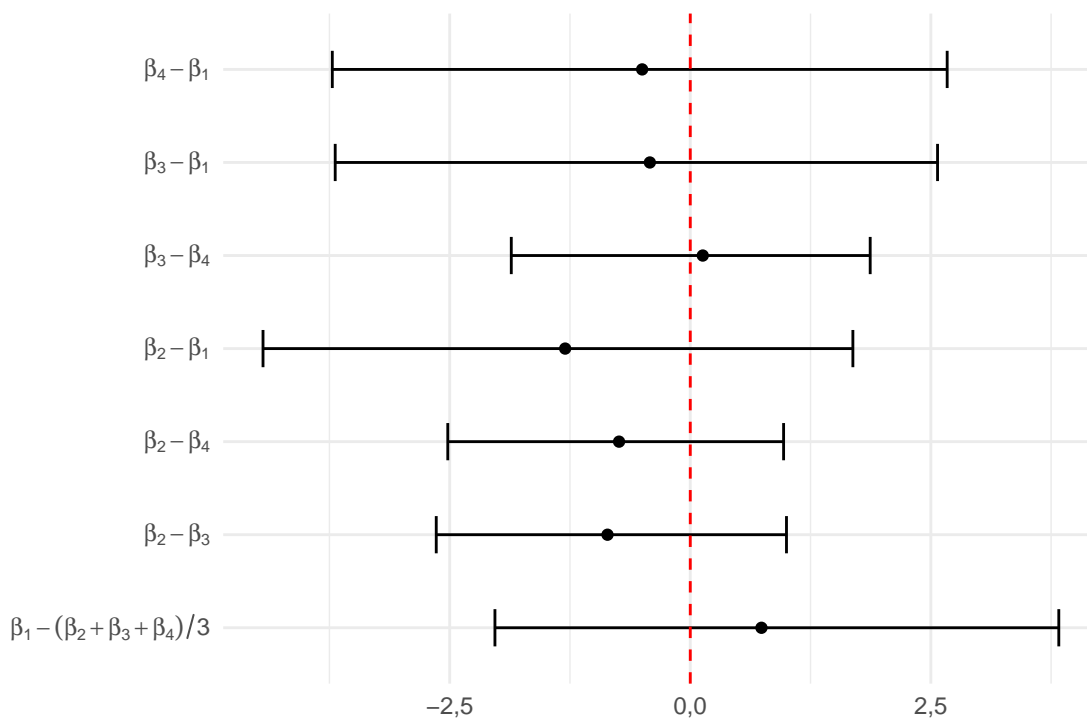


Figura 12. Intervalos de 95% de credibilidade para os contrastes do modelo M5

Podemos observar também que os resultados dos contraste utilizando o modelo M5 com efeitos aleatórios foi diferente, se comparado ao modelo 2 em que as cepas foram consideradas de efeito fixo sobre a variável resposta severidade. Isso ocorre por conta das hipóteses que foram testadas pelo pesquisador: no primeiro modelo o intuito era testar a média dos grupos das cepas específicas eram diferentes, já no segundo caso, o foco passa a ser estudar as médias das populações dos grupos das cepas. Essa metodologia que foi aplicada, pode ser

utilizada para diversos estudos, pois dependendo do que o pesquisador esteja interessado, ele facilmente escolhera o modelo que ira responder o que lhe interessa.

Como foi evidenciado no trabalho, os modelos (60) e (63) são recomendados quando o pesquisador tem o intuito de estudar se tais fatores como fixo e assim avaliar se os mesmos causam a mesma influência na variável resposta. Talvez se Liu et al. (2021) usar nossa metodologia na análise do experimento dela, possa ser que o modelo capte uma diferença as três cepas do fungo *Colletotrichum truncatum*, porque nossa metodologia precisa apenas que a variável resposta esteja contida no intervalo (0,1), que é o caso do experimento dela.

Além disso a nossa metodologia não se limita apenas na análise da severidade na área da fitopatologia, pode ser utilizada para avaliar uma gama de experimentos, como por exemplo: taxa de crescimento de doenças, taxa de mortalidade, índice acadêmico, proporção de doenças em diversas culturas, proporção de gás inalado ou utilizado, propagação de vírus, proporção de pessoas desempregadas, entre outras. É importante ressaltar, que com a utilização da inferência bayesiana podemos incorporar uma informação a priori sobre as covariáveis ou fatores, caso o pesquisador tenha.

Outro ponto positivo dessa metodologia é acomodar bem o desbalanceamento natural, que normalmente acontece nos experimentos de diversas áreas, a maiorias dos modelos apresentados nesse trabalho podem analisar experimentos com desbalanceamento, em estudos com dois fatores aninhados. Nos dados que foram utilizados, existe um desbalanceamento, que é observado nos grupos, pois cada grupo não tem a mesma quantidade de cepas, é o modelo conseguiu se ajustar bem a esses dados, mesmo com o desbalanceamento.

