

RESEARCH ARTICLE

Teste Wald para avaliação de parâmetros de regressão e dispersão em modelos multivariados de covariância linear generalizada[†]

Author One*¹ | Author Two^{2,3} | Author Three³¹Org Division, Org Name, State name, Country name²Org Division, Org Name, State name, Country name³Org Division, Org Name, State name, Country name**Correspondence**

*Corresponding author name, This is sample corresponding address. Email: authorone@gmail.com

Present Address

This is sample for present address text this is sample for present address text

Summary

250 PALAVRAS

KEYWORDS:

keyword1, keyword2, keyword3, keyword4, keyword5, keyword6

1 | INTRODUCTION

1.1 | Estudos clínicos

Ensaio clínico representa um dos procedimentos mais comuns utilizados para avaliar se um novo tratamento é seguro, efetivo ou até mesmo melhor que um tratamento padrão. Tratam-se de procedimentos amplamente utilizados em pesquisa médica com o objetivo de avaliar intervenções sobre um grupo de indivíduos.

According to¹ "a clinical trial is a planned experiment designed to assess the efficacy of a treatment in man by comparing the outcomes in a group of patients treated with the test with those observed in a comparable group of patients receiving a control treatment where patients in both groups are enrolled, treated, and followed over the same time period".

In this type of study participants are randomly assigned to a treatment or control group (or to multiple treatment groups)² e o objetivo consiste em avaliar se existe diferença significativa entre os grupos.

1.2 | Problemas na área médica que envolvem múltiplas respostas

Often in medical studies, multiple outcomes are taken for a group of patients, por exemplo:³ avalia o effect of community health worker support on clinical outcomes of low-income patients;⁴ estudou o effect of a workplace wellness program on employee health and economic outcomes;⁵ presents the results of a five-year clinical and echocardiographic outcomes from the nordic aortic valve intervention (NOTION);⁶ avalia o effect of Home-Based Exercise and Weight Loss Programs on Breast Cancer-Related Lymphedema Outcomes;⁷ test the association between certain exclusion criteria and trial characteristics. Consequently, the joint analysis of multiple outcomes studies has been of increased interest in the medical and statistical literature.

[†]This is an example for title footnote.

⁰**Abbreviations:** ANA, anti-nuclear antibodies; APC, antigen-presenting cells; IRF, interferon regulatory factor

1.3 | Introdução ao Dataset

Similarly to the aforementioned articles, we are interested in the analysis of multiple outcomes in the context of longitudinal data analysis. The study we shall describe was conducted com o objetivo de avaliar o efeito do uso de probióticos no controle de vícios e compulsões alimentares em pacientes submetidos à cirurgia bariátrica. Trata-se de um problema com duas variáveis resposta: um escore que caracteriza compulsão e o número de sintomas apresentados que caracterizam vício.

Neste estudo, um conjunto de indivíduos foi dividido em dois grupos: um deles recebeu um placebo e o outro recebeu o tratamento. Além disso os indivíduos foram avaliados ao longo tempo; deste modo o delineamento gera observações que não são independentes, já que medidas tomadas em um mesmo indivíduo tendem a ser correlacionadas.

1.4 | McGLM

É importante notar que trata-se de um problema com duas variáveis resposta, não gaussianas, com observações não independentes, tendo em vista que medidas tomadas em um mesmo indivíduo são correlacionadas. Portanto trata-se de um problema que técnicas de modelagem tradicionais seriam de difícil aplicação. Se faz necessario uso de alguma metodologia que comporte os requisitos do problema.

In this article, we adopt the multivariate covariance generalized linear models (McGLM) framework⁸, which provides um ambiente para modelar múltiplas respostas não gaussianas simultaneamente com uma flexible and interpretable modelling of the covariance structure. The within outcomes covariance matrix is specified for each marginal outcome using a linear combination of known matrices, while the joint covariance matrix is specified using the generalized Kronecker product^{9,8}.

1.5 | Importância de testes de hipóteses em regressão

Quando trabalha-se com modelos de regressão, um interesse comum aos analistas é o de verificar se a ausência de determinada variável explicativa do modelo geraria uma perda no ajuste. Deste modo, uma conjectura de interesse é avaliar se há evidência suficiente nos dados para afirmar que determinada variável explicativa não possui efeito sobre a resposta. Isto é feito por meio de testes de hipóteses. Neste contexto, três testes de hipóteses são comuns: o teste da razão de verossimilhanças, o teste Wald e o teste do multiplicador de lagrange, também conhecido como teste escore.¹⁰ descreve a formulação geral dos três testes. Todos eles são baseados na função de verossimilhança dos modelos.

1.6 | ANOVA e MANOVA

Para o caso dos modelos lineares tradicionais existem técnicas como a análise de variância (ANOVA), proposta inicialmente por¹¹. Segundo¹², a ANOVA é um dos métodos estatísticos mais amplamente usados para testar hipóteses e que está presente em praticamente todos os materiais introdutórios de estatística. O objetivo da técnica é a avaliação do efeito de cada uma das variáveis explicativas sobre a resposta. Isto é feito por meio da comparação via testes de hipóteses entre modelos com e sem cada uma das variáveis explicativas. Logo, tal procedimento permite que seja possível avaliar se a retirada de cada uma das variáveis gera um modelo significativamente pior quando comparado ao modelo com a variável. Para o caso multivariado estende-se a técnica de análise de variância (ANOVA) para a análise de variância multivariada¹³, a MANOVA. E dentre os testes de hipóteses multivariados já discutidos na literatura, destacam-se o λ de Wilk's¹⁴, traço de Hotelling-Lawley^{15,16}, traço de Pillai¹⁷ e maior raiz de Roy¹⁸.

1.7 | Testes de comparações múltiplas

Complementar às ANOVAs e MANOVAs estão os testes de comparações múltiplas. Tais procedimentos são utilizados quando a análise de variância aponta como conclusão a existência de efeito significativo dos parâmetros associados a uma variável categórica, ou seja, há ao menos uma diferença significativa entre os níveis de um fator. Com isso, o teste de comparações múltiplas é mais um procedimento baseado em testes de hipóteses, utilizado para determinar onde estão estas diferenças. Por exemplo, suponha que há no modelo uma variável categórica X de três níveis: A, B e C. A análise de variância mostrará se há efeito da variável X no modelo, isto é, se os valores da resposta estão associados aos níveis de X , contudo este resultado não nos mostrará se os valores da resposta diferem de A para B, ou de A para C, ou ainda se B difere de C. Para detectar tais diferenças empregam-se os testes de comparações múltiplas. Dentre os testes discutidos na literatura encontram-se o teste de Dunnett,

Tukey, t de student (LSD), Scott-Knott, dentre outros.¹⁹ discute diversos procedimentos para fins de comparações múltiplas. Já²⁰ trata de procedimentos de comparações múltiplas em modelos lineares.

1.8 | Objetivo do trabalho

Contudo, para o contexto de múltiplas respostas não gaussianas, são escassas as alternativas para testes de hipóteses. Com isso, por se tratar de uma classe de modelos flexível e com alto poder de aplicação a problemas práticos, nosso objetivo geral é o desenvolvimento de testes de hipóteses para os McGLMs. Temos os seguintes objetivos específicos: propor a utilização do teste Wald para realização de testes de hipóteses gerais sobre parâmetros de regressão e dispersão de McGLMs, possibilitando também a geração de quadros de análises de variância, análises de variância multivariadas e testes de comparações múltiplas para problemas de regressão com múltiplas respostas não gaussianas. Outro objetivo é avaliar as propriedades e comportamento dos testes propostos com base em estudos de simulação e avaliar o potencial de aplicação das metodologias discutidas com base na aplicação ao já mencionado conjuntos de dados. No problema em questão, técnicas de modelagem tradicionais seriam de difícil aplicação. Contudo trata-se de um problema de possível análise via McGLM e testes de hipóteses podem ser empregados para avaliar o efeito da interação entre momento e uso do probiótico sobre vício e compulsão alimentar.

1.9 | Organização do trabalho (seções)

We present the data set in section 2. In section 3 we present a revisão da estrutura geral e estimação dos parâmetros de um McGLM, baseado nas ideias de⁸. In section 4 é apresentada nossa proposta com os detalhes do teste Wald para avaliar suposições sobre parâmetros de um McGLM. In section 5 apresentamos os resultados da avaliação de performance do teste proposto com base em um estudo de simulação. In the section 6 we apply the model to the data and present the main results of uso de probióticos no controle de vícios e compulsões alimentares em pacientes sujeitos a cirurgia bariátrica. Finally, the main results are discussed in section 7, including some directions for future investigations.

2 | DATASET

2.1 | Contexto

Trata-se de um ensaio clínico randomizado, duplo-cego, controlado por placebo, realizado com pacientes submetidos ao Bypass Gástrico em Y de Roux (RYGB) no período de abril de 2018 a dezembro de 2019. O estudo foi aprovado pelo Comitê de Ética em Pesquisa da Pontifícia Universidade Católica do Paraná (PUCPR) (nº 4.252.808) e registrado pelo Registro Brasileiro de Ensaios Clínicos - REBEC (nºRBR-4x3gqp). A pesquisa foi explicada a cada participante antes de sua participação e, daqueles que concordaram, foi obtido o consentimento informado por escrito.

A divisão dos grupos (placebo ou probiótico) foi feita de forma aleatória. Os critérios de inclusão dos indivíduos no estudo foram: adultos (18-59 anos) que fariam RYGB, com índice de massa corporal (IMC) ≥ 35 kg/m² e que não usaram antibióticos antes da cirurgia.

Foram retirados do estudo pacientes que foram submetidos a outras técnicas cirúrgicas ou reoperação, tiveram complicações pós-cirúrgicas, fizeram antibioticoterapia concomitante ao uso de probiótico/placebo ou não usaram os comprimidos de probiótico/placebo por mais de nove dias consecutivos.

