

Carolina Sales da Silva

**Uma avaliação de testes de hipóteses em
modelos de regressão beta prime**

João Pessoa

2022

Carolina Sales da Silva

Uma avaliação de testes de hipóteses em modelos de regressão beta prime

Trabalho de conclusão de curso em Estatística da Universidade Federal da Paraíba, como requisito parcial para obtenção do título de Bacharel em Estatística.

Orientadora: Profa. Dra. Tarciana Liberal Pereira

Universidade Federal da Paraíba – UFPB

Centro de Ciências Exatas e da Natureza

Departamento de Estatística

Curso de Bacharelado em Estatística

João Pessoa

2022

Catálogo na publicação
Seção de Catalogação e Classificação

S586a Silva, Carolina Sales da.

Uma avaliação de testes de hipóteses em modelos de regressão beta prime / Carolina Sales da Silva. - João Pessoa, 2022.

53 f. : il.

Orientação: Tarciana Liberal Pereira.

TCC (Curso de Bacharelado em Estatística) - UFPB/CCEN.

1. Teste de hipótese. 2. Modelo de regressão Beta Prime. 3. Simulação de Monte Carlo. I. Pereira, Tarciana Liberal. II. Título.

UFPB/CCEN

CDU 311(043.2)



DEPARTAMENTO DE ESTATÍSTICA
UNIVERSIDADE FEDERAL DA PARAÍBA

Trabalho de Conclusão de Curso de Estatística intitulado Uma avaliação de testes de hipóteses em modelos de regressão beta prime de autoria de Carolina Sales da Silva, aprovado pela banca examinadores constituída pelos seguintes professores:

Profa. Dra. Tarciana Liberal Pereira de Araujo
(Professora Orientadora)
Universidade Federal da Paraíba

Prof. Dr. Hemílio Fernandes Campos Coelho
(Professor Examinador)
Universidade Federal da Paraíba

Profa. Dra. Maria Lídia Coco Terra
(Professora Examinadora)
Universidade Federal da Paraíba

João Pessoa, 2022.

Agradecimentos

Agradeço primeiramente a Deus, pela proteção e sabedoria, sempre iluminando meu caminho.

Aos meus pais, Cecília e Carlos Alberto, por todo incentivo e dedicação para me proporcionar uma boa educação e não medirem esforços para tornar meus objetivos possíveis. Em especial à minha mãe, um exemplo de força, fé e altruísmo. Obrigada por todos os conselhos, paciência, compreensão nos momentos difíceis e, principalmente, todo amor. Acredito que nunca conseguirei expressar em palavras quão grata eu sou. Muito obrigada.

Ao meu melhor amigo e namorado, Adriano, por todo amor, dedicação e companherismo. Por sempre acreditar em mim, mesmo quando eu duvidava de mim mesma. Pela motivação e incentivo em durante todo percurso desse trabalho. Obrigada por ter sido meu porto seguro nos momentos difíceis e pela paciência e dedicação de quem ama.

Ao meu irmão, Hugo, por me proteger e ter estado ao meu lado nos diversos momentos difíceis que passamos juntos.

A toda a minha gigante família, que só cresceu durante meu percurso no curso de graduação, muito obrigada por sempre torcerem e acreditarem em mim.

A minha prima, Isabel Cristina, pelo cuidado e solidariedade. Por ter me acolhido como uma filha e me dado o suporte necessário para que pudesse construir meu futuro em uma cidade nova. Por isso, e muito mais, eu agradeço de coração.

A minha prima Maria Luiza, pela amizade dos últimos anos. Obrigada por nunca deixar eu me sentir sozinha e estar comigo nos momentos que mais precisava, se tornou uma verdadeira amiga.

Aos colegas e amigos que conheci durante o curso de graduação, por toda amizade, motivação, aprendizado e momentos maravilhosos durante todos esses anos. Acredito que sem saber, me motivaram e incentivaram a continuar tornando tudo isso possível, muito obrigada. Em especial Bianca, Leticia, Nicollas, Wanusa, Cecilia, Franciele, Manel, Kleber e Marcos.