Todo roteiro que foi implementado no software R para obtenção dos resultados relacionado aos modelos mostrados nesta seção, podem ser visto no Apêndice II.

5 CONCLUSÕES

No presente trabalho podemos concluir que todos os objetivos foram alcançados, pois os modelos propostos servem para análise de dados de severidade causada por cepas do fungo na cultura da soja e com os modelos também foi possível responder o segundo objetivo que era testar se havia diferença nos níveis de severidade causada pelos grupos genéticos de cepas.

Além disso, os modelos propostos são de fácil implementação na linguagem de programação R (R Core Team, 2021) por meio do pacote *rstan* (Stan Development Team, 2020) para dados provenientes de experimentos sem ou com desbalanceamento. As análises considerando efeitos de cepa fixos ou aleatórios podem levar, como no caso em questão, a conclusões diferentes.

Podemos concluir baseados nos resultados que a metodologia proposta pode ser utilizada para avaliação da severidade e outras variáveis que estão no intervalo de variação limitado ao $(0,1)$, tanto na cultura da soja como foi mostrado no presente trabalho ou em outras culturas.

5.1 Trabalhos futuros

As ideias sugeridas para trabalhos futuros serão:

- i). Extensão desta metodologia, para analisar dados provenientes de um experimento em que cada repetição contenha réplicas.
- ii). Testar outras distribuições como priori para os parâmetros.
- iii). Testar a sensibilidade do modelo a diferentes distribuições a priori.

REFERÊNCIAS

- AGRAWAL, A. K.; YANG, T.-T. Numerical simulations of low-btu coal gas combustion in a gas turbine combustor. In: **Turbo Expo: Power for Land, Sea, and Air**. New York, Estados Unidos: American Society of Mechanical Engineers, 1991. v. 79009, p. V003T06A003.
- AKAIKE, H. A new look at the statistical model identification. **IEEE transactions on automatic control**, v. 19, n. 6, p. 716–723, 1974.
- AMORIM, L. **Avaliação de doenças**. São Paulo: Editora Ceres, 1995. v. 3. 234–235 p.
- ARANDA-ORDAZ, F. J. On two families of transformations to additivity for binary response data. **Biometrika**, v. 68, n. 2, p. 357–363, 08 1981. ISSN 0006-3444.
- BONAT, W. H.; JR, P. J. R.; ZEVIANI, W. M. Likelihood analysis for a class of beta mixed models. **Journal of Applied Statistics**, Taylor & Francis, v. 42, n. 2, p. 252–266, 2015.
- BOUFLEUR, T. R.; CIAMPI-GUILLARDI, M.; TIKAMI, Í.; ROGÉRIO, F.; THON, M. R.; SUKNO, S. A.; MASSOLA JÚNIOR, N. S.; BARONCELLI, R. Soybean anthracnose caused by *Colletotrichum* species: Current status and future prospects. **Molecular Plant Pathology**, v. 22, n. 4, p. 393–409, 2021. ISSN 13643703.
- BOX, G. E.; TIAO, G. C. **Bayesian inference in statistical analysis**. Hoboken, New Jersey, U.S.: John Wiley & Sons, 2011. v. 40.
- BROOKS, S.; SMITH, J.; VEHTARI, A.; PLUMMER, M.; STONE, M.; ROBERT, C.; TITTERINGTON, D.; NELDER, J.; ATKINSON, A.; DAWID, A.; LAWSON, A.; CLARK, A.; BERNARDO, J.; SAHU, S.; RICHARDSON, S.; GREEN, P.; BURNHAM, K.; DEIORIO, M.; ROBERT, C.; DRAPER, D.; GELFAND, A.; TREVISANI, M.; HODGES, J.; LEE, Y.; DE LUNA, X.; MENG, X. Discussion on the paper by Spiegelhalter, Best, Carlin and van der Linde. **Journal of the Royal Statistical Society. Series B: Statistical Methodology**, Wiley-Blackwell, v. 64, n. 4, p. 616–639, 2002. ISSN 1369-7412.
- CARLIM B.; LOUIS, T. Bayes and empirical bayes methods for data analysis essays on item response theory. **Chapman and Hall**, New York, 2001.
- CHIB, S.; GREENBERG, E. Understanding the Metropolis-Hastings Algorithm. **The American Statistician**, American Statistical Association, Taylor & Francis, Ltd., v. 49, n. 4, p. 327–335, 1995. ISSN 00031305.
- CRIBARI-NETO, F.; ZEILEIS, A. Beta regression in r. **Journal of statistical software**, v. 34, p. 1–24, 2010.
- DIAO, Y.; ZHANG, C.; XU, J.; LIN, D.; LIU, L.; MTUNG'E, O. G.; LIU, X. Genetic differentiation and recombination among geographic populations of the fungal pathogen *Colletotrichum truncatum* from chili peppers in China. **Evolutionary applications**, v. 8, n. 1, p. 108–118, 2015.