Ambos os grupos receberam as mesmas orientações alimentares após a cirurgia, foram acompanhados pela mesma equipe cirúrgica (médico, nutricionista e psicólogo) e tiveram o mesmo número de consultas pré-agendadas antes e após a cirurgia, seguindo o protocolo estabelecido pela instituição onde o estudo foi realizado.

No sétimo dia de pós-operatório, os participantes foram orientados a ingerir dois comprimidos mastigáveis/dia de placebo ou comprimido probiótico Flora Vantage, 5 bilhões de *Lactobacillus acidophilus* NCFM ®Strain e 5 bilhões de *Bifidobacterium lactis* Bi-07 ® da Bariatric Advantage (Aliso Viejo, CA, EUA) por 90 dias.

Os indivíduos foram avaliados em 3 momentos. A primeira avaliação (T0) foi realizada na primeira consulta, aproximadamente 10 dias antes da cirurgia. As avaliações de acompanhamento foram realizadas aproximadamente três meses (T1) e um ano de pós-operatório (T2). Foram realizadas avaliações clínicas e antropométricas, bem como os questionários autoaplicáveis foram entregues aos participantes a cada encontro.

A avaliação de compulsão alimentar foi feita com base na escala de compulsão alimentar (BES), uma das ferramentas mais usadas para medir a compulsão alimentar. Trata-se de um questionário em formato de escala likert de 16 itens, elaborado de acordo com o Manual Diagnóstico e Estatístico de Transtornos Mentais (3ª edição)² por ?. Os indivíduos foram orientados a selecionar a opção que melhor representasse sua resposta e o escore final foi a soma dos pontos de cada item, este escore varia de 0 a 46.

Para avaliação de vício alimentar foi utilizada a escala de vício alimentar (YFAS), um questionário que busca detectar sintomas de comportamentos alimentares aditivos. O YFAS foi baseado nos critérios de dependência de substâncias do Manual Diagnóstico e Estatístico IV – Revisão de Texto (DSM-IV-TR)² e endossado para alimentos altamente processados. Este questionário foi desenvolvido por ?. O questionário é uma combinação de 25 opções em escala Likert e a opção de avaliação utilizada foi o número de sintomas de vício.

2.2 | Desenho experimental e coleta de dados

A amostra final é formada por 71 indivíduos, dos quais 33 pertencem ao grupo placebo e 38 ao grupo tratamento. Se todos estes indivíduos fossem avaliados nos 3 momentos definidos no estudo, o conjunto de dados teria 213 observações. Contudo, ao longo do estudo, diversos indivíduos não compareceram às consultas, o que faz com que haja dados faltantes no conjunto de dados. Após tratamento dos dados e exclusão de observações faltantes restaram 184 observações.

É importante notar que trata-se de um problema com duas variáveis resposta: um escore que caracteriza compulsão e o número de sintomas apresentados que caracterizam vício. Além disso, as observações não são independentes, tendo em vista que medidas tomadas em um mesmo indivíduo são correlacionadas. Portanto trata-se de um problema que técnicas de modelagem tradicionais seriam de difícil aplicação mas é um cenário ideal para resolução via McGLM e no qual testes de hipóteses são de extrema importância para avaliar o efeito da interação entre momento e uso do probiótico sobre vício e compulsão alimentar.

2.3 | Conjunto de dados

Para fins de análise, o escore que caracteriza compulsão e o número de sintomas apresentados que caracterizam vício foram transformados para a escala unitária, considerando que tratam-se de variáveis restritas. O objetivo da análise é avaliar o efeito de momento e grupo nas métricas de vício e compulsão. O conjunto de dados contém as seguintes variáveis:

- id: variável identificadora de indivíduo.
- momento: variável identificadora de momento (T0, T1, T2).
- grupo: variável identificadora de grupo (placebo, probiótico)
- YFAS: proporção de sintomas que caracterizam vício.
- BES: proporção de escore de compulsão.

2.4 | Análise exploratória

A análise gráfica apresentada na ?? mostra, em (a) e (d) que ambas as variáveis de interesse apresentam considerável assimetria à direita. Os boxplots das métricas em função de grupo apresentados em (b) e (e) mostram sensíveis diferenças entre o grupo placebo e probiótico para ambas as respostas. Já os boxplots das métricas em função dos momentos de avaliação, apresentados em (c) e (f), evidenciam que para ambas as métricas os valores eram mais altos no momento T0, havendo considerável redução no momento T1. Quando comparamos T1 e T2, para YFAS parece que há um leve aumento na última avaliação; já para BES, T1 e T2 não parecem diferir.

FIGURA

Ainda de forma exploratória podemos avaliar o comportamento das métricas de vício e compulsão por meio da avaliação das medidas descritivas por momento e por grupo, apresentadas na Table 1. É possível verificar a redução de indivíduos ao longo do tempo, algo comum em estudos prospectivos. Quanto às medidas, nota-se que ambos os grupos (placebo ou probiótico) apresentam médias mais altas no momento T0 do que nos demais momentos. Portanto, existe uma clara redução das métricas quando comparado ao pré operatório. Quando comparamos os momentos pós-operatório (T1 e T2) verificamos que as medidas

de YFAS para o grupo placebo e probiótico e BES no grupo placebo apresentam, em média, um aumento das métricas no último momento de avaliação. Já as medidas de BES no grupo probiótico apresentam queda.

3 | MCGLM

Considere $\mathbf{Y}_{N \times R} = \{\mathbf{Y}_1, \dots, \mathbf{Y}_R\}$ uma matriz de variáveis resposta e $\mathbf{M}_{N \times R} = \{\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_R\}$ uma matriz de valores esperados. A matriz de variância e covariância para cada resposta $r, r = 1, \dots, R$ é denotada por $\boldsymbol{\Sigma}_r$, tem dimensão $N \times N$. Além disso, é necessária uma matriz de correlação $\boldsymbol{\Sigma}_b$, de ordem $R \times R$, que descreve a correlação entre as variáveis resposta.

Os McGLMs⁸ são definidos por:

$$\begin{aligned} E(\mathbf{Y}) &= \mathbf{M} = \{g_1^{-1}(\mathbf{X}_1 \boldsymbol{\beta}_1), \dots, g_R^{-1}(\mathbf{X}_R \boldsymbol{\beta}_R)\} \\ \text{Var}(\mathbf{Y}) &= \mathbf{C} = \boldsymbol{\Sigma}_R \overset{G}{\otimes} \boldsymbol{\Sigma}_b, \end{aligned} \quad (1)$$

em que $\boldsymbol{\Sigma}_R \overset{G}{\otimes} \boldsymbol{\Sigma}_b = \text{Bdiag}(\tilde{\boldsymbol{\Sigma}}_1, \dots, \tilde{\boldsymbol{\Sigma}}_R)(\boldsymbol{\Sigma}_b \otimes \mathbf{I})\text{Bdiag}(\tilde{\boldsymbol{\Sigma}}_1^\top, \dots, \tilde{\boldsymbol{\Sigma}}_R^\top)$ é o produto generalizado de Kronecker (author?)⁹, a matriz $\tilde{\boldsymbol{\Sigma}}_r$ denota a matriz triangular inferior da decomposição de Cholesky da matriz $\boldsymbol{\Sigma}_r$. O operador $\text{Bdiag}()$ denota a matriz bloco-diagonal e \mathbf{I} uma matriz identidade $N \times N$. As funções $g_r()$ são as tradicionais funções de ligação. \mathbf{X}_r denota uma matriz de delineamento $N \times k_r$. $\boldsymbol{\beta}_r$ denota um vetor $k_r \times 1$ de parâmetros de regressão.

Para variáveis resposta contínuas, binárias, binomiais, proporções ou índices a matriz de variância e covariância $\boldsymbol{\Sigma}_r$ é dada por:

$$\boldsymbol{\Sigma}_r = \mathbf{V}(\boldsymbol{\mu}_r; p_r)^{1/2} (\boldsymbol{\Omega}(\boldsymbol{\tau}_r)) \mathbf{V}(\boldsymbol{\mu}_r; p_r)^{1/2}. \quad (2)$$

No caso de variáveis resposta que sejam contagens a matriz de variância e covariância para cada variável resposta fica dada por:

$$\boldsymbol{\Sigma}_r = \text{diag}(\boldsymbol{\mu}_r) + \mathbf{V}(\boldsymbol{\mu}_r; p_r)^{1/2} (\boldsymbol{\Omega}(\boldsymbol{\tau}_r)) \mathbf{V}(\boldsymbol{\mu}_r; p_r)^{1/2}, \quad (3)$$

em que $\mathbf{V}(\boldsymbol{\mu}_r; p_r) = \text{diag}(\vartheta(\boldsymbol{\mu}_r; p_r))$ denota uma matriz diagonal na qual as entradas são dadas pela função de variância $\vartheta(\cdot; p_r)$ aplicada aos elementos do vetor $\boldsymbol{\mu}_r$.

Diferentes escolhas de funções de variância $\vartheta(\cdot; p_r)$ implicam em diferentes suposições a respeito da distribuição da variável resposta. No pacote ‘mcglm’ existem 3 opções de funções de variância: função de variância potência, função de dispersão Poisson–Tweedie e função de variância binomial.

A função de variância potência caracteriza a família Tweedie de distribuições, é dada por $\vartheta(\cdot; p_r) = \mu_r^{p_r}$, na qual destacam-se a distribuições: Normal ($p = 0$), Poisson ($p = 1$), gama ($p = 2$) e Normal inversa ($p = 3$) ? ?.

A função de dispersão Poisson–Tweedie ? visa contornar a inflexibilidade da utilização da função de variância potência para respostas que caracterizam contagens. A função de dispersão é dada por $\vartheta(\cdot; p) = \mu + \tau \mu^p$ em que τ é o parâmetro de dispersão. Temos assim uma rica classe de modelos para lidar com respostas que caracterizam contagens, uma vez que muitas distribuições importantes aparecem como casos especiais, tais como: Hermite ($p = 0$), Neyman tipo A ($p = 1$), binomial negativa ($p = 2$) e Poisson–inversa gaussiana ($p = 3$).