Aos meus amigos, Roberta, Lucas e Lorrana, apesar da distância ter possibilitado alguns poucos momentos juntos, estes momentos são sempre muito felizes e especiais para mim.

Aos professores do Departamento de Estatística da Universidade Federal da Paraíba, agradeço de coração por todo aprendizado que vocês me proporcionaram durante o curso, por toda compreensão, conselhos e suas contribuições à minha formação como estatística.

Em especial gostaria de agradecer aos professores Maria Lúcia, Hemílio, Tatiene, Luiz, Pedro Rafael, Eufrásio, Ana Hermínia, Rodrigo, Everlane, Marcelo, Isabel e Ana Flávia.

Em especial ao professor Hemílio, por ter me dado a oportunidade de ser sua aluna PIBIC. Obrigada por despertar minha admiração pela profissão de estatística. Obrigada por sempre continuar acreditando em mim, pela compreensão nos momentos difíceis, apoio e ajuda incondicional.

Um agradecimento mais que especial a minha orientadora, professora Tarciana. Em momentos que pensei que não seria possível terminar o curso de graduação, me incentivou e não me deixou desistir. Obrigado pelas horas de dedicação e paciência ao longo desses anos como minha professora e orientadora. Minha gratidão por todos os ensinamentos, conselhos, pela oportunidade, confiança e motivação.

Aos professores Hemílio e Maria Lúcia pela prontidão em participar da banca desse trabalho.

Ao CNPQ pelo apoio financeiro durante a graduação.

Resumo

O modelo de regressão beta prime (BOURGUIGNON; SANTOS-NETO; CASTRO, 2018) é bastante útil para modelar dados positivos assimétricos utilizados em diversas áreas do conhecimento, tais como: economia, administração, engenharias, biologia, agronomia, saúde, sociologia, entre outras. A análise de modelos de regressão tem o intuito de explicar a relação entre as variáveis do estudo. Para isso, através de testes de hipóteses verificamos se há evidência estatística de acordo com os dados amostrais de que as hipóteses feitas sobre os parâmetros populacionais são verdadeiras. Sabendo disso, neste trabalho o interesse é testar hipóteses sobre os parâmetros do modelo de regressão beta prime. Dessa forma, nós avaliamos os desempenhos dos principais testes de hipótese para o modelo de regressão beta prime considerando os submodelos da média e da precisão em pequenas e grandes amostras, sendo uma alternativa aos modelos lineares generalizados. São considerados os testes da razão de verossimilhanças, escore e Wald usuais, além da razão de verossimilhança sinalizada e gradiente. Através de simulações de Monte Carlo, em amostras finitas, avaliamos o desempenho dos testes por meio das propriedades de tamanho e poder. De acordo com os resultados das simulações foi possível obter evidências de que os testes escore, gradiente e da razão de verossimilhança sinalizada apresentaram os melhores desempenhos para os cenários abordados no estudo.

Palavras – chave: Teste de hipótese, Modelo de regressão Beta Prime, Simulação de Monte Carlo.

Abstract

The beta prime regression model BOURGUIGNON; SANTOS-NETO; CASTRO (2018) is very useful for modeling asymmetric positive data used in several areas of knowledge, such as: economics, administration, engineering, biology, agronomy, health, sociology and others. The analysis of regression models pursue to explain the relationship among the study variables. To do this, through tests of hypotheses, we investigate if there is statistical evidence according to the sample data of that the assumptions made about the population parameters are true. In this work, the interest is to test hypotheses about the parameters of the regression model beta prime. We evaluate the performance of the main hypothesis tests for the beta prime regression model considering the mean and precision sub-models in small and large samples, being an alternative to generalized linear models. By considering the likelihood ratio test, the usuals score and Wald tests, in addition to the signed likelihood ratio and gradient. Through Monte Carlo simulations, in finite samples, we evaluated the performance of the tests through the properties of size and power. According to the simulation results was possible to obtain evidence that the score, gradient and signaled likelihood ratio tests presented the best performances for the scenarios addressed in the study.

Keywords: Beta prime regression model, hypothesis test, Monte Carlo simulation.