DUANE, S.; KENNEDY, A.; PENDLETON, B. J.; ROWETH, D. Hybrid Monte Carlo. **Physics Letters B**, v. 195, n. 2, p. 216–222, sep 1987. ISSN 03702693.

EMBRAPA. **Soja em números (safra 2020/21)**. 2022. Acesso em: 09 Fev. 2022. Disponível em: <<https://www.embrapa.br/soja/cultivos/soja1/dados-economicos>>.

FERRARI, S. L.; CRIBARI-NETO, F. Beta regression for modelling rates and proportions. **Journal of Applied Statistics**, v. 31, n. 7, p. 799–815, 2004. ISSN 02664763.

FIESP. **Soja e suas riquezas – história**. 2022. Acesso em: 09 Fev. 2022. Disponível em: <<https://www.fiesp.com.br/sindimilho/sobre-o-sindmilho/curiosidades/soja-e-suas-riquezas-historia/>>.

FLÁVIA, R.; GLADIEUX, P.; MASSOLA JUNIOR, N. S.; CIAMPI-GUILLARDI, M. Multiple introductions without admixture of *Colletotrichum truncatum* associated with soybean anthracnose in Brazil. **Phytopathology**, v. 109, n. 4, p. 681–689, 2019. ISSN 0031949X.

GELMAN, A. Two simple examples for understanding posterior p-values whose distributions are far from uniform. **Electronic Journal of Statistics**, v. 7, p. 2595–2602, 1 2013. ISSN 1935-7524.

GELMAN, A.; CARLIN, J. B.; STERN, H. S.; DUNSON, D. B.; VEHTARI, A.; RUBIN, D. B. **Bayesian Data Analysis**. 3. ed. New York: Chapman and Hall/CRC, 2013. 675 p. ISBN 9780429113079.

GEMAN, S.; GEMAN, D. Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images. **IEEE Transactions on pattern analysis and machine intelligence**, IEEE, n. 6, p. 721–741, 1984.

GOLENKO-GINZBURG, D. On the distribution of activity time in pert. **Journal of the Operational Research Society**, Taylor & Francis, v. 39, n. 8, p. 767–771, 1988.

GRAHAM, V.; HOLLANDS, K. A method to generate synthetic hourly solar radiation globally. **Solar Energy**, v. 44, n. 6, p. 333–341, 1990. ISSN 0038-092X.

HARROP-WILLIAMS, K. Random nature of soil porosity and related properties. **Journal of Engineering Mechanics**, v. 115, n. 5, p. 1129–1133, 1989.

HASTINGS, W. K. Monte Carlo sampling methods using Markov chains and their applications. **Biometrika**, v. 57, n. 1, p. 97–109, 1970. ISSN 1464-3510.

HOFFMAN, M. D.; GELMAN, A. The no-u-turn sampler: Adaptively setting path lengths in hamiltonian monte carlo. **Journal of Machine Learning Research**, v. 15, n. 47, p. 1593–1623, 2014.

JANARDAN, K.; PADMANABHAN, G. Double bounded beta distribution for hydrologic variables. In: **Proc. 17th Annu. Pittsburgh Conference (part 3)**. [S.l.: s.n.], 1986. v. 17, p. 1107–1111.

JEFFREYS, H. **The theory of probability**. 3. ed. Oxford, Inglaterra: Oxford University Press (OUP), 1998.

JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous univariate distributions**. 2. ed. New York: John Wiley & sons, 1995. v. 2. ISBN 978-0-471-58494-0.

KARAVIAS, F.; MYERS, A. L. Molecular thermodynamics of adsorption from gas mixtures: composition of adsorbed phase from gravimetric data. **Chemical engineering science**, Elsevier, v. 47, n. 6, p. 1441–1451, 1992.

LESAFFRE, E.; LAWSON, A. B. **Bayesian Methods in Biostatistics (Statistics in Practice)**. 1. ed. Nova Jersey: John Wiley & Sons, 2012. ISBN 9780470018231.

LIU, B.; STEIN, L.; COCHRAN, K.; TOIT, L. J. du; FENG, C.; CORRELL, J. C. Three new fungal leaf spot diseases of spinach in the united states and the evaluation of fungicide efficacy for disease management. **Plant Disease**, American Phytopathological Society, v. 105, p. 316–323, 2 2021. ISSN 19437692.

