

Análise comportamental de ovelhas submetidas à intervenção humana usando GAMLSS

Lineu Alberto Cavazani de Freitas ¹

Cesar Augusto Taconeli ²

José Luiz Padilha da Silva ³

Priscilla Regina Tamioso ⁴

Carla Forte Maiolino Molento ⁵

1 Introdução

O presente estudo é uma análise dos dados obtidos em experimento sobre comportamento de ovelhas descrito em [Tamioso et al. \(2017\)](#). Na ocasião foram observadas 20 ovelhas pertencentes a 2 linhagens genéticas, que foram submetidas à escovação por um humano que lhes era familiar. Verificou-se as proporções de tempo em que os animais permaneciam em determinadas posturas, o que constitui uma variável resposta restrita ao intervalo $[0,1]$. O modelo de regressão mais conhecido para dados restritos ao intervalo $(0,1)$ é o modelo beta ([FERRARI; CRIBARI-NETO, 2004](#)).

Todavia, caso a variável resposta em questão assuma valores iguais a 0 e/ou 1 a distribuição beta se torna inviável pois o suporte desta é o intervalo aberto $(0,1)$. Neste contexto, as distribuições beta inflacionadas surgem como uma solução viável. Casos deste tipo são tratados em estudos de aplicação como o descrito em [Pereira, Souza e Cribari-Neto \(2014\)](#). Em [Ospina e Ferrari \(2010\)](#) são propostas variações da distribuição beta que permitem modelar dados nos intervalos $[0,1)$, $(0,1]$ ou $[0,1]$.

No presente estudo considerou-se a modelagem usando GAMLSS. Esta metodologia permite não só a modelagem da média como há a possibilidade de incluir efeitos aleatórios e de modelar os demais parâmetros da distribuição em função de covariáveis. O objetivo da modelagem foi verificar o efeito da linhagem genética, escovação e isolamento na proporção do tempo em que ovelhas permaneciam com as orelhas levantadas e assimétricas. A variável resposta assume valores no intervalo $[0, 1]$, o que motivou o uso da distribuição beta inflacionada.

¹PET-Est. UFPR. e-mail: lialcafre@gmail.com

²DEST UFPR.

³DEST UFPR.

⁴LABEA UFPR.

⁵LABEA UFPR.

2 Material e Métodos

No experimento sob análise, vinte ovelhas classificadas como reativas ou não reativas ao isolamento social temporário foram submetidas à escovação por um humano familiar. O experimento foi realizado na fazenda experimental INRA La Fage, Roqueford, França, em setembro de 2015. As ovelhas tinham 15 meses de idade, eram não gestantes e não amamentavam quando foram observadas.

O experimento foi conduzido em três sessões experimentais: na primeira tinha-se uma grade de metal separando o animal testado dos demais animais, sem distância entre eles; na segunda havia duas grades de metal separando o animal testado dos demais animais a uma distância de 1,7 m, sem período de adaptação para o distanciamento dos animais; na terceira sessão os animais voltaram a ser separados por apenas uma grade, sem distanciamento dos demais animais.

As sessões de testes ocorreram dois dias após a fase de adaptação dos animais ao equipamento e aos humanos e, em cada sessão, as ovelhas foram observadas em 3 momentos distintos: fase de pré escovação, com duração de 2 minutos e 30 segundos; fase de escovação, com duração de 3 minutos; e pós escovação, com duração de 2 minutos e 30 segundos.

Os dados coletados dizem respeito às proporções de tempo em que dado animal permaneceu em uma série de posturas, tratando-se, portanto, de respostas entre 0 e 1 e de natureza contínua. Ao todo 9 respostas foram observadas, dentre as quais a proporção do tempo em que o animal permaneceu em pé, movimentando a cauda, comendo ou ruminando, entre outras posturas.

A resposta selecionada para análise foi a proporção do tempo em que o animal permaneceu com as orelhas levantadas ou assimétricas. Como se trata de uma variável restrita ao intervalo $[0, 1]$, a distribuição de probabilidades selecionada para modelar esses dados deve ter suporte adequado, sendo assim, a distribuição beta inflacionada é uma das candidatas para a modelagem. Embora o suporte da distribuição beta não seja adequado, pode-se utilizar a correção proposta em [Smithson e Verkuilen \(2006\)](#) para trazer todas as respostas para o suporte da distribuição beta. A transformação sugerida é $y_c = \frac{y(n+1)+0,5}{n}$.