Por fim, a função de variância binomial, dada por $\vartheta(\cdot; p_r) = \mu_r^{p_{r1}}(1 - \mu_r)^{p_{r2}}$ é indicada quando a variável resposta é binária, restrita a um intervalo ou quando tem-se o número de sucessos em um número de tentativas.

É possível notar que o parâmetro de potência p aparece em todas as funções de variância discutidas. Este parâmetro tem especial importância pois trata-se de um índice que distingue diferentes distribuições de probabilidade importantes no contexto de modelagem e, por esta razão, pode ser utilizado como uma ferramenta para seleção automática da distribuição de probabilidade que mais se adequa ao problema.

A matriz de dispersão $\boldsymbol{\Omega}(\boldsymbol{\tau})$ descreve a parte da covariância dentro de cada variável resposta que não depende da estrutura média, isto é, a estrutura de correlação entre as observações da amostra. Baseando-se nas ideias de ? e ?, (author?)⁸ propuseram modelar a matriz de dispersão através de um preditor linear matricial combinado com uma função de ligação de covariância dada por:

$$h\{\boldsymbol{\Omega}(\boldsymbol{\tau}_r)\} = \tau_{r0} \mathbf{Z}_0 + \dots + \tau_{rD} \mathbf{Z}_D, \quad (4)$$

em que $h()$ é a função de ligação de covariância, \mathbf{Z}_{rd} com $d = 0, \dots, D$ são matrizes que representam a estrutura de covariância presente em cada variável resposta r e $\boldsymbol{\tau}_r = (\tau_{r0}, \dots, \tau_{rD})$ é um vetor $(D + 1) \times 1$ de parâmetros de dispersão.

No pacote ‘mcglm’ estão disponíveis 3 funções de ligação de covariância: identidade, inversa e exponencial-matriz. A especificação da função de ligação de covariância é discutida por ? e é possível selecionar combinações de matrizes para se obter os mais conhecidos modelos da literatura para dados longitudinais, séries temporais, dados espaciais e espaço-temporais. Maiores detalhes são discutidos por ?.

Deste modo, os McGLM configuram uma estrutura geral para análise via modelos de regressão para dados não gaussianos com múltiplas respostas em que não se faz suposições quanto à independência das observações. A classe é definida por 3 funções (de ligação, de variância e de covariância) além de um preditor linear e um preditor linear matricial para cada resposta sob análise.

3.1 | Estimação e inferência

Os McGLMs são ajustados baseados no método de funções de estimação descritos em detalhes por (author?)⁸ e ?. Nesta seção é apresentada uma visão geral do algoritmo e da distribuição assintótica dos estimadores baseados em funções de estimação.

As suposições de segundo momento dos McGLM permitem a divisão dos parâmetros em dois conjuntos: $\theta = (\beta^\top, \lambda^\top)^\top$. Desta forma, $\beta = (\beta_1^\top, \dots, \beta_R^\top)^\top$ é um vetor $K \times 1$ de parâmetros de regressão e $\lambda = (\rho_1, \dots, \rho_{R(R-1)/2}, p_1, \dots, p_R, \tau_1^\top, \dots, \tau_R^\top)^\top$ é um vetor $Q \times 1$ de parâmetros de dispersão. Além disso, $\mathcal{Y} = (\mathbf{Y}_1^\top, \dots, \mathbf{Y}_R^\top)^\top$ denota o vetor empilhado de ordem $NR \times 1$ da matriz de variáveis resposta \mathbf{Y}_{NR} e $\mathcal{M} = (\mu_1^\top, \dots, \mu_R^\top)^\top$ denota o vetor empilhado de ordem $NR \times 1$ da matriz de valores esperados \mathbf{M}_{NR} .

Para estimação dos parâmetros de regressão é utilizada a função quasi-score², representada por

$$\psi_\beta(\beta, \lambda) = \mathbf{D}^\top \mathbf{C}^{-1}(\mathcal{Y} - \mathcal{M}), \quad (5)$$

em que $\mathbf{D} = \nabla_\beta \mathcal{M}$ é uma matriz $NR \times K$, e ∇_β denota o operador gradiente. Utilizando a função quasi-score a matriz $K \times K$ de sensibilidade de ψ_β é dada por

$$\mathbf{S}_\beta = E(\nabla_\beta \psi_\beta) = -\mathbf{D}^\top \mathbf{C}^{-1} \mathbf{D}, \quad (6)$$

enquanto que a matriz $K \times K$ de variabilidade de ψ_β é escrita como

$$\mathbf{V}_\beta = \text{VAR}(\psi_\beta) = \mathbf{D}^\top \mathbf{C}^{-1} \mathbf{D}. \quad (7)$$

Para os parâmetros de dispersão é utilizada a função de estimação de Pearson, definida da forma

$$\psi_{\lambda_i}(\beta, \lambda) = \text{tr}(\mathbf{W}_{\lambda_i}(\mathbf{r}^\top \mathbf{r} - \mathbf{C})), i = 1, \dots, Q, \quad (8)$$

em que $\mathbf{W}_{\lambda_i} = -\frac{\partial \mathbf{C}^{-1}}{\partial \lambda_i}$ e $\mathbf{r} = (\mathcal{Y} - \mathcal{M})$. A entrada (i, j) da matriz de sensibilidade $Q \times Q$ de ψ_λ é dada por

$$\mathbf{S}_{\lambda_{ij}} = E\left(\frac{\partial}{\partial \lambda_i} \psi_{\lambda_j}\right) = -\text{tr}(\mathbf{W}_{\lambda_i} \mathbf{C} \mathbf{W}_{\lambda_j} \mathbf{C}). \quad (9)$$

Já a entrada (i, j) da matriz de variabilidade $Q \times Q$ de ψ_λ é definida por

$$\mathbf{V}_{\lambda_{ij}} = \text{Cov}(\psi_{\lambda_i}, \psi_{\lambda_j}) = 2\text{tr}(\mathbf{W}_{\lambda_i} \mathbf{C} \mathbf{W}_{\lambda_j} \mathbf{C}) + \sum_{l=1}^{NR} k_l^{(4)} (\mathbf{W}_{\lambda_i})_{ll} (\mathbf{W}_{\lambda_j})_{ll}, \quad (10)$$

em que $k_l^{(4)}$ denota a quarta cumulante de \mathcal{Y}_l . No processo de estimação dos McGLM são usadas as versões empíricas.

Para se levar em conta a covariância entre os vetores β e λ , (author?)⁸ obtiveram as matrizes de sensibilidade e variabilidade cruzadas, denotadas por $\mathbf{S}_{\lambda\beta}$, $\mathbf{S}_{\beta\lambda}$ e $\mathbf{V}_{\lambda\beta}$, mais detalhes em⁸. As matrizes de sensibilidade e variabilidade conjuntas de ψ_β e ψ_λ são denotados por

$$\mathbf{S}_\theta = \begin{bmatrix} \mathbf{S}_\beta & \mathbf{S}_{\beta\lambda} \\ \mathbf{S}_{\lambda\beta} & \mathbf{S}_\lambda \end{bmatrix} \text{ e } \mathbf{V}_\theta = \begin{bmatrix} \mathbf{V}_\beta & \mathbf{V}_{\lambda\beta} \\ \mathbf{V}_{\lambda\beta} & \mathbf{V}_\lambda \end{bmatrix}. \quad (11)$$

Seja $\hat{\theta} = (\hat{\beta}^\top, \hat{\lambda}^\top)^\top$ o estimador baseado em funções de estimação de θ . Então, a distribuição assintótica de $\hat{\theta}$ é

$$\hat{\theta} \sim N(\theta, \mathbf{J}_\theta^{-1}), \quad (12)$$

em que \mathbf{J}_θ^{-1} é a inversa da matriz de informação de Godambe, dada por $\mathbf{J}_\theta^{-1} = \mathbf{S}_\theta^{-1} \mathbf{V}_\theta \mathbf{S}_\theta^{-\top}$, em que $\mathbf{S}_\theta^{-\top} = (\mathbf{S}_\theta^{-1})^\top$.

Para resolver o sistema de equações $\psi_\beta = 0$ e $\psi_\lambda = 0$ faz-se uso do algoritmo Chaser modificado, proposto por ?, que fica definido como

$$\begin{aligned}\beta^{(i+1)} &= \beta^{(i)} - S_\beta^{-1} \psi_\beta(\beta^{(i)}, \lambda^{(i)}), \\ \lambda^{(i+1)} &= \lambda^{(i)} \alpha S_\lambda^{-1} \psi_\lambda(\beta^{(i+1)}, \lambda^{(i)}).\end{aligned}\quad (13)$$

4 | TESTE WALD PARA MCGLMS

Este capítulo é dedicado à apresentação de nossa proposta: o uso do teste Wald para avaliação dos parâmetros de McGLMs. Vale lembrar que nos McGLMs existem parâmetros de regressão, dispersão, potência e correlação e que, neste trabalho, estamos interessados na avaliação dos parâmetros de regressão e dispersão. Cada conjunto de parâmetros possui uma interpretação prática bastante relevante de tal modo que por meio dos parâmetros de regressão é possível identificar as explicativas relevantes, por meio dos parâmetros de dispersão é possível avaliar o impacto da correlação entre unidades do conjunto de dados, por meio dos parâmetros de potência é possível identificar qual distribuição de probabilidade melhor se adequa ao problema de acordo com a função de variância e por meio dos parâmetros de correlação é possível avaliar a associação entre respostas.