MAFFETT, A. L.; WACKERMAN, C. C. The modified beta density function as a model for synthetic aperture radar clutter statistics. **IEEE Transactions on geoscience and remote sensing**, IEEE, v. 29, n. 2, p. 277–283, 1991.

MANCO, O. C. U. **Modelos de regressão beta com efeitos aleatórios normais e não normais para dados longitudinais**. Tese (Doutorado) — Universidade de São Paulo, 2013.

MARTÍNEZ-MINAYA, J.; CONESA, D.; LÓPEZ-QUÍLEZ, A.; MIRA, J. L.; VICENT, A. Modeling inoculum availability of plurivorosphaerella nawae in persimmon leaf litter with bayesian beta regression. **Phytopathology®**, Am Phytopath Society, v. 111, n. 7, p. 1184–1192, 2021.

MCNALLY, R. J. Q. Maximum likelihood estimation of the parameters of the prior distributions of three variables that strongly influence reproductive performance in cows. **Biometrics**, [Wiley, International Biometric Society], v. 46, n. 2, p. 501–514, 1990. ISSN 0006341X, 15410420.

METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equation of state calculations by fast computing machines. **The Journal of Chemical Physics**, v. 21, n. 6, p. 1087–1092, 1953. ISSN 0021-9606.

NEAL, R. M. **Handbook of Markov Chain Monte Carlo**. New York: Chapman and Hall/CRC, 2011. 113–162 p. ISBN 9780429138508.

NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear models. **Journal of the Royal Statistical Society. Series A (General)**, v. 135, n. 3, p. 370, 1972. ISSN 00359238.

NISHIO, M.; ARAKAWA, A. Performance of hamiltonian Monte Carlo and no-u-turn sampler for estimating genetic parameters and breeding values. **Genetics selection evolution**, BioMed Central, v. 51, n. 1, p. 1–12, 2019. ISSN 12979686.

OLIVOTO, T. **pliman: Tools for plant image analysis**. 2021. R package version 0.3.0.9000. Disponível em: <<https://github.com/TiagoOlivoto/pliman>>.

OSPINA, P. L. E. **Regressão beta**. Tese (Doutorado) — Tese (Doutorado em Estatística), 2007.

PANDEY, A. K.; BASANDRAI, A. K.; BASANDRAI, D.; BODDEPALLI, V. N.; RATHORE, A.; ADAPALA, G.; NAIR, R. M. Field-relevant new sources of resistance to anthracnose caused by *colletotrichum truncatum* in a mungbean mini-core collection. **Plant Disease**, American Phytopathological Society, v. 105, 7 2021. ISSN 19437692.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2021. Disponível em: <<https://www.R-project.org/>>.

RAHMAN, S.; KHALLAT, M.; SALAMEH, Z. Characterization of insolation data for use in photovoltaic system analysis models. **Energy**, v. 13, n. 1, p. 63–72, 1988. ISSN 0360-5442.

ROBERT, C.; CASELLA, G. A short history of Markov Chain Monte Carlo: subjective recollections from incomplete data. **Statistical Science**, Institute of Mathematical Statistics, v. 26, n. 1, Feb 2011. ISSN 0883-4237.

ROGÉRIO, F. **Population genetics of *Colletotrichum truncatum* associated with soybean anthracnose**. Tese (Doutorado) — Universidade de São Paulo, 2019.

ROGÉRIO, F.; MASSOLA JUNIOR, N. S. Population genetics of *Colletotrichum truncatum* associated with soybean anthracnose. 2019.

SCHMIDT, A. M.; MORAES, C. P. de; MIGON, H. S. A hierarchical dynamic beta regression model of school performance in the brazilian mathematical olympiads for public schools. **arXiv preprint arXiv:1507.00565**, 2015.

SCHWARZ, G. Estimating the dimension of a model. **The annals of statistics**, p. 461–464, 1978.

SHARMA, G.; PINNAKA, A. K.; SHENOY, B. D. Infra-specific diversity of *Colletotrichum truncatum* associated with chilli anthracnose in India based on microsatellite marker analysis. **Archives of Phytopathology and Plant Protection**, Taylor & Francis, v. 47, n. 20, p. 2509–2523, 2014.

SHARMA, S.; GUPTA, G.; RAMTEKE, R. *Colletotrichum truncatum* [(Schw.) Andrus & WD Moore], the causal agent of anthracnose of soybean [*Glycine max* (L.) Merrill]—A Review. **Soybean**, v. 9, p. 31–52, 2011.

SILVA, C. Q. da; MIGON, H. S.; CORREIA, L. T. Dynamic bayesian beta models. **Computational Statistics & Data Analysis**, Elsevier, v. 55, n. 6, p. 2074–2089, 2011.