Lista de ilustrações

Figura 1 – Gráfico da função densidade de probabilidade da distribuição <i>BP</i> considerando $\phi = 1$ e distintos valores para μ	19
Figura 2 – Gráfico da função densidade de probabilidade da distribuição <i>BP</i> considerando $\phi = 10$ e distintos valores para μ	19
Figura 3 – Gráfico da função densidade de probabilidade da distribuição <i>BP</i> considerando $\phi = 100$ e distintos valores para μ	20

Lista de tabelas

Tabela 1 – Cenários considerados na simulação.	31
Tabela 2 – Taxa de rejeição da hipótese nula (%) para a o cenário I, submodelo de μ	32
Tabela 3 – Taxa de rejeição da hipótese nula (%) para a o cenário II, submodelo de μ	33
Tabela 4 – Taxa de rejeição da hipótese nula (%) para a o cenário III, submodelo de μ	33
Tabela 5 – Taxa de rejeição da hipótese nula (%) para a o cenário IV, submodelo de μ	34
Tabela 6 – Taxa de rejeição da hipótese nula (%) para a o cenário I, submodelo de ϕ com $r = 1$	34
Tabela 7 – Taxa de rejeição da hipótese nula (%) para a o cenário II, submodelo de ϕ com $r = 1$	35
Tabela 8 – Taxa de rejeição da hipótese nula (%) para a o cenário III, submodelo de ϕ com $r = 1$	35
Tabela 9 – Taxa de rejeição da hipótese nula (%) para a o cenário IV, submodelo de ϕ com $r = 1$	36
Tabela 10 – Taxa de rejeição da hipótese nula (%) para a o cenário I, submodelo de ϕ com $r = 2$	36
Tabela 11 – Taxa de rejeição da hipótese nula (%) para a o cenário II, submodelo de ϕ com $r = 2$	37
Tabela 12 – Taxa de rejeição da hipótese nula (%) para a o cenário III, submodelo de ϕ com $r = 2$	37
Tabela 13 – Taxa de rejeição da hipótese nula (%) para a o cenário IV, submodelo de ϕ com $r = 2$	38
Tabela 14 – Poder dos testes (%), submodelo de μ para a o cenário I.	40
Tabela 15 – Poder dos testes (%), submodelo de μ para a o cenário II.	41
Tabela 16 – Poder dos testes (%), submodelo de μ para a o cenário III.	42
Tabela 17 – Poder dos testes (%), submodelo de μ para a o cenário IV.	43
Tabela 18 – Poder dos testes (%), submodelo de ϕ com $r = 1$ para a o cenário I.	43
Tabela 19 – Poder dos testes (%), submodelo de ϕ com $r = 1$ para a o cenário II.	44
Tabela 20 – Poder dos testes (%), submodelo de ϕ para a o cenário III.	44
Tabela 21 – Poder dos testes (%), submodelo de ϕ com $r = 1$ para a o cenário IV.	45
Tabela 22 – Poder dos testes (%), submodelo de ϕ com $r = 2$ para a o cenário I.	45
Tabela 23 – Poder dos testes (%), submodelo de ϕ com $r = 2$ para a o cenário II.	46
Tabela 24 – Poder dos testes (%), submodelo de ϕ com $r = 2$ para a o cenário III.	46

Tabela 25 – Poder dos testes (%), submodelo de ϕ com $r = 2$ para a o cenário IV. . 47

Lista de abreviaturas e siglas

BP	Beta Prime
MLGs	Modelos lineares generalizados
FDP	Função densidade de probabilidade
EMV	Estimadores de máxima verossimilhança
RV	Razão de verossimilhanças
R	Razão de verossimilhanças Sinalizada
S	Escore
W	Wald
G	Gradiente

Sumário

	Lista de ilustrações	9
1	INTRODUÇÃO	14
1.1	Organização do trabalho	15
2	DISTRIBUIÇÃO BETA PRIME	16
3	MODELO DE REGRESSÃO BETA PRIME	21
4	TESTES DE HIPÓTESES	23
4.1	Introdução	23
4.2	Teste da Razão de Verossimilhanças	23
4.3	Teste da Razão de Verossimilhanças Sinalizada	25
4.4	Teste Escore	25
4.5	Teste Wald	26
4.6	Teste Gradiente	27
5	TESTES DE HIPÓTESES PARA O MODELO DE REGRESSÃO BETA PRIME	28
6	ESTUDO DE SIMULAÇÃO	30
6.1	Simulação de Tamanho	31
6.2	Simulação de Poder	38
7	CONSIDERAÇÕES FINAIS	48
7.1	Conclusão	48
	REFERÊNCIAS	49
A	APÊNDICE	51
A.1	Vetor Escore	51
A.2	Matriz de Informação de Fisher	52