Portanto, para ajuste dos modelos, foram utilizadas a distribuição beta e beta inflacionada. Adicionalmente, incluíram-se dois efeitos aleatórios aos modelos: de animal e de animal dentro de sessão. A inserção desses efeitos deve-se à necessidade de incorporar as correlações entre as medidas num mesmo animal e do animal numa mesma sessão.

Os modelos ajustados continham os efeitos fixos de: sessão, momento experimental, linhagem genética além dos efeitos aleatórios de animal e animal dentro de sessão experimental. Considerou-se também a inclusão das interações de segunda ordem para os fatores de efeito fixo. Na distribuição beta inflacionada foi especificado ainda um preditor para o parâmetro de inflação contendo os efeitos aditivos de sessão, momento e linhagem.

Todas as análises foram realizadas utilizando o software R versão 3.4.1 ([R Core Team, 2017](#)) e, para ajuste dos modelos, foi usado o pacote `gamlss` ([STASINOPOULOS; RIGBY et al., 2007](#)).

3 Resultados e discussões

3.1 Análise descritiva

Na Figura 1 é feita uma breve análise descritiva dos dados em estudo. O gráfico superior ao lado esquerdo mostra que trata-se de uma variável inflacionada em um (72 ocorrências) e com frequências menores nas demais regiões do suporte. Os demais gráficos apresentam boxplots da proporção de tempo versus sessão, linhagem e momento. Nota-se que há uma maior simetria na sessão 1 que nas demais e que os animais menos reativos ao isolamento social permanecem mais tempo com as orelhas levantadas ou assimétricas que aqueles considerados reativos.

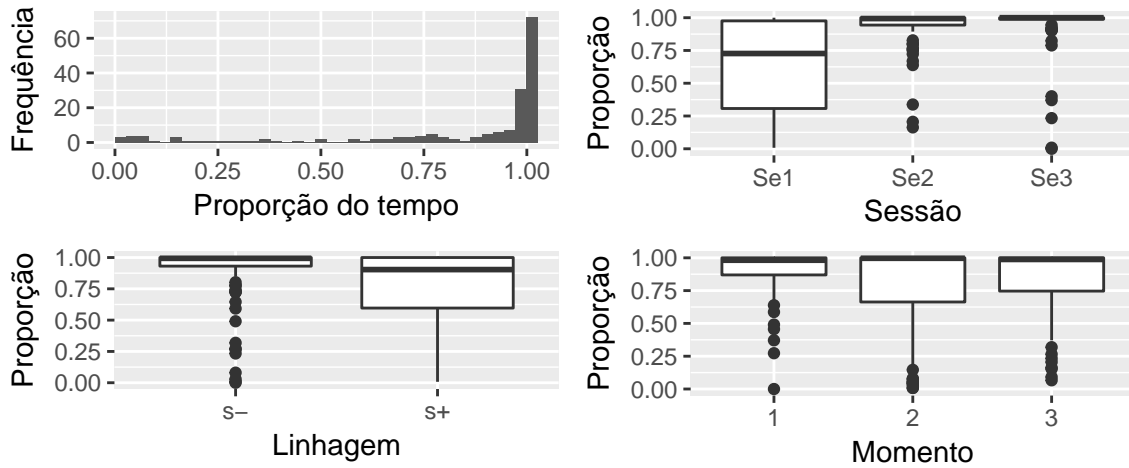


Figura 1: Análise descritiva dos dados sobre a proporção do tempo com orelha levantada ou assimétrica

3.2 Ajustes dos modelos de regressão

O modelo inicialmente ajustado com distribuição beta inflacionada foi especificado da seguinte forma:

$$Y_{ijkl}|u_j, v_{jk} \sim \text{Beta Inflacionada}(\mu_{ijkl}, \sigma, \nu, \tau_{ikl}),$$

em que μ_{ijkl} representa a média de y_{ijkl} , sendo $i = 1, 2$ referente à linhagem; $j = 1, 2, \dots, 20$ aos animais; $k = 1, 2, 3$ à sessão e $l = 1, 2, 3$ ao momento experimental, σ representa o parâmetro de dispersão, ν e τ_{ikl} os parâmetros de inflação.