SILVA JÚNIOR, M. B. da; RESENDE, M. L. V. de; POZZA, E. A.; RESENDE, A. R.; VASCONCELOS, V. A. M.; MONTEIRO, A. C. A.; SILVEIRA, G. C. D.; BOTELHO, D. M. dos S. Phosphites for the management of anthracnose in soybean pods. **Journal of**

Plant Pathology, Springer Science and Business Media Deutschland GmbH, v. 103, p. 611–617, 5 2021. ISSN 22397264.

SMITHSON, M.; VERKUILEN, J. A better lemon squeezer? maximum-likelihood regression with beta-distributed dependent variables. **Psychological methods**, American Psychological Association, v. 11, n. 1, p. 54, 2006.

SPIEGELHALTER, D. J.; BEST, N. G.; CARLIN, B. P.; LINDE, A. V. D. The deviance information criterion: 12 years on. **Journal of the Royal Statistical Society. Series B**, v. 76, n. 3, p. 485–493, 2014. ISSN 13697412, 14679868.

Stan Development Team. **RStan: the R interface to Stan**. 2020. R package version 2.21.2. Disponível em: <<http://mc-stan.org/>>.

VEHTARI, A.; GELMAN, A.; GABRY, J. Practical bayesian model evaluation using leave-one-out cross-validation and WAIC. **Statistics and Computing**, Springer Science and Business Media LLC, v. 27, n. 5, p. 1413–1432, 2016. ISSN 1573-1375.

VERKUILEN, J.; SMITHSON, M. Mixed and mixture regression models for continuous bounded responses using the beta distribution. **Journal of Educational and Behavioral Statistics**, Sage Publications Sage CA: Los Angeles, CA, v. 37, n. 1, p. 82–113, 2012.

WATANABE, S. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. **Journal of Machine Learning Research**, v. 11, p. 3571–3594, 2010.

XIANG, Q.; ANDERSEN, S. L.; PERLS, T. T.; SEBASTIANI, P. Studying the interplay between apolipoprotein e and education on cognitive decline in centenarians using bayesian beta regression. **Frontiers in genetics**, Frontiers, p. 1673, 2021.

ZERBETO, A. P. **Modelos de mistura beta mistos sob abordagem bayesiana**. Tese (Doutorado) — Universidade de São Paulo, 2018.

APÊNDICES

Apêndice I

```
require(pliman)
imagem <- soja.jpg
healthy <- imagem_sadia.jpg
symptoms <- imagem_com_doenca.jpg
background <- imagem_do_fundo.jpg

symptomatic_area(img = imagem,
  img_healthy      = healthy,
  img_symptoms     = symptoms,
  img_background   = background)
```

Apêndice II

```
#' title: "DISSERTAÇÃO"
#' author: "ANTONIO LEOPOLDO CARDOSO SABINO"
# Limpando a memória
rm(list=ls(all=T))
# Este valor alto (999) evita o retorno dos números em notação científica.
options(scipen = 999, OutDec = ',')
#'
# Bibliotecas utilizadas -----

require(readxl)
require(ggplot2)
require(loo)
require(rstan)
require(xtable)

# Dados utilizados -----

dados_media_rep <- read.csv("dados_media_soja1.csv", sep = ';', header = T)
```



```

dados_media_rep <- transform(dados_media_rep, tratamento=factor(tratamento),
                             severidade=as.numeric(severidade),
                             grupo=factor(grupo))

# modelos considerando as cepas de efeito fixo -----

codigo_stan1 <- "
data {
  int<lower=0> N;
  int<lower=0> NI1;
  matrix[N, NI1] X1;
  vector[N] y;
}
parameters {
  vector[NI1] beta;
  real<lower=0> phi;
}
transformed parameters{
  vector[N] mu;
  mu = inv_cloglog(X1*beta);
}
model {
  phi ~ lognormal(0,10);
  beta ~ normal(0,10);
  y ~ beta_proportion(inv_cloglog(X1*beta),phi);
}
generated quantities {
  real y_rep[N];
  vector[N] log_lik;
  for (n in 1:N) {
    log_lik[n] = beta_proportion_lpdf(y[n] | inv_cloglog(X1[n]*beta), phi);
  }
  y_rep = beta_proportion_rng(inv_cloglog(X1*beta), phi);
}
"