É importante notar que como temos que as estimativas dos parâmetros seguem distribuição normal, qualquer forma quadrática obtida por meio do estimador ou combinações lineares (tal como a estatística de Wald) é também uma forma quadrática que tem distribuição qui-quadrado. Portanto, nossa proposta não é uma simples adaptação, mas é um procedimento assintoticamente válido com respaldo na teoria da distribuição de formas quadráticas. Mais sobre a distribuição de formas quadráticas pode ser visto nos trabalhos de ?, ? e ?.

4.1 | Hipóteses e estatística de teste

Um McGLM possui vetor de parâmetros $\theta = (\beta^\top, \lambda^\top)^\top$, em que $\beta = (\beta_1^\top, \dots, \beta_R^\top)^\top$ é um vetor $K \times 1$ de parâmetros de regressão e $\lambda = (\rho_1, \dots, \rho_{R(R-1)/2}, p_1, \dots, p_R, \tau_1^\top, \dots, \tau_R^\top)^\top$ é um vetor $Q \times 1$ de parâmetros de dispersão.

Considere θ^* o vetor $h \times 1$ de parâmetros desconsiderando os parâmetros de correlação, ou seja, θ^* refere-se apenas a parâmetros de regressão, dispersão ou potência. As estimativas de dos parâmetros de θ^* são dadas por $\hat{\theta}^*$. De maneira similar, considere J^{*-1} a inversa da matriz de informação de Godambe desconsiderando os parâmetros de correlação, de dimensão $h \times h$.

Seja L uma matriz de especificação de hipóteses a serem testadas, de dimensão $s \times h$ e c um vetor de dimensão $s \times 1$ com os valores sob hipótese nula, em que s denota o número de restrições. As hipóteses a serem testadas podem ser escritas como:

$$H_0 : L\theta^* = c \text{ vs } H_1 : L\theta^* \neq c. \quad (14)$$

Desta forma, a generalização da estatística do teste Wald para verificar a validade de uma hipótese sobre parâmetros de um McGLM fica dada por:

$$W = (L\hat{\theta}^* - c)^T (L J^{*-1} L^T)^{-1} (L\hat{\theta}^* - c),$$

em que $W \sim \chi_s^2$, ou seja, independente do número de parâmetros nas hipóteses, a estatística de teste W é um único valor que segue assintoticamente distribuição χ^2 com graus de liberdade dados pelo número de restrições, isto é, o número de linhas da matriz L , denotado por s .

Em geral, cada coluna da matriz L corresponde a um dos h parâmetros de θ^* e cada linha a uma restrição. Sua construção consiste basicamente em preencher a matriz com 0, 1 e eventualmente -1 de tal modo que o produto $L\theta^*$ represente corretamente as hipóteses de interesse. A correta especificação de L permite testar qualquer parâmetro individualmente ou até mesmo formular hipóteses para diversos parâmetros.

Em um contexto prático, após a obtenção das estimativas dos parâmetros do modelo podemos estar interessados em três tipos de hipóteses: a primeira delas diz respeito a quando o interesse está em avaliar se existe evidência que permita afirmar que apenas um único parâmetro é igual a um valor postulado; a segunda delas ocorre quando há interesse em avaliar se existe evidência para afirmar que um conjunto de parâmetros é igual a um vetor de valores postulado; já a terceira hipótese diz respeito a situações em que o analista está interessado em saber se a diferença entre os efeitos de duas variáveis é igual a 0, isto é, se o efeito das variáveis sobre a resposta é o mesmo.

Para fins de ilustração dos tipos de hipóteses mencionadas, considere a situação em que deseja-se investigar se uma variável numérica X_1 possui efeito sobre duas variáveis respostas, denotadas por Y_1 e Y_2 . Para tal tarefa coletou-se uma amostra com

N observações e para cada observação registrou-se os valores de X_1 , Y_1 e Y_2 . Com base nos dados coletados ajustou-se um McGLM bivariado, com preditor dado por:

$$g_r(\mu_r) = \beta_{r0} + \beta_{r1}X_1, r = 1, 2, \quad (15)$$

em que o índice r denota a variável resposta, $r = 1, 2$; β_{r0} representa o intercepto; β_{r1} um parâmetro de regressão associado a uma variável X_1 . Considere que cada resposta possui apenas um parâmetro de dispersão τ_{r0} e que os parâmetros de potência foram fixados. Portanto, trata-se de um problema em que há duas variáveis resposta e apenas uma variável explicativa. Considere que as unidades em estudo são independentes, logo $Z_0 = I$.

Neste cenário poderiam ser perguntas de interesse: existe efeito da variável X_1 apenas sobre a primeira resposta? Ou apenas sobre a segunda resposta? É possível que a variável X_1 possua efeito sobre as duas respostas ao mesmo tempo? É possível que o efeito da variável seja o mesmo para ambas as respostas? Todas essas perguntas podem ser respondidas por meio de testes de hipóteses sobre os parâmetros do modelo e especificadas por meio da Equation 14. Nas subseções a seguir são apresentados os elementos para responder a cada uma destas perguntas.

4.1.1 | Exemplo 1: hipótese para um único parâmetro

Considere o primeiro tipo de hipótese: há interesse em avaliar se existe efeito da variável X_1 apenas sobre primeira resposta. A hipótese pode ser escrita da seguinte forma:

$$H_0 : \beta_{11} = 0 \text{ vs } H_1 : \beta_{11} \neq 0. \quad (16)$$

Esta mesma hipótese pode ser reescrita na notação mais conveniente para aplicação da estatística do teste Wald, tal como na Equation 14 em que:

- $\theta^{*T} = [\beta_{10} \ \beta_{11} \ \beta_{20} \ \beta_{21} \ \tau_{11} \ \tau_{21}]$.
- $L = [0 \ 1 \ 0 \ 0 \ 0 \ 0]$.
- $c = [0]$, é o valor sob hipótese nula.

Note que o vetor θ^* possui seis elementos, consequentemente a matriz L contém seis colunas (uma para cada elemento) e uma linha, pois apenas um único parâmetro está sendo testado. Essa única linha é composta por zeros, exceto a coluna referente ao parâmetro de interesse que recebe 1. É simples verificar que o produto $L\theta^*$ representa a hipótese de interesse inicialmente postulada na Equation 16. Com isso, a distribuição assintótica do teste é χ_1^2 .

4.1.2 | Exemplo 2: hipótese para múltiplos parâmetros

Suponha agora que o interesse neste problema genérico não seja mais testar o efeito da variável explicativa apenas em uma resposta. Suponha que o interesse seja avaliar se existe evidência suficiente para afirmar que há efeito da variável explicativa X_1 em ambas as respostas simultaneamente. Neste caso teremos que testar 2 parâmetros: β_{11} , que associa X_1 à primeira resposta; e β_{21} , que associa X_1 à segunda resposta. Podemos escrever a hipótese da seguinte forma:

$$H_0 : \beta_{r1} = 0 \text{ vs } H_1 : \beta_{r1} \neq 0, \quad (17)$$

ou, de forma equivalente:

$$H_0 : \begin{pmatrix} \beta_{11} \\ \beta_{21} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ vs } H_1 : \begin{pmatrix} \beta_{11} \\ \beta_{21} \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

As hipóteses na forma da Equation 14 possui os seguintes elementos:

- $\theta^{*T} = [\beta_{10} \ \beta_{11} \ \beta_{20} \ \beta_{21} \ \tau_{11} \ \tau_{21}]$.
- $L = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$.
- $c = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, é o valor sob hipótese nula.

O vetor θ^* se mantém com seis elementos e a matriz L com seis colunas. Neste caso estamos testando dois parâmetros, portanto a matriz L possui duas linhas. Novamente, essas linhas são compostas por zeros, exceto nas colunas referentes ao parâmetro de interesse. É simples verificar que o produto $L\theta^*$ representa a hipótese de interesse inicialmente postulada na Equation 17. Com isso, a distribuição assintótica do teste é χ^2_2 .

4.1.3 | Exemplo 3: hipótese de igualdade de parâmetros

Suponha que a hipótese de interesse não envolve testar se o valor do parâmetro é igual a um valor postulado mas sim verificar se, no caso deste problema genérico, o efeito da variável X_1 é o mesmo independente da resposta. Nesta situação formaríamos uma hipótese de igualdade entre os parâmetros ou, em outros termos, se a diferença dos efeitos é nula:

$$H_0 : \beta_{11} - \beta_{21} = 0 \text{ vs } H_1 : \beta_{11} - \beta_{21} \neq 0, \quad (18)$$

na notação da Equation 14 os elementos das hipóteses são:

- $\theta^{*T} = [\beta_{10} \ \beta_{11} \ \beta_{20} \ \beta_{21} \ \tau_{11} \ \tau_{21}]$.
- $L = [0 \ 1 \ 0 \ -1 \ 0 \ 0]$.
- $c = [0]$, é o valor sob hipótese nula.

Como existe apenas uma hipótese, a matriz L possui apenas uma linha. Para a matriz L ser corretamente especificada no caso de uma hipótese de igualdade é necessário colocar 1 na coluna referente a um parâmetro, e -1 na coluna referente ao outro parâmetro, de tal modo que o produto $L\theta^*$ represente a hipótese de interesse inicialmente postulada, neste caso o produto $L\theta^*$ gera exatamente a mesma hipótese especificada na Equation 18 e a distribuição assintótica do teste é χ^2_1 .

4.1.4 | Exemplo 4: hipótese sobre parâmetros de regressão ou dispersão para respostas sob mesmo preditor

A Equation 15 descreve um modelo bivariado genérico. É importante notar que neste exemplo ambas as respostas estão sujeitas ao mesmo preditor. Na prática, quando se trata dos McGLMs, preditores diferentes podem ser especificados entre variáveis respostas. Deste modo, o que foi exposto na subsubsection 4.1.2 serve para qualquer caso em que haja interesse em testar hipóteses sobre mais de um parâmetro do modelo, sejam eles na mesma resposta ou em respostas diferentes, independente dos preditores entre respostas.