O preditor da média foi especificado, inicialmente, por:

$$\text{logito}(\mu_{ijkl}) = \alpha^{(1)} + \beta_i^{(1)} + \gamma_k^{(1)} + \theta_l^{(1)} + (\beta\gamma)_{ik}^{(1)} + (\beta\theta)_{il}^{(1)} + (\gamma\theta)_{kl}^{(1)} + u_j + v_{jk};$$

em que $\beta_i^{(1)}$, $\gamma_k^{(1)}$ e $\theta_l^{(1)}$ representam os efeitos de linhagem, sessão e momento; $(\beta\gamma)_{ik}^{(1)}$, $(\beta\theta)_{il}^{(1)}$ e $(\gamma\theta)_{kl}^{(1)}$ as correspondentes interações de segunda ordem; u_i e v_{ij} são os efeitos aleatórios do modelo: $u_j \sim N(0, \sigma_U^2)$ e $v_{jk} \sim N(0, \sigma_V^2)$

O preditor da inflação em 1 fica dado por:

$$\text{logito}(\tau_{ikl}) = \alpha^{(2)} + \beta_i^{(2)} + \gamma_k^{(2)} + \theta_l^{(2)},$$

em que $\beta_i^{(2)}$, $\gamma_k^{(2)}$ e $\theta_l^{(2)}$ os efeitos de linhagem, sessão e momento no parâmetro de inflação. A especificação do modelo beta é a mesma do modelo beta inflacionado, porém sem os parâmetros τ e ν .

A escolha entre as distribuições beta e beta inflacionada foi conduzida por meio de diagnóstico dos ajustes e comportamento dos resíduos. A análise de resíduos nos modelos GAMLSS é realizada com base em resíduos quantílicos aleatorizados (DUNN; SMYTH, 1996), cuja principal característica é apresentarem distribuição Normal padrão caso o modelo ajustado esteja correto.

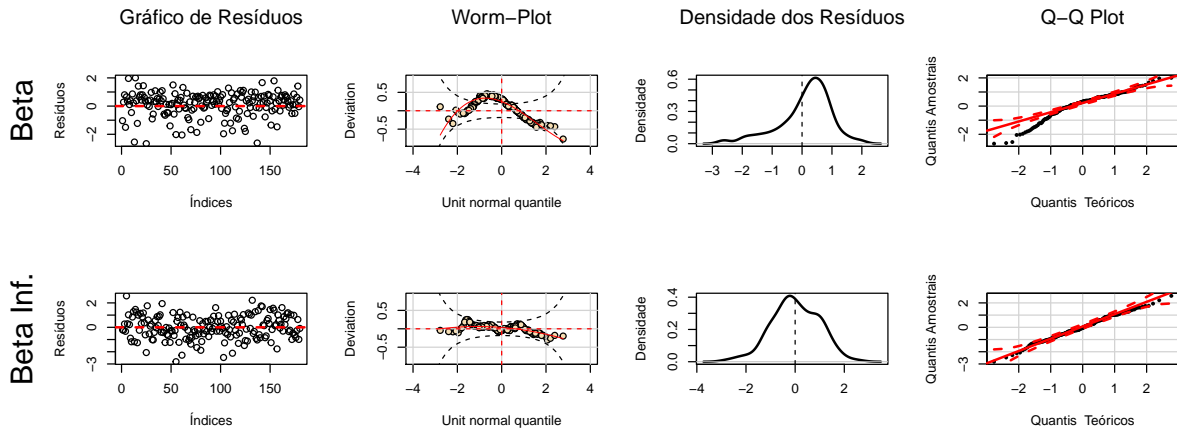


Figura 2: Análise de resíduos dos modelo

Os gráficos de resíduos apresentados na Figura 1 mostram os resíduos dispersos entre ± 2 e sem padrão sistemático para ambas as distribuições. Contudo há considerável assimetria e fuga de normalidade na distribuição dos resíduos para o modelo beta, enquanto que para o modelo beta inflacionado os resíduos apresentam distribuição aparentemente normal.

No geral, a análise de resíduos indica que modelo beta inflacionado está melhor ajustado.

tado aos dados. A fim de obter o modelo mais simples possível que não diferisse estatisticamente do original, foram testados novos modelos, obtidos mediante exclusão de termos do modelo original. Os modelos foram então comparados via teste de razão de verossimilhanças. Com base nos resultados do teste, o modelo mais simples que não difere estatisticamente do modelo completo é aquele sem a interação entre sessão e linhagem ($p=0,98$).