```

```

X <- model.matrix(severidade~-1+tratamento,dados_media_rep)
y <- dados_media_rep$severidade
N <- length(y)
NI <- ncol(X)
lista_dados <- list(X1=X1,X2=X2,y=y,N=N,NI1=NI1,NI2=NI2)

modelo1 <- stan(model_code = codigo_stan1, data = lista_dados,
               warmup = 100, iter = 3000, chains = 3,
               cores = 10, thin = 5,algorithm = "HMC")

codigo_stan2 <- "
data {
  int<lower=0> N;
  int<lower=0> NI1;
  int<lower=0> NI2;
  matrix[N, NI1] X1;
  matrix[N, NI2] X2;
  vector[N] y;
}
parameters {
  vector[NI1] beta;
  vector[NI2] lambda;
}
transformed parameters{
  vector[N] mu;
  vector[N] phi;
  mu = inv_cloglog(X1*beta);
  phi = exp(X2*lambda);
}
model {
  lambda ~ normal(0,10);
  beta ~ normal(0,10);
  y ~ beta_proportion(inv_cloglog(X1*beta),exp(X2*lambda));
}
generated quantities {
  real y_rep[N];

```

```

vector[N] log_lik;
for (n in 1:N) {
  log_lik[n] = beta_proportion_lpdf(y[n] | inv_cloglog(X1[n]*beta),
  exp(X2[n]*lambda));
}
y_rep = beta_proportion_rng(inv_cloglog(X1*beta), exp(X2*lambda));
}
"

```

```

X1 <- model.matrix(severidade~-1+tratamento,dados_media_rep)
X2 <- model.matrix(severidade~-1+tratamento,dados_media_rep)
y <- dados_media_rep$severidade
N <- length(y)
NI1 <- ncol(X1)
NI2 <- ncol(X2)
lista_dados <- list(X1=X1,X2=X2,y=y,N=N,NI1=NI1,NI2=NI2)

```

```

modelo2 <- stan(model_code = codigo_stan2, data = lista_dados,
               warmup = 100, iter = 3000, chains = 3,
               cores = 10, thin = 5,algorithm = "HMC")

```

```

### obtenção do pd valor para cada modelo
ppc <- function(mod,y,estatis){
  y <- y ## variável resposta
  matyrep <- rstan::extract(mod,'y_rep')$y_rep
  K <- nrow(matyrep)
  a <- 0
  for (i in 1:K) {
    a <- I(estatis(matyrep[i,])>=estatis(y))+a
  }
  b <- a/K
  b
}

```

```

modelos <- list(modelo1, modelo2)

```

```

pd_modelos <- NULL
for (i in 1:2) {
  pd_modelos[i]=ppc(modelos[[i]],y,mean)
}

# Critérios WAIC e LOOIC
criterio_waic <- NULL
for (i in 1:2) {
  criterio_waic[i]=waic(extract_log_lik(modelos[[i]]))["waic"]
}
criterio_loo <- NULL
for (i in 1:2) {
  criterio_loo[i]=loo(modelos[[i]])['looic']
}

# gráficos da densidade dos valores preditos a posteriori
bayesplot::ppc_dens_overlay(y,
  rstan::extract(modelo1,'y_rep')$y_rep)+
  xlim(0,.15)+ylim(0,30)
bayesplot::ppc_dens_overlay(y,
  rstan::extract(modelo2,'y_rep')$y_rep)+
  xlim(0,.15)+ylim(0,30)

# intervalos de credibilidade dos valores preditos
bayesplot::ppc_intervals(cloglog(y),
  cloglog(rstan::extract(modelo1,'y_rep')$y_rep),
  prob = .95,prob_outer = .95)+xlab("Observações")+
  ylab("cloglog(severidade)")+ylim(-25,0)
bayesplot::ppc_intervals(cloglog(y),
  cloglog(rstan::extract(modelo2,'y_rep')$y_rep),
  prob = .95,prob_outer = .95)+xlab("Observações")+
  ylab("cloglog(severidade)")+ylim(-25,0)

# intervalos de credibilidade dos parâmetros
rstan::plot(modelo2,pars='beta')
rstan::plot(modelo2,pars='lambda')

```

```

# contrastes do modelo 2
C <- as.matrix(read_xlsx('MC.xlsx',col_names=F))
betas <- rstan::extract(modelo2,'beta')$beta
contraste_de_grupos_de_betas <- C%*%t(betas)

# modelos utilizando dados com dois fatores aninhados -----

# modelo M1

X1 <- model.matrix(severidade~-1+grupo,dados_media_rep)
X2 <- model.matrix(severidade~-1+tratamento,dados_media_rep)
y <- dados_media_rep$severidade
N <- length(y)
NI1 <- ncol(X1)
NI2 <- ncol(X2)
lista_dados <- list(X1=X1,X2=X2,y=y,N=N,NI1=NI1,NI2=NI2)
M1 <- stan(model_code = codigo_stan2, data = lista_dados , warmup = 100,
           iter = 3000, chains = 3, cores = 10,
           thin = 5,algorithm = "HMC")

print(M1)

# modelo M2

codigo_stan3 <- "
data {
  int<lower=0> N;
  int<lower=0> NI;
  int<lower=0> NJ;
  matrix[N, NI] X;
  matrix[N, NJ] Z;
  vector[N] y;