Contudo, nos casos em que as respostas estão sujeitas a preditores idênticos e as hipóteses sobre os parâmetros não se alteram de resposta para resposta, uma especificação alternativa do procedimento é utilizando o produto Kronecker para testar uma mesma hipótese sobre múltiplas respostas tal como utilizado em ?.

Suponha que, neste exemplo, as hipóteses de interesse seguem sendo escritas tal como na Equation 17. Contudo, como se trata de um modelo bivariado com mesmo preditor para as duas respostas, a hipótese de interesse é igual entre respostas e envolve apenas parâmetros de regressão, torna-se conveniente escrever a matriz L como o produto Kronecker de duas matrizes: uma matriz G e uma F , ou seja, $L = G \otimes F$. Desta forma, a matriz G tem dimensão $R \times R$ e especifica as hipóteses referentes às respostas, já a matriz F especifica as hipóteses entre variáveis e tem dimensão $s' \times h'$, em que s' é o número de restrições lineares, ou seja, o número de parâmetros testados para uma única resposta, e h' é o número total de coeficientes de regressão ou dispersão da resposta. Portanto, a matriz L tem dimensão $(s'R \times h)$.

Em geral, a matriz G é uma matriz identidade de dimensão igual ao número de respostas analisadas no modelo. Enquanto que a matriz F equivale a uma matriz L caso houvesse apenas uma única resposta no modelo e apenas parâmetros de regressão ou dispersão. Utilizamos o produto Kronecker destas duas matrizes para garantir que a hipótese descrita na matriz F seja testada nas R respostas do modelo.

Assim, considerando que se trata do caso em que se pode reescrever as hipóteses por meio da decomposição da matriz L , os elementos do teste ficam dados por:

- $\beta^T = [\beta_{10} \ \beta_{11} \ \beta_{20} \ \beta_{21}]$: os parâmetros de regressão do modelo.
- $G = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$: matriz identidade com dimensão dada pelo número de respostas.

- $F = [0 \ 1]$: equivalente a um L para uma única resposta.
- $L = G \otimes F = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$: matriz de especificação das hipóteses sobre todas as respostas.
- $c = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, é o valor sob hipótese nula.

Deste modo, o produto $L\beta$ representa a hipótese de interesse inicialmente postulada. Neste caso, a distribuição assintótica do teste é χ^2_2 . O procedimento é facilmente generalizado quando há interesse em avaliar uma hipótese sobre os parâmetros de dispersão e esta especificação é bastante conveniente para a geração de quadros de análise de variância.

4.2 | ANOVA e MANOVA via teste Wald

Com base na proposta de utilização do teste Wald para McGLMs, buscamos propor neste trabalho três diferentes procedimentos para geração de quadros de ANOVA e MANOVA para parâmetros de regressão, seguimos a nomenclatura tipos I, II e III. Além disso, buscamos propor também um procedimento análogo à uma ANOVA e MANOVA para avaliação dos parâmetros de dispersão de um dado modelo. No caso das ANOVAs gera-se um quadro para cada variável resposta. Para as MANOVAs apenas um quadro é gerado, por isso, para que seja possível realizar as MANOVAs, as respostas do modelo devem estar sujeitas ao mesmo preditor.

Para fins de ilustração dos testes feitos por cada tipo de análise de variância proposta, considere a situação em que deseja-se investigar se duas variáveis numéricas denotadas por X_1 e X_2 possuem efeito sobre duas variáveis resposta denotadas por Y_1 e Y_2 . Para tal tarefa coletou-se uma amostra com N observações e para cada observação foram registrados os valores de X_1 , X_2 , Y_1 e Y_2 . Com base nos dados coletados ajustou-se um modelo bivariado, com preditor dado por:

$$g_r(\mu_r) = \beta_{r0} + \beta_{r1}X_1 + \beta_{r2}x_2 + \beta_{r3}X_1X_2.$$

em que o índice r denota a variável resposta, $r = 1, 2$; β_{r0} representa o intercepto; β_{r1} um parâmetro de regressão associado à variável X_1 , β_{r2} um parâmetro de regressão associado à variável X_2 e β_{r3} um parâmetro de regressão associado à interação entre X_1 e X_2 . Considere que as unidades em estudo são independentes, portanto cada resposta possui apenas um parâmetro de dispersão τ_{r0} associado a uma matriz $Z_0 = I$. Além disso considere que os parâmetros de potência foram fixados.

4.2.1 | ANOVA e MANOVA tipo I

Nossa proposta de análise de variância do tipo I para os McGLMs realiza testes sobre os parâmetros de regressão de forma sequencial. Neste cenário, os seguintes testes seriam efetuados:

1. Testa se todos os parâmetros são iguais a 0.
2. Testa se todos os parâmetros, exceto intercepto, são iguais a 0.
3. Testa se todos os parâmetros, exceto intercepto e os parâmetros referentes a X_1 , são iguais a 0.
4. Testa se todos os parâmetros, exceto intercepto e os parâmetros referentes a X_1 e X_2 , são iguais a 0.

Cada um destes testes seria uma linha do quadro de análise de variância. No caso da ANOVA seria gerado um quadro por resposta, no caso da MANOVA um quadro em que as hipóteses são testadas para ambas as respostas. Este procedimento pode ser chamado de sequencial pois a cada linha é acrescentada uma variável. Em geral, justamente por esta sequencialidade, se torna difícil interpretar os efeitos das variáveis pela análise de variância do tipo I. Em contrapartida, as análises do tipo II e III testam hipóteses que são, geralmente, de mais interesse.

4.2.2 | ANOVA e MANOVA tipo II

Nossa análise de variância do tipo II efetua testes similares ao último teste da análise de variância sequencial. Em um modelo sem interação o que é feito é, em cada linha, testar o modelo completo contra o modelo sem uma variável. Deste modo se

torna melhor interpretável o efeito daquela variável sobre o modelo completo, isto é, o impacto na qualidade do modelo caso retirássemos determinada variável.

Caso haja interações no modelo, é testado o modelo completo contra o modelo sem o efeito principal e qualquer efeito de interação que envolva a variável. Considerando o preditor exemplo, a análise de variância do tipo II faria os seguintes testes:

1. Testa se o intercepto é igual a 0.
2. Testa se os parâmetros referentes a X_1 são iguais a 0. Ou seja, é avaliado o impacto da retirada de X_1 do modelo. Neste caso retira-se a interação pois nela há X_1 .
3. Testa se os parâmetros referentes a X_2 são iguais a 0. Ou seja, é avaliado o impacto da retirada de X_2 do modelo. Neste caso retira-se a interação pois nela há X_2 .
4. Testa se o efeito de interação é 0.

Note que, nas linhas em que se busca entender o efeito de X_1 e X_2 , a interação também é avaliada, pois retira-se do modelo todos os parâmetros que envolvem aquela variável.

4.2.3 | ANOVA e MANOVA tipo III

Na análise de variância do tipo II são feitos testes comparando o modelo completo contra o modelo sem todos os parâmetros que envolvem determinada variável (sejam efeitos principais ou interações). Já nossa análise de variância do tipo III considera o modelo completo contra o modelo sem determinada variável, seja ela efeito principal ou de interação. Deste modo, cuidados devem ser tomados nas conclusões pois uma variável não ter efeito constatado como efeito principal não quer dizer que não haverá efeito de interação. Considerando o preditor exemplo, a análise de variância do tipo III faria os seguintes testes:

1. Testa se o intercepto é igual a 0.
2. Testa se os parâmetros de efeito principal referentes a X_1 são iguais a 0. Ou seja, é avaliado o impacto da retirada de X_1 nos efeitos principais do modelo. Neste caso, diferente do tipo II, nada se supõe a respeito do parâmetro de interação, por mais que envolva X_1 .
3. Testa se os parâmetros de efeito principal referentes a X_2 são iguais a 0. Ou seja, é avaliado o impacto da retirada de X_2 nos efeitos principais do modelo. Novamente, diferente do tipo II, nada se supõe a respeito do parâmetro de interação, por mais que envolva X_2 .
4. Testa se o efeito de interação é 0.

Note que nas linhas em que se testa o efeito de X_1 e X_2 mantém-se o efeito da interação, diferentemente do que é feito na análise de variância do tipo II. É importante notar que as análises de variância do tipo II e III tal como foram propostas nesse trabalho geram os mesmos resultados quando aplicadas a modelos sem efeitos de interação. Além disso, generalizamos o procedimento tipo III para lidar com parâmetros de dispersão.

4.3 | Teste de comparações múltiplas via teste Wald

Quando a ANOVA aponta para efeito significativo de uma variável categórica, costuma ser de interesse do pesquisador avaliar quais dos níveis diferem entre si. Para isso são empregados os testes de comparações múltiplas. Na literatura existem diversos procedimentos para efetuar tais testes, muitos deles descritos em (author?)¹⁹.

Tal tipo de situação pode ser avaliada utilizando o teste Wald. Através da correta especificação da matriz L , é possível avaliar hipóteses sobre qualquer possível contraste entre os níveis de uma determinada variável categórica. Portanto, é possível usar a estatística de Wald para efetuar também testes de comparações múltiplas.

O procedimento é baseado basicamente em 3 passos. O primeiro deles é obter a matriz de combinações lineares dos parâmetros do modelo que resultam nas médias ajustadas. Com esta matriz é possível gerar a matriz de contrastes, dada pela subtração duas a duas das linhas da matriz de combinações lineares. Por fim, basta selecionar as linhas de interesse desta matriz e usá-las como matriz de especificação de hipóteses do teste Wald, no lugar da matriz L .