1 Introdução

Os modelos de regressão são amplamente utilizados em diversas áreas do conhecimento para análise de dados. O objetivo desta técnica é obter uma equação que explique satisfatoriamente a relação entre uma variável resposta e uma ou mais variáveis explicativas, possibilitando fazer predição de valores da variável de interesse. Porém, em muitas situações práticas, algumas das suposições necessárias para utilização dos modelos clássicos de regressão não são satisfeitas, como normalidade e homocedasticidade. Este fato gerou a necessidade de criar novos modelos de regressão, como os modelos lineares generalizados (MLGs) introduzidos na literatura por NELDER; WEDDERBURN (1972).

Neste sentido, BOURGUIGNON; SANTOS-NETO; CASTRO (2018) propuseram um modelo de regressão adaptado para situações em que a variável de interesse é contínua e pertence aos reais positivos. Neste modelo foi proposta uma nova parametrização em termos dos parâmetros de média e precisão em que a variável resposta apresenta distribuição beta prime (*BP*). Além disso, a variável de interesse relaciona-se com as outras variáveis por meio dos parâmetros de média e precisão, como nos modelos lineares generalizados.

Um dos objetivos principais da análise de modelos de regressão é explicar a relação entre as variáveis do estudo de modo que podemos identificar o quanto variáveis independentes influenciam as variáveis dependentes e, através dessa relação, realizar previsões dos valores da variável de interesse. Porém, antes de realizar qualquer análise sobre a modelagem estatística, é importante se certificar que o modelo foi bem ajustado para que não acarrete em inferências imprecisas a respeito dos dados (QUEIROZ, 2011). Sendo assim, uma etapa usualmente realizada após a estimação dos parâmetros é a realização de testes de hipóteses para confirmar ou refutar hipóteses sobre eles. Os testes mais aplicados na literatura são o teste da razão de verossimilhanças, teste de Wald. Além desses testes, este trabalho aborda os testes da razão de verossimilhança sinalizada e gradiente.

Segundo QUEIROZ (2011), os testes citados anteriormente são realizados a partir de valores críticos válidos para grandes amostras, o que pode conduzir a probabilidades de erro tipo I distorcidas em amostras finitas. Portanto, ao escolher um teste de hipótese para avaliação de um modelo de regressão, é importante que o mesmo não apresente distorções de tamanho e seja poderoso quando avaliado em amostras finitas. Dessa forma, este trabalho tem como objetivo avaliar numericamente o desempenho dos testes de hipóteses para o modelo de regressão *BP*, considerado um modelo de regressão novo na literatura. Por meio de simulações de Monte Carlo, avaliamos o desempenho dos testes através das

taxas de rejeição da hipótese nula (tamanho) e das taxas de rejeição da hipótese alternativa (poder). Além disso, foram considerados testes de hipóteses sobre os vetores de parâmetros que indexam os submodelos da média e da precisão.

1.1 Organização do trabalho

Este trabalho é composto por mais seis capítulos, como segue:

- No Capítulo 2 apresentamos a distribuição *BP*;
- No Capítulo 3 abordamos o modelo de regressão *BP*;
- No Capítulo 4 apresentamos os testes de hipóteses considerados;
- No Capítulo 5 apresentamos os testes de hipóteses propostos para o modelo de regressão *BP*;
- No Capítulo 6 avaliamos os testes abordados em amostras finitas, considerando diferentes cenários, por meio de simulações de Monte Carlo;
- Por fim, no Capítulo 7 são apresentadas as considerações finais.

2 Distribuição Beta Prime

A distribuição beta prime (*BP*) é uma distribuição de probabilidade absolutamente contínua com suporte nos reais positivos inicialmente abordada por KEEPING (1962) e MCDONALD (1984).