```

```

vector[NJ] zeros;
matrix[NJ,NI-1] G;
}

parameters {
vector[NI] beta;
vector[NJ] b;
real<lower=0> phi;
vector<lower=0>[NI-1] sigma;
}

transformed parameters{
vector[N] mu;
mu = inv_cloglog(X*beta+Z*b);
}

model {
sigma ~ lognormal(0,10);
b ~ multi_normal(zeros,diag_matrix(G*sigma));
phi ~ lognormal(0,10);
beta ~ normal(0, 10);
y ~ beta_proportion(inv_cloglog(X*beta+Z*b),phi);
}

generated quantities {
real y_rep[N];
vector[N] log_lik;
for (n in 1:N) {
log_lik[n] = beta_proportion_lpdf(y[n] | inv_cloglog(X[n]*beta+Z[n]*b),
phi);
}
y_rep = beta_proportion_rng(inv_cloglog(X*beta+Z*b),phi);
}

"

X <- model.matrix(severidade~-1+grupo,dados_media_rep)
Z <- model.matrix(severidade~-1+tratamento,dados_media_rep)[,-1]
NI <- ncol(X)

```

```

NJ <- ncol(Z)
y <- dados_media_rep$severidade
N <- length(y)
zeros <- rep(0,NJ)
G <- matrix(rep(c(1,0,1,0,1),c(5,13,4,13,4)),ncol = 3,byrow = F)
lista_dados <- list(X=X,Z=Z,y=y,N=N,NI=NI,NJ=NJ,zeros=zeros,G=G)
M2 <- stan(model_code = codigo_stan3, data = lista_dados, warmup = 3000,
           iter = 15000, chains = 3, cores = 10,
           thin = 20,algorithm = "HMC")

print(M2)

# modelo M3

codigo_stan4 <- "
data {
  int<lower=0> N;
  int<lower=0> NJ;
  int<lower=0> NI1;
  int<lower=0> NI2;
  matrix[N, NI1] X1;
  matrix[N, NI2] X2;
  matrix[N, NJ] Z;
  vector[N] y;
  vector[NJ] zeros;
  matrix[NJ,NI1-1] G;
}

parameters {
  vector[NJ] b;
  vector<lower=0>[NI1-1] sigma;
  vector[NI1] beta;
  vector[NI2] lambda;
}

transformed parameters{

```

```

vector[N] mu;
vector[N] phi;
  mu = inv_cloglog(X1*beta+Z*b);
  phi = exp(X2*lambda);
}
model {
  sigma ~ lognormal(0,10);
  b ~ multi_normal(zeros,diag_matrix(G*sigma));
  lambda ~ normal(0,10);
  beta ~ normal(0,10);
  y ~ beta_proportion(inv_cloglog(X1*beta+Z*b),exp(X2*lambda));
}
generated quantities {
  real y_rep[N];
  vector[N] log_lik;
  for (n in 1:N) {
    log_lik[n] = beta_proportion_lpdf(y[n] | inv_cloglog(X1[n]*beta+Z[n]*b),
    exp(X2[n]*lambda));
  }
  y_rep = beta_proportion_rng(inv_cloglog(X1*beta+Z*b),exp(X2*lambda));
}
"
X1 <- model.matrix(severidade~-1+grupo,dados_media_rep)
X2 <- model.matrix(severidade~-1+tratamento,dados_media_rep)
Z <- model.matrix(severidade~-1+tratamento,dados_media_rep)[,-1]
NI1 <- ncol(X)
NI2 <- ncol(X2)
NJ <- ncol(Z)
y <- dados_media_rep$severidade
N <- length(y)
zeros <- rep(0,NJ)
G <- matrix(rep(c(1,0,1,0,1),c(5,13,4,13,4)),ncol = 3,byrow = F)
lista_dados <- list(X1=X1,X2=X2,Z=Z,y=y,N=N,NI1=NI1,
  NI2=NI2,NJ=NJ,zeros=zeros,G=G)
M3 <- stan(model_code = codigo_stan4,data = lista_dados,
  warmup = 3000, iter = 15000, chains = 3,