Por exemplo, suponha que há uma variável resposta Y sujeita a uma variável explicativa X de 4 níveis: A, B, C e D. Para avaliar o efeito da variável X , ajustou-se um modelo dado por:

$$g(\mu) = \beta_0 + \beta_1[X = B] + \beta_2[X = C] + \beta_3[X = D].$$

Nesta parametrização o primeiro nível da variável categórica é mantido como categoria de referência e, para os demais níveis, mede-se a mudança para a categoria de referência; este é o chamado contraste de tratamento. Neste contexto β_0 representa a média ajustada do nível A, enquanto que β_1 representa a diferença de A para B, β_2 representa a diferença de A para C e β_3 representa a diferença de A para D. Com esta parametrização é possível obter o valor predito para qualquer uma das categorias de tal modo que se o indivíduo pertencer à categoria A, β_0 representa o predito; se o indivíduo pertencer à categoria B, $\beta_0 + \beta_1$ representa o predito; para a categoria C, $\beta_0 + \beta_2$ representa o predito e, por fim, para a categoria D, $\beta_0 + \beta_3$ representa o predito.

Matricialmente, estes resultados podem ser descritos como

$$\mathbf{K}_0 = \begin{matrix} A \\ B \\ C \\ D \end{matrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

Note que o produto $\mathbf{K}_0\boldsymbol{\beta}$ gera o vetor de preditos para cada nível de X . Por meio da subtração das linhas da matriz de combinações lineares \mathbf{K}_0 podemos gerar uma matriz de contrastes \mathbf{K}_1

$$\mathbf{K}_1 = \begin{matrix} A - B \\ A - C \\ A - D \\ B - C \\ B - D \\ C - D \end{matrix} \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

Para proceder um teste de comparações múltiplas basta selecionar os contrastes desejados nas linhas da matriz \mathbf{K}_1 e utilizar estas linhas como matriz de especificação de hipóteses do teste Wald. Por fim, como usual em testes de comparações múltiplas, é recomendada a correção dos valores-p por meio da correção de Bonferroni.

Para efetuação deste procedimento para os McGLMs devemos lembrar que trata-se de uma classe de modelos multivariados. E tal como ocorre no caso das análises de variância, para os testes de comparações múltiplas existem duas possibilidades: testes para uma única resposta e testes para múltiplas respostas.

Na prática, se o interesse for um teste de comparações múltiplas multivariado, existe a necessidade de todas as respostas estarem sujeitas a um mesmo preditor e basta expandir a matriz de contrastes utilizando o produto Kronecker, seguindo uma ideia muito similar ao exposto na subsubsection 4.1.4. No caso de um teste de comparações múltiplas para cada resposta, basta selecionar o vetor de estimativas e a partição correspondente ao vetor da matriz \mathbf{J}_θ^{-1} para a resposta específica e proceder com o teste. Desta forma é possível chegar a um simples e útil procedimento de comparações múltiplas para quando há um McGLM com variáveis explicativas categóricas e há interesse em determinar quais níveis diferem entre si.

5 | SIMULATION STUDIES

Com o objetivo de avaliar o poder do teste Wald em testes de hipóteses sobre parâmetros de McGLMs, foram executados estudos de simulação. Nestas simulações avaliamos o comportamento da proposta para três distribuições de probabilidade: Normal, Poisson e Bernoulli. Simulamos cenários univariados e também trivariados com diferentes tamanhos amostrais para verificar as propriedades dos testes sobre parâmetros de regressão e dispersão.

Para simular conjuntos de dados univariados foram usadas bibliotecas padrões do R. Para simular conjuntos de dados com múltiplas respostas seguindo distribuição Normal, foi usada a biblioteca R *mvtnorm*[?],[?]. Para as outras distribuições foi utilizado o método NORTA[?] implementado na biblioteca R *NORTARA*[?].

5.1 | Parâmetros de regressão

Para avaliação de hipóteses sobre parâmetros de regressão foram considerados tamanhos amostrais de 50, 100, 250, 500 e 1000. Foram gerados 500 conjuntos de dados para cada tamanho amostral simulando uma situação com uma variável explicativa categórica de 4 níveis. Para distribuição Normal os parâmetros de regressão usados foram: $\beta_0 = 5$, $\beta_1 = 0$, $\beta_2 = 0$, $\beta_3 = 0$. Para a distribuição Poisson os parâmetros de regressão usados foram: $\beta_0 = 2, 3$, $\beta_1 = 0$, $\beta_2 = 0$, $\beta_3 = 0$. E para a distribuição Bernoulli os parâmetros de regressão usados foram: $\beta_0 = 0, 5$, $\beta_1 = 0$, $\beta_2 = 0$, $\beta_3 = 0$. Os valores foram escolhidos de tal modo que o coeficiente de variação para distribuição Normal fosse de 20%, as contagens para Poisson fossem próximas de 10 e a probabilidade de sucesso da Bernoulli fosse aproximadamente 0,6. Foram avaliados cenários univariados e trivariados com estas características. Para os cenários trivariados, existem 4 parâmetros por resposta que seguem as configurações descritas. Para cada amostra gerada foi ajustado um McGLM nos quais as funções de ligação e variância para cada distribuição são apresentadas na Table 2.

Em todos os casos o preditor matricial para a matriz de variância e covariância foi especificado de forma a explicitar que as observações são independentes dentro de cada resposta. A correlação entre respostas no caso trivariado é dada pela matriz Σ_b descrita na Equation 19.

$$\Sigma_b = \begin{bmatrix} 1 & 0,75 & 0,5 \\ 0,75 & 1 & 0,25 \\ 0,5 & 0,25 & 1 \end{bmatrix} \quad (19)$$

Com os modelos ajustados, o procedimento consistiu em variar as hipóteses testadas sobre os parâmetros simulados. Consideramos 20 diferentes hipóteses baseadas em um decréscimo em β_0 e distribuição deste decréscimo nos demais β s da hipótese nula. O decréscimo para respostas seguindo distribuição Normal foi de 0,15; para distribuição Poisson o decréscimo foi de 0,05; e para distribuição Bernoulli o decréscimo foi de 0,25. Estes valores foram escolhidos levando em conta o afastamento desejado das hipóteses testadas na escala da resposta e é importante notar que estes valores são diferentes devido ao impacto da função de ligação usada em cada configuração de modelo e também devido às propriedades dos parâmetros das distribuições.

Para cada ponto avaliamos o percentual de rejeição da hipótese nula. A ideia é verificar o que ocorre quando afastamos as hipóteses nulas dos reais valores dos parâmetros. Espera-se que no primeiro ponto haja um percentual de rejeição baixo, pois a hipótese nula corresponde aos reais valores dos parâmetros. Para os demais pontos espera-se que o percentual de rejeição aumente gradativamente, pois as hipóteses afastam-se cada vez mais dos valores originalmente simulados. As hipóteses testadas em cada cenário estão disponíveis no apêndice desta dissertação.

Para representar graficamente os resultados tomamos a distância euclidiana de cada vetor de hipóteses com relação ao vetor usado para simular os dados. Adicionalmente dividimos o vetor de distâncias pela maior distância para obter distâncias padronizadas entre 0 e 1, independente dos parâmetros de regressão. Os resultados são apresentados na ??.

FIGURA

De modo geral, quanto mais distante a hipótese é dos valores inicialmente simulados, maior é o percentual de rejeição. Como esperado, os menores percentuais foram observados na hipótese igual aos valores simulados. Nos cenários univariados o percentual de rejeição foi próximo de 5% quando a hipótese era igual aos valores simulados mesmo com tamanhos amostrais reduzidos. Para os cenários trivariados, no menor tamanho amostral avaliado, o percentual de rejeição não excedeu 10% e em tamanhos amostrais iguais a 500 o percentual de rejeição foi próximo de 5%. Também como esperado, foi possível verificar que conforme aumenta-se o tamanho amostral, o percentual de rejeição aumenta para hipóteses pouco diferentes dos valores simulados dos parâmetros.

5.2 | Parâmetros de dispersão

Para avaliação de hipóteses sobre parâmetros de dispersão foram considerados os mesmos tamanhos amostrais: 50, 100, 250, 500 e 1000. Contudo, os conjuntos de dados simulam uma situação em que cada unidade amostral fornece 5 medidas ao conjunto de dados. Foram gerados 500 conjuntos de dados para cada tamanho amostral e distribuição. Para distribuição Normal foram simulados vetores com média 5 e desvio padrão igual a 1. Para distribuição Poisson foram simuladas contagens com taxa igual a 10. Para distribuição Bernoulli foram simulados vetores de uma variável dicotômica com probabilidade de sucesso igual a 0,6.

Em todos os casos, os parâmetros de dispersão para gerar os conjuntos de dados foram fixados em $\tau_0 = 1$, $\tau_1 = 0$ e não foi incluído efeito de variáveis explicativas. Foram avaliados cenários univariados e trivariados com estas características. Para

cada amostra gerada foi ajustado um McGLM com funções de ligação e variância tal como descrito na Table 2. Nos cenários trivariados a correlação entre respostas é dada pela Equation 19.

Neste caso, como o objetivo é avaliar a correlação dentro das respostas, é necessário especificar um preditor matricial. O objetivo é testar hipóteses sobre os parâmetros de dispersão associados a este preditor matricial.

Com os modelos ajustados, o procedimento consistiu em variar as hipóteses testadas sobre os parâmetros simulados. Consideramos 20 diferentes hipóteses baseadas em um decréscimo sucessivo de 0,02 em τ_0 e acréscimo de 0,02 em τ_1 para cada hipótese nula testada. Para cada ponto avaliamos o percentual de rejeição da hipótese nula. A ideia é afastar sucessivamente a hipótese dos valores simulados e avaliar se conforme afastamos a hipótese dos valores verdadeiros, o percentual de rejeição aumenta. As hipóteses testadas estão disponíveis no apêndice.

Do mesmo modo que foi feito para os parâmetros de regressão, foi tomada a distância euclidiana de cada vetor de hipóteses com relação ao vetor usado para simular os dados; e o vetor de distâncias foi padronizado para obter distâncias entre 0 e 1. Os resultados são apresentados na ??.