A distribuição ficou conhecida na literatura como distribuição beta inversa ou distribuição beta do tipo dois devido a importantes estudos desenvolvidos por ABERNATHY; SMITH (1993) e MCDONALD; XU (1995), respectivamente. Porém, ao longo dos anos poucos trabalhos foram desenvolvidos envolvendo esta distribuição. Dos trabalhos existentes na literatura podemos citar, por exemplo, MCDONALD (1987) que apresentou e discutiu algumas propriedades desta distribuição e obteve os estimadores de máxima verossimilhança para os parâmetros. Adicionalmente, em MCDONALD; BUTLER (1990) e, mais recente, em TULUPYEV et al. (2013) a distribuição foi utilizada no contexto de modelos de regressão para variáveis aleatórias positivas. Além disso, segundo alguns autores recentes como LEÃO et al. (2021), a distribuição BP pode ser utilizada em estudos de análise de sobrevivência para mensurar dados sobre o tempo de ocorrência de eventos de interesse. No entanto, esses trabalhos consideraram a parametrização usual da distribuição *BP*, em função dos parâmetros de forma.

Seja Y uma variável aleatória que segue distribuição *BP*, $Y \sim BP(\alpha, \beta)$ cuja função densidade de probabilidade (fdp) (KEEPING,1962) é dada por

$$f(y; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} y^{\alpha-1} (1+y)^{-(\alpha+\beta)}, \quad y > 0 \quad (2.1)$$

em que $\alpha > 0$, $\beta > 0$ são parâmetros de forma e $B(\alpha, \beta)$ é a função beta dada por

$$\int_0^\infty t^{\alpha-1} (1+t)^{-(\alpha+\beta)} dt = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)},$$

e $\Gamma(\cdot)$ a função gama definida por

$$\Gamma(a) = \int_0^\infty e^{-u} u^{a-1} du.$$

A função de distribuição acumulada da distribuição *BP* é da forma

$$F(y; \alpha, \beta) = I_{\frac{y}{1+y}}(\alpha, \beta),$$

em que

$$I_x(\alpha, \beta) = \frac{B(x; \alpha, \beta)}{B(\alpha, \beta)}$$

é a função beta incompleta regularizada com

$$B(x; \alpha, \beta) = \int_0^x t^{\alpha-1} (1+t)^{-(\alpha+\beta)} dt, \quad y > 0.$$

A função densidade de probabilidade da distribuição *BP* (2.1) é decrescente com $f(y; \alpha, \beta) \rightarrow \infty$ ao passo que $y \rightarrow 0$, se $0 < \alpha < 1$; quando $\alpha = 1$, $f(y; \alpha, \beta)$ é decrescente com moda em $y = 0$, no caso em que $\alpha > 1$, $f(y; \alpha, \beta)$ cresce e depois decresce com moda igual a $(\alpha - 1)/(\beta + 1)$. Adicionalmente para $0 < \alpha < 1$, $f(y; \alpha, \beta)$ é côncava; se $1 < \alpha \leq 2$, $f(y; \alpha, \beta)$ é convexa para baixo e, em seguida, para cima com ponto de inflexão x_2 ; e para $\alpha > 2$, $f(y; \alpha, \beta)$ é côncava para cima e então para baixo e novamente para cima com pontos de inflexão em x_1 e x_2 , em que

$$x_1 = \frac{(\alpha - 1)(\alpha + 2) - \sqrt{(\alpha - 1)(\beta + 2)(\alpha + \beta)}}{(\beta + 2)(\beta + 1)}$$

e

$$x_2 = \frac{(\alpha - 1)(\alpha + 2) + \sqrt{(\alpha - 1)(\beta + 2)(\alpha + \beta)}}{(\beta + 2)(\beta + 1)}.$$

O r -ésimo momento de uma variável aleatória Y que segue distribuição *BP* pode ser descrito da forma

$$E[Y^r] = \frac{B(\alpha + r, \beta - r)}{B(\alpha, \beta)}, \quad -\alpha < r < \beta.$$