```



```
cores = 10, thin = 20, algorithm = "HMC")

print(M3)

# modelo M4

codigo_stan5 <- "
data {
  int<lower=0> N;
  int<lower=0> NI;
  int<lower=0> NJ;
  matrix[N, NI] X;
  matrix[N, NJ] Z;
  vector[N] y;
}

parameters {
  vector[NI] beta;
  vector[NJ] b;
  real<lower=0> phi;
  real<lower=0> sigma;
}

transformed parameters{
  vector[N] mu;
  mu = inv_cloglog(X*beta+Z*b);
}

model {
  sigma ~ lognormal(0,10);
  b ~ normal(0,sigma);
  phi ~ lognormal(0,10);
  beta ~ normal(0, 10);
  y ~ beta_proportion(inv_cloglog(X*beta+Z*b),phi);
}

generated quantities {
```

```

real y_rep[N];
vector[N] log_lik;
  for (n in 1:N) {
    log_lik[n] = beta_proportion_lpdf(y[n] | inv_cloglog(X[n]*beta+Z[n]*b)
    ,phi);
  }
  y_rep = beta_proportion_rng(inv_cloglog(X*beta+Z*b),phi);
}
"

X <- model.matrix(severidade~-1+grupo,dados_media_rep)
X2 <- model.matrix(severidade~-1+tratamento,dados_media_rep)
Z <- model.matrix(severidade~-1+tratamento,dados_media_rep)
NI1 <- ncol(X)
NI2 <- ncol(X2)
NJ <- ncol(Z)
y <- dados_media_rep$severidade
N <- length(y)
lista_dados <- list(X=X,Z=Z,y=y,N=N,NI=NI1,NJ=NJ)
M4 <- stan(model_code = codigo_stan5, data = lista_dados,
          warmup = 3000, iter = 10000, chains = 3,
          cores = 10, thin = 20,algorithm = "HMC")

print(M4)

# modelo M5

codigo_stan6 <- "
data {
  int<lower=0> N;
  int<lower=0> NJ;
  int<lower=0> NI1;
  int<lower=0> NI2;
  matrix[N, NI1] X1;
  matrix[N, NI2] X2;

```

```

matrix[N, NJ] Z;
vector[N] y;
}

parameters {
  vector[NJ] b;
  real<lower=0> sigma;
  vector[NI1] beta;
  vector[NI2] lambda;
}

transformed parameters{
  vector[N] mu;
  vector[N] phi;
  mu = inv_cloglog(X1*beta+Z*b);
  phi = exp(X2*lambda);
}

model {
  sigma ~ lognormal(0,10);
  b ~ normal(0,sigma);
  lambda ~ normal(0,10);
  beta ~ normal(0,10);
  y ~ beta_proportion(inv_cloglog(X1*beta+Z*b),exp(X2*lambda));
}

generated quantities {
  real y_rep[N];
  vector[N] log_lik;
  for (n in 1:N) {
    log_lik[n] = beta_proportion_lpdf(y[n]|inv_cloglog(X1[n]*beta+Z[n]*b),
    exp(X2[n]*lambda));
  }
  y_rep = beta_proportion_rng(inv_cloglog(X1*beta+Z*b),exp(X2*lambda));
}

"
X1 <- model.matrix(severidade~-1+grupo,dados_media_rep)
X2 <- model.matrix(severidade~-1+tratamento,dados_media_rep)

```

```

Z <- model.matrix(severidade~-1+tratamento,dados_media_rep)
NI1 <- ncol(X1)
NI2 <- ncol(X2)
NJ <- ncol(Z)
y <- dados_media_rep$severidade
N <- length(y)
lista_dados <- list(X1=X1,X2=X2,Z=Z,y=y,N=N,NI1=NI1,NI2=NI2,NJ=NJ)

M5 <- stan(model_code = codigo_stan6, data = lista_dados,
           warmup = 3000, iter = 10000, chains = 3,
           cores = 10, thin = 20, algorithm = "HMC")

print(M5)

### obtenção do pd valor para cada modelo

modelos <- list(M1,M2,M3,M4,M5)

pd_modelos <- NULL
for (i in 1:2) {
  pd_modelos[i]=ppc(modelos[[i]],y,mean)
}

# Critérios WAIC e LOOIC

criterio_waic <- NULL
for (i in 1:5) {
  criterio_waic[i]=waic(extract_log_lik(modelos[[i]]))["waic"]
}

criterio_loo <- NULL
for (i in 1:5) {
  criterio_loo[i]=loo(modelos[[i]])['looic']
}

```

```
# gráfico da densidade dos valores preditos a posteriori
bayesplot::ppc_dens_overlay(y,
  rstan::extract(M5,'y_rep')$y_rep)+
  xlim(0,.15)+ylim(0,30)

# gráficos dos intervalos de credibilidade dos parâmetros betas
rstan::plot(M5,pars='beta')

# contrastes
C <- as.matrix(read_xlsx('MC1.xlsx',col_names=F))
betas <- rstan::extract(M5,'beta')$beta
contraste_de_grupos_de_betas <- C%*%t(betas)
```