FIGURA

Assim como observado para os parâmetros de regressão, o comportamento dos gráficos mostra que, quanto mais distante a hipótese é dos valores inicialmente simulados, maior é o percentual de rejeição, e os menores percentuais são observados em hipóteses próximas aos valores simulados. Na primeira hipótese testada, para os cenários univariados, percentuais de rejeição próximos a 8% foram observados no menor tamanho amostral avaliado. A partir de tamanhos amostrais iguais a 250 o percentual de rejeição foi próximo de 5%. Já para os casos trivariados, no menor tamanho amostral, o percentual de rejeição excedia 10% no menor tamanho amostral; para tamanhos amostrais maiores, os percentuais ficaram em torno de 7%. Também verificou-se para os parâmetros de dispersão que conforme aumenta-se o tamanho amostral, o percentual de rejeição aumenta para hipóteses pouco diferentes dos valores simulados dos parâmetros.

6 | APPLICATION

6.1 | Especificação do modelo

Para análise dos dados foi ajustado um modelo multivariado com os efeitos fixos das variáveis momento e grupo. Adicionalmente, foi incluído no modelo o efeito da interação entre estas duas variáveis explicativas.

Como já mencionado, trata-se de um experimento em que as observações não são independentes pois medidas tomadas em um mesmo indivíduo são correlacionadas e esta correlação deve ser especificada no modelo.

Ambas as respostas foram tratadas como proporções, por este motivo foi utilizada a função de ligação logito com função de variância binomial. Adicionalmente, estimou-se o parâmetro de potência para todas as respostas em análise. Os preditores lineares são dados por

$$g_r(\mu_r) = \beta_{0r} + \beta_{1r}T1 + \beta_{2r}T2 + \beta_{3r}Probiotico + \beta_{4r}T1 * Probiotico + \beta_{5r}T2 * Probiotico,$$

em que o índice r refere-se às variáveis respostas do estudo (1 para YFAS, 2 para BES). Foram consideradas categorias de referência o grupo placebo e o momento T0. β_{0r} representa o intercepto, β_{1r} o efeito do momento T1, β_{2r} o efeito do momento T2, β_{3r} o efeito de probiótico. Os parâmetros β_{4r} e β_{5r} referem-se à interação entre momento e grupo, de tal forma que β_{4r} representa o efeito da interação entre T1 e probiótico, e β_{5r} representa o efeito da interação entre T2 e probiótico.

Os preditores matriciais, iguais para ambas as respostas, são dados por $h\{\mathbf{\Omega}(\boldsymbol{\tau})\} = \tau_0\mathbf{Z}_0 + \tau_1\mathbf{Z}_1$. A função $h(\cdot)$ utilizada foi a identidade, τ_0 e τ_1 representam os parâmetros de dispersão, \mathbf{Z}_0 representa uma matriz identidade de ordem, 184×184 e \mathbf{Z}_1 representa uma matriz de dimensão 184×184 especificada de forma a explicitar que as medidas provenientes do mesmo indivíduo são correlacionadas.

Para exemplificar a forma do preditor matricial, vamos considerar 3 indivíduos: A, B e C. Suponha que o indivíduo A compareceu às 3 consultas, portanto temos informações deste indivíduo em T0, T1 e T2. O indivíduo B compareceu em T0 e T1. Já o indivíduo C compareceu apenas em T0. Deste modo temos 3 indivíduos e 6 observações. Logo \mathbf{Z}_0 é uma matriz identidade

6×6 e Z_1 é uma espécie de matriz bloco diagonal em que o tamanho dos blocos varia de acordo com o número de medidas para cada indivíduo. Neste cenário o preditor matricial tem a forma

$$h\{\Omega(\tau)\} = \tau_0 \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} + \tau_1 \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (20)$$

6.2 | Resultados do ajuste

Com o propósito de verificar a qualidade do ajuste do modelo, foi feita a análise de resíduos do modelo. A análise mostra que os resíduos de Pearson para YFAS e BES apresentam média 0 e desvio padrão próximo de 1. Na ?? são exibidos os histogramas dos resíduos de Pearson por resposta, a distribuição dos resíduos é aproximadamente simétrica com a maior parte dos dados entre -2 e 2.

Na ?? são exibidos os resíduos versos preditos do modelo. Os resultados mostram que não parece haver qualquer tipo de relação entre resíduos e preditos. De forma geral, o modelo parece estar razoavelmente bem ajustado aos dados.

FIGURA

FIGURA

As estimativas dos parâmetros, intervalos de confiança assintóticos com 95% de confiança e valores-p da hipótese de nulidade dos parâmetros são mostrados na Table 3. Adicionalmente, a ?? mostra os valores preditos para cada combinação dos fatores para uma melhor interpretação dos resultados.

FIGURA

6.3 | Testes de hipóteses

Até este ponto foi apresentada uma análise padrão, com os resultados usuais da análise de um modelo de regressão. Indo mais além nesta análise, podemos fazer uso dos resultados e implementações deste trabalho.

Podemos optar por uma análise de variância multivariada do tipo II para avaliar a importância das variáveis no problema. O resultado, apresentado na tabela Table 4 aponta para existência significativa do efeito de momento e ausência de efeito de grupo, indicando que para ambas as respostas as métricas se alteram ao longo do tempo mas sem alteração entre grupos.

A fim de avaliar os resultados por resposta, podemos utilizar uma análise de variância univariada do tipo II. Os resultados são apresentados na Table 5. Considerando um nível de significância de 0,01, existe evidência que aponta para efeito de momento.

Como as análises de variâncias apontaram para efeito significativo de variáveis categóricas, podemos explorar quais níveis diferem entre si. A Table 6 apresenta as comparações duas a duas entre momentos. Os resultados mostram que, para ambas as respostas existem diferenças entre o primeiro versus segundo e primeiro versus terceiro momento, mas os dois últimos momentos não diferem entre si.

A tabela Table 7 apresenta as comparações entre grupos para cada momento para ambas as respostas. Os resultados apontam para a ausência de diferença entre grupos em cada momento.

No modelo, incluímos a informação de que existem medidas que foram tomadas em um mesmo indivíduo. Esta informação é declarada no preditor matricial que estima um parâmetro de dispersão associado à matriz que aponta a relação entre os indivíduos. Uma hipótese de interesse pode ser avaliar se existe evidência para crer que, neste problema, as medidas tomadas em um mesmo indivíduo são de fato correlacionadas. Para isso podemos postular hipóteses sobre os parâmetros de dispersão. Tal como nas análises de variância, isso pode ser feito por resposta ou para ambas as respostas simultaneamente.

A tabela Table 8 apresenta os resultados do teste multivariado, ou seja, avalia a hipótese de que em ambas as respostas as medidas sejam correlacionadas. Os resultados apontam que não há evidência para crer que as medidas tomadas em um mesmo indivíduo apresentam correlação.

7 | CONCLUDING REMARKS

O objetivo deste trabalho foi desenvolver procedimentos para realizar testes de hipóteses sobre parâmetros de McGLMs baseados na estatística de Wald. McGLMs contam com parâmetros de regressão, dispersão, potência e correlação; cada conjunto de parâmetros possui uma interpretação prática bastante relevante no contexto de análise de problemas com potenciais múltiplas respostas em função de um conjunto de variáveis explicativas.

Com base na proposta de utilização do teste Wald para McGLMs, desenvolvemos procedimentos para testes de hipóteses lineares gerais, geração de quadros de ANOVA e MANOVA para parâmetros de regressão e dispersão e também testes de comparações múltiplas.

As propriedades dos testes foram avaliadas com base em estudos de simulação. Foram considerados cenários univariados e bivariados com diferentes distribuições de probabilidade para as respostas e diferentes tamanhos amostrais. A ideia desta etapa do trabalho foi gerar conjuntos de dados com parâmetros de regressão e dispersão fixados e testar hipóteses sobre parâmetros de modelos ajustados com estes dados.

Em um primeiro momento, testamos a hipótese de que os parâmetros eram realmente iguais aos fixados na simulação. Em seguida afastamos gradativamente as hipóteses dos valores simulados a fim de verificar se, à medida que afasta-se a hipótese dos verdadeiros valores, o percentual de rejeição aumenta.

De modo geral, em todos os casos foi possível observar que quanto mais distante a hipótese é dos valores inicialmente simulados, maior é o percentual de rejeição. Tal como esperado, os menores percentuais foram observados na hipótese igual aos valores simulados e também foi possível verificar que conforme aumenta-se o tamanho amostral, o percentual de rejeição aumenta para hipóteses pouco diferentes dos valores simulados dos parâmetros.

Sendo assim, os resultados das simulações mostraram que o teste Wald pode ser usado para avaliar hipóteses sobre parâmetros de regressão e dispersão de McGLMs, o que permite uma melhor interpretação do efeito das variáveis e estruturas de delineamento em contextos práticos.

Adicionalmente, fizemos a aplicação das metodologias propostas a um conjunto de dados real em que o objetivo é avaliar o efeito do uso de probióticos no controle de vícios e compulsões alimentares. Trata-se de um problema com duas variáveis resposta: um escore que caracteriza compulsão e o número de sintomas apresentados que caracterizam vício.

Neste estudo, um conjunto de indivíduos foi dividido em dois grupos: um deles recebeu um placebo e o outro recebeu o tratamento. Além disso os indivíduos foram avaliados ao longo tempo; deste modo o delineamento gera observações que não são independentes, já que medidas tomadas em um mesmo indivíduo tendem a ser correlacionadas.

Este é um problema em que técnicas de modelagem tradicionais seriam de difícil aplicação. Contudo trata-se de um problema de possível análise via McGLM e testes de hipóteses podem ser empregados para avaliar o efeito da interação entre momento e uso do probiótico sobre vício e compulsão alimentar.