Par.	μ			τ		
	exp(Est.)	IC(95%)	p-valor	exp(Est.)	IC(95%)	p-valor
α	4,11	(2,54;6,66)	<0,01	0,11	(0,04;0,30)	<0,01
γ_{s2}	2,39	(1,21;4,73)	0,01	6,52	(2,46;17,27)	<0,01
γ_{s3}	1,78	(0,81;3,90)	0,15	17,16	(6,31;46,72)	<0,01
θ_{dur}	0,30	(0,15;0,59)	<0,01	2,76	(1,16;6,56)	0,02
θ_{dep}	0,66	(0,31;1,37)	0,26	1,46	(0,62;3,44)	0,39
β_{reat}	0,66	(0,37;1,19)	0,17	0,40	(0,20;0,81)	0,01
$(\gamma\theta)_{se2:dur}$	5,12	(1,84;14,22)	<0,01			
$(\gamma\theta)_{se3:dur}$	1,88	(0,62;5,69)	0,26			
$(\gamma\theta)_{se2:dep}$	1,20	(0,48;3,00)	0,69			
$(\gamma\theta)_{se3:dep}$	1,72	(0,54;5,43)	0,36			
$(\beta\theta)_{reat:dep}$	0,25	(0,11;0,59)	<0,01			
$(\beta\theta)_{reat:dur}$	0,58	(0,25;1,32)	0,20			

Tabela 1: Resumo do modelo ajustado.

Os resultados mostrados na Tabela 1 podem ser interpretados como razões de chances e mostram que na pré escovação, a proporção do tempo que os animais permanecem com as orelhas levantadas ou assimétricas é maior que durante a escovação para animais da linhagem reativa. Na sessão 2, nota-se que a proporção do tempo na postura é maior durante a escovação do que na fase de pré escovação. Nota-se ainda que a proporção do tempo na postura é menor durante a escovação do que antes para animais da linhagem menos reativa na sessão 1. Além disso, a proporção é maior na sessão 2 do que na sessão 1 na fase de pré escovação. As estimativas dos componentes de variância foram 0,000001 e 0,81 para os efeitos aleatórios de animal e animal dentro de sessão, respectivamente.

Para o parâmetro de inflação da distribuição os efeitos de sessão e linhagem se mostraram significativos; os resultados indicam que animais na sessão 2 e 3 são mais propensos a passarem todo o tempo com as orelhas levantadas ou assimétricas que os animais na primeira sessão, além disso, animais reativos tendem a passar todo o tempo na postura analisada.

4 Conclusões

Este trabalho teve como objetivo analisar o efeito de escovação, isolamento e linhagem em ovelhas submetidas à escovação por um humano familiar. A variável sob análise foi a proporção do tempo em que os animais permaneceram com as orelhas levantadas ou assimétricas. Para modelagem considerou-se a classe GAMLSS e as distribuições beta inflacionada e beta. Para viabilidade de aplicação da distribuição beta foi aplicada uma transformação para restringir a resposta ao intervalo (0,1). Os resultados mostraram que o modelo beta inflacionado proporcionou melhor ajuste aos dados. Notou-se que as variáveis sessão, tempo e as interações entre sessão com tempo e tempo com linhagem mostraram-se significativas para o tempo. Quanto à probabilidade do animal passar todo o tempo na postura, as variáveis sessão, tempo e linhagem foram significativos

A metodologia aqui abordada não é comumente utilizada para análise de dados em estudos similares. Contudo, trata-se de um método extremamente flexível no que diz respeito às distribuições implementadas e recursos adicionais para modelagem estatística, tornando o GAMLSS uma útil e eficiente ferramenta para análise de dados de comportamento animal.

Referências

- DUNN, P. K.; SMYTH, G. K. Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, Taylor & Francis, v. 5, n. 3, p. 236–244, 1996. [4](#)
- FERRARI, S.; CRIBARI-NETO, F. Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, Taylor & Francis, v. 31, n. 7, p. 799–815, 2004. [1](#)
- OSPINA, R.; FERRARI, S. L. Inflated beta distributions. *Statistical Papers*, Springer, v. 51, n. 1, p. 111, 2010. [1](#)
- PEREIRA, T. L.; SOUZA, T. C.; CRIBARI-NETO, F. Uma avaliação da eficiência do gasto público nas regiões do Brasil. *Ciência e Natura*, Universidade Federal de Santa Maria, v. 36, 2014. [1](#)
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2017. Disponível em: <https://www.R-project.org/>. [3](#)
- 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. [2](#)
- STASINOPOULOS, D. M.; RIGBY, R. A. et al. Generalized additive models for location scale and shape (gamlss) in R. *Journal of Statistical Software*, v. 23, n. 7, p. 1–46, 2007. [3](#)
- TAMIOSO, P. R. et al. Does emotional reactivity influence behavioral and cardiac responses of ewes submitted to brushing? *Behavioural Processes*, p. np, 2017. [1](#)