Para $r \in \mathbf{N}$ e $r < \beta$ o r -ésimo momento pode ser simplificado por

$$E[Y^r] = \prod_{i=1}^r \frac{\alpha + i - 1}{\beta - i}.$$

Dessa forma, a esperança e a variância de uma variável aleatória Y que segue uma distribuição *BP* são expressas da seguinte forma

$$E[Y] = \frac{\alpha}{\beta - 1}, \quad \beta > 1, \quad \text{e} \quad \text{Var}(Y) = \frac{\alpha(\alpha + \beta - 1)}{(\beta - 2)(\beta - 1)^2}, \quad \beta > 2.$$

A distribuição *BP* apresenta relação com algumas outras distribuições, que são:

- Se $Y \sim F(2\alpha, 2\beta)$, então $\frac{\alpha}{\beta}Y \sim BP(\alpha, \beta)$;

- Se $Z \sim B(\alpha, \beta)$, então $\frac{Z}{1-Z} \sim BP(\alpha, \beta)$;
- Se $V \sim Gama(\alpha, 1)$ e $W \sim Gama(\beta, 1)$, então $\frac{V}{W} \sim BP(\alpha, \beta)$;
- Se Y tem distribuição Lomax (Pareto tipo II), com parâmetros de escala α e de forma λ , então $\frac{Y}{\lambda} \sim BP(1, \alpha)$;
- Se Y tem distribuição Pareto com mínimo y_1 e parâmetro de forma α , então $Y - y_1 \sim BP(1, \alpha)$;
- Se Y tem distribuição Pareto padrão tipo IV com parâmetro de forma α e parâmetro de desigualdade γ , então $Y^{\frac{1}{\alpha}} \sim BP(1, \gamma)$.

Em BOURGUIGNON; SANTOS-NETO; CASTRO (2018) a distribuição BP foi estudada e utilizada em um contexto de modelo de regressão baseado em uma nova parametrização, em que são consideradas duas estruturas de regressão para a média e precisão separadamente. O modelo de regressão proposto é conveniente para modelar dados assimétricos e é uma alternativa aos MLG's. Neste estudo, os autores propuseram uma parametrização da distribuição BP em termo dos parâmetros da média e da precisão, considerando $\mu = \frac{\alpha}{\beta-1}$ e $\phi = \beta - 2$, sendo assim $\alpha = \mu(1 + \phi)$ e $\beta = 2 + \phi$. Dessa forma, considerando essa nova parametrização, temos que a média e variância são dadas por

$$E(Y) = \mu \quad \text{e} \quad Var(Y) = \frac{\mu(1 + \mu)}{\phi}$$

respectivamente.

Portanto, de acordo com a nova parametrização, uma variável aleatória Y que segue distribuição BP com média μ e parâmetro de precisão ϕ é denotada por $Y \sim BP(\mu, \phi)$. Conforme visto acima, ϕ pode ser interpretado como um parâmetro de precisão uma vez que, para μ fixo, quanto maior o valor de ϕ , menor a variância de Y . A função densidade de probabilidade, considerando a nova parametrização, pode ser escrita da seguinte forma

$$f(y; \mu, \phi) = \frac{y^{\mu(\phi+1)-1} (1+y)^{-[\mu(\phi+1)+\phi+2]}}{B(\mu(1+\phi), \phi+2)}, \quad y > 0, \quad (2.2)$$

em que $\mu > 0$ e $\phi > 0$.

As Figuras 1, 2 e 3 mostram gráficos da função densidade em (2.2) para alguns valores de parâmetros. Dessa forma, fica evidente que a distribuição é muito flexível e pode ser uma alternativa interessante para outras distribuições com suporte nos reais positivos.

No entanto, é importante citar que a distribuição BP pode ser escrita como pertencente à família exponencial das distribuições de acordo com a observação feita por BOURGUIGNON; SANTOS-NETO; CASTRO, 2018.

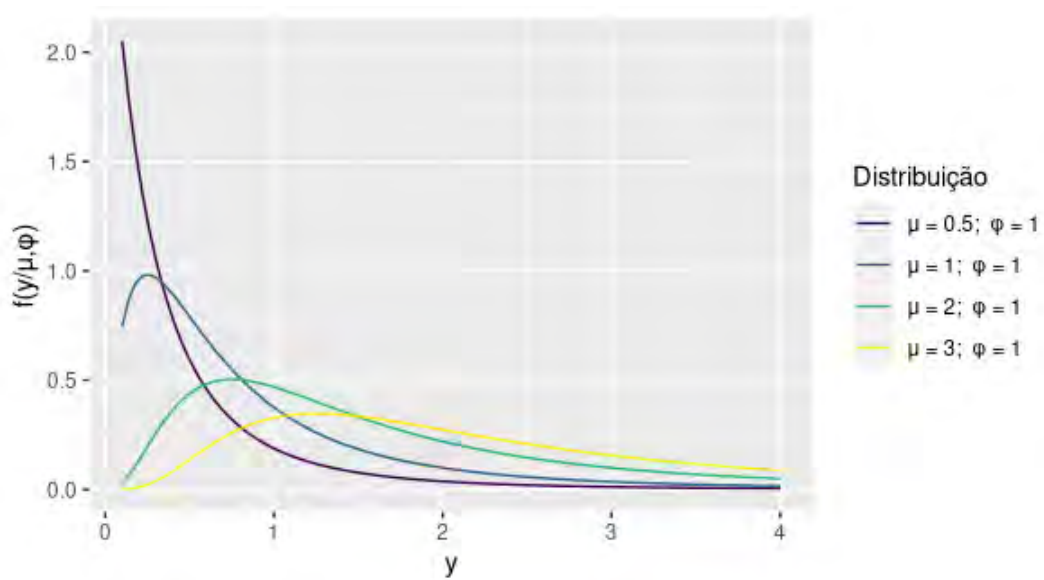


Figura 1 – Gráfico da função densidade de probabilidade da distribuição *BP* considerando $\phi = 1$ e distintos valores para μ

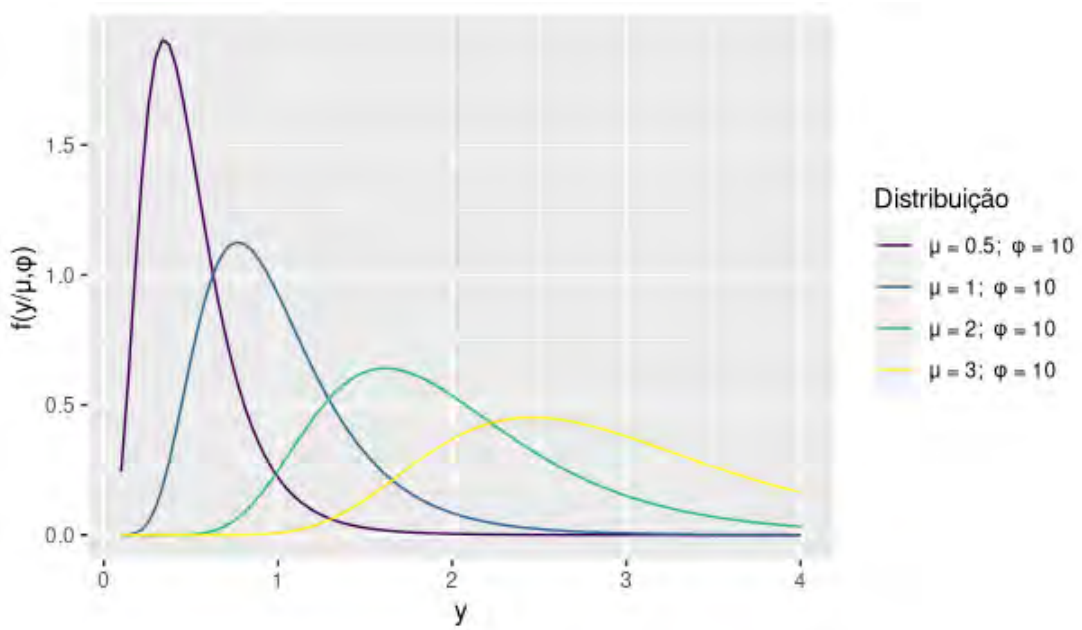


Figura 2 – Gráfico da função densidade de probabilidade da distribuição *BP* considerando $\phi = 10$ e distintos valores para μ .