Os resultados, baseados nos testes propostos neste trabalho indicam que existe evidência que aponta para efeito de momento, ou seja, vício e compulsão alimentar alteram-se ao longo do tempo. Os testes de comparações múltiplas indicam que, para ambas as respostas, existem diferenças entre o primeiro versus segundo e primeiro versus terceiro momento, mas os dois últimos momentos não diferem entre si. Os resultados também apontam para a ausência de diferença entre grupos em cada momento. Uma avaliação dos parâmetros de dispersão mostra que não há evidência para crer que as medidas tomadas em um mesmo indivíduo apresentam correlação.

Algumas limitações deste trabalho dizem respeito a casos não explorados nos estudos de simulação, tais como: avaliação do desempenho dos testes ao definir hipóteses lineares que combinem parâmetros de diferentes tipos, impacto de um número diferente de observações por indivíduos em problemas longitudinais ou de medidas repetidas, impacto no poder do teste conforme o número de parâmetros testados aumenta e o comportamento do teste em problemas multivariados com distribuições de probabilidade diferentes das exploradas.

Possíveis extensões deste trabalho que seguem na linha de avaliação de parâmetros de McGLMs para um melhor entendimento do impacto dos elementos em problemas de modelagem são: explorar correções de valores-p de acordo com o tamanho das hipóteses testadas, explorar procedimentos além do teste Wald (como o teste Escore e o teste da razão de verossimilhanças), implementar novos procedimentos para comparações múltiplas, adaptar a proposta para lidar com contrastes alternativos aos usuais, explorar procedimentos para seleção automática de covariáveis (backward elimination, forward selection, stepwise selection) e também seleção de covariáveis por meio de inclusão de penalização no ajuste por complexidade (similar a ideia de regressão por splines).

Todos estes procedimentos foram implementados na linguagem R e complementam as funcionalidades existentes na biblioteca *mcglm*.

ACKNOWLEDGMENTS

This is acknowledgment text.[?] Provide text here. This is acknowledgment text. Provide text here. This is acknowledgment text. Provide text here. This is acknowledgment text. Provide text here. This is acknowledgment text. Provide text here. This is acknowledgment text. Provide text here. This is acknowledgment text. Provide text here. This is acknowledgment text. Provide text here.

Author contributions

This is an author contribution text. This is an author contribution text. This is an author contribution text. This is an author contribution text. This is an author contribution text.

Financial disclosure

None reported.

Conflict of interest

The authors declare no potential conflict of interests.

SUPPORTING INFORMATION

The following supporting information is available as part of the online article:

Figure S1. 500 hPa geopotential anomalies for GC2C calculated against the ERA Interim reanalysis. The period is 1989–2008.

Figure S2. The SST anomalies for GC2C calculated against the observations (OIst).

References

1. Meinert C, Tonascia S. *Clinical Trials: Design, Conduct, and Analysis*. Monographs in epidemiology and biostatistics Oxford University Press . 1986.
2. Hannan EL. Randomized clinical trials and observational studies: guidelines for assessing respective strengths and limitations. *JACC: Cardiovascular Interventions* 2008; 1(3): 211–217.
3. Kangovi S, Mitra N, Norton L, et al. Effect of community health worker support on clinical outcomes of low-income patients across primary care facilities: a randomized clinical trial. *JAMA internal medicine* 2018; 178(12): 1635–1643.
4. Song Z, Baicker K. Effect of a workplace wellness program on employee health and economic outcomes: a randomized clinical trial. *Jama* 2019; 321(15): 1491–1501.
5. Thyregod HGH, Ihlemann N, Jørgensen TH, et al. Five-year clinical and echocardiographic outcomes from the Nordic Aortic Valve Intervention (NOTION) randomized clinical trial in lower surgical risk patients. *Circulation* 2019; 139(4): 2714–23.
6. Schmitz KH, Troxel AB, Dean LT, et al. Effect of home-based exercise and weight loss programs on breast cancer-related lymphedema outcomes among overweight breast cancer survivors: The WISER Survivor randomized clinical trial. *JAMA oncology* 2019; 5(11): 1605–1613.

7. Duma N, Kothadia SM, Azam TU, et al. Characterization of comorbidities limiting the recruitment of patients in early phase clinical trials. *The Oncologist* 2019; 24(1): 96–102.
8. Bonat WH, Jørgensen B. Multivariate covariance generalized linear models. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 2016; 65(5): 649–675.
9. Martinez-Beneito MA. A general modelling framework for multivariate disease mapping. *Biometrika* 2013; 100(3): 539–553.
10. Engle RF. Wald, likelihood ratio, and Lagrange multiplier tests in econometrics. *Handbook of econometrics* 1984; 2: 775–826.
11. Fisher RA, Mackenzie WA. Studies in crop variation. II. The manurial response of different potato varieties. *The Journal of Agricultural Science* 1923; 13(3): 311–320. doi: 10.1017/S0021859600003592
12. St L, Wold S, others . Analysis of variance (ANOVA). *Chemometrics and intelligent laboratory systems* 1989; 6(4): 259–272.
13. Smith H, Gnanadesikan R, Hughes J. Multivariate analysis of variance (MANOVA). *Biometrics* 1962; 18(1): 22–41.
14. Wilks SS. Certain generalizations in the analysis of variance. *Biometrika* 1932: 471–494.
15. Lawley D. A generalization of Fisher's z test. *Biometrika* 1938; 30(1/2): 180–187.
16. Hotelling H. A generalized T test and measure of multivariate dispersion. In: University of California Press. ; 1951: 23–41.
17. Pillai K, others . Some new test criteria in multivariate analysis. *The Annals of Mathematical Statistics* 1955; 26(1): 117–121.
18. Roy SN. On a heuristic method of test construction and its use in multivariate analysis. *The Annals of Mathematical Statistics* 1953: 220–238.
19. Hsu J. *Multiple comparisons: theory and methods*. CRC Press . 1996.
20. Bretz F, Hothorn T, Westfall P. Multiple comparison procedures in linear models. In: Springer. 2008 (pp. 423–431).

Grupo	Momento	n	YFAS	BES
			Média (desvio padrão)	Média (desvio padrão)
Placebo	T0	33	0,37 (0,26)	0.24 (0,20)
Placebo	T1	32	0,11 (0,15)	0.09 (0,10)
Placebo	T2	22	0,16 (0,15)	0.10 (0,12)
Probiótico	T0	38	0,49 (0,24)	0.32 (0,18)
Probiótico	T1	37	0,09 (0,12)	0.10 (0,08)
Probiótico	T2	22	0,10 (0,14)	0.07 (0,09)

TABLE 1 Número de indivíduos, média e desvio padrão para YFAS e BES para cada combinação de grupo e momento.

Distribuição	Função de ligação	Função de variância
Normal	Identidade	Constante
Poisson	Logarítmica	Tweedie
Bernoulli	Logito	Binomial

TABLE 2 Funções de ligação e variância utilizadas nos modelos para cada distribuição simulada.

Parâmetro	YFAS			BES		
	Estimativa	Intervalo de confiança	Valor-p	Estimativa	Intervalo de confiança	Valor-p
β_0	-0,54	(-0,87;-0,22)	<0,01	-1,13	(-1,44;-0,83)	<0,01
β_1	-1,55	(-2,17;-0,94)	<0,01	-1,16	(-1,62;-0,69)	<0,01
β_2	-1,13	(-1,75;-0,51)	<0,01	-1,05	(-1,58;-0,52)	<0,01
β_3	0,49	(0,05;0,93)	0,0284	0,37	(-0,03;0,77)	0,0733
β_4	-0,73	(-1,60;0,14)	0,0	-0,33	(-0,96;0,30)	0,3081
β_5	-0,98	(-1,93;-0,03)	0,0429	-0,80	(-1,58;-0,02)	0,0449
τ_0	0,18	(0,01;0,35)	0,0411	0,17	(0,00;0,34)	0,0458
τ_1	0,01	(-0,02;0,04)	0,5718	0,04	(-0,01;0,10)	0,1357
p	0,91	(0,47;1,34)	<0,05	1,23	(0,77;1,68)	<0,05

TABLE 3 Estimativas dos parâmetros, intervalos com 95% de confiança e valores-p do modelo.

Variável	Graus de liberdade	Qui-quadrado	Valor-p
Intercepto	2	53,1581	<0,01
Momento	8	139,0161	<0,01
Grupo	6	8,4928	0,2042
Momento*Grupo	4	6,9923	0,1363

TABLE 4 Análise de variância multivariada tipo II.

Variável	Graus de liberdade	YFAS		BES	
		Qui-quadrado	Valor-p	Qui-quadrado	Valor-p
Intercepto	1	10,6128	<0,01	53,1473	<0,01
Momento	4	102,9875	<0,01	99,5681	<0,01
Grupo	3	6,6837	0,0827	5,3083	0,1506
Momento*Grupo	2	5,5984	0,0609	4,2477	0,1196

TABLE 5 Análise de variância univariada do tipo II.

Contraste	Graus de liberdade	Qui-quadrado	Valor-p
T0-T1	2	97,9874	<0,01
T0-T2	2	67,2462	<0,01
T1-T2	2	2,4730	0,8712

TABLE 6 Comparações duas a duas entre momentos para ambas as respostas.

Contraste	Graus de liberdade	Qui-quadrado	Valor-p
T0:Placebo-T0:Probiótico	2	5,5819	0.9204
T1:Placebo-T1:Probiótico	2	0,6096	1
T2:Placebo-T2:Probiótico	2	1,7645	1

TABLE 7 Comparações duas a duas entre grupos para cada momento para ambas as respostas.

Variável	Graus de liberdade	Qui-quadrado	Valor-p
τ_0	2	7,1936	<0,01
τ_1	2	2,3201	0,3135

TABLE 8 Análise de variância multivariada do tipo III para parâmetros de dispersão.

How to cite this article: Williams K., B. Hoskins, R. Lee, G. Masato, and T. Woollings (2016), A regime analysis of Atlantic winter jet variability applied to evaluate HadGEM3-GC2, *Q.J.R. Meteorol. Soc.*, 2017;00:1–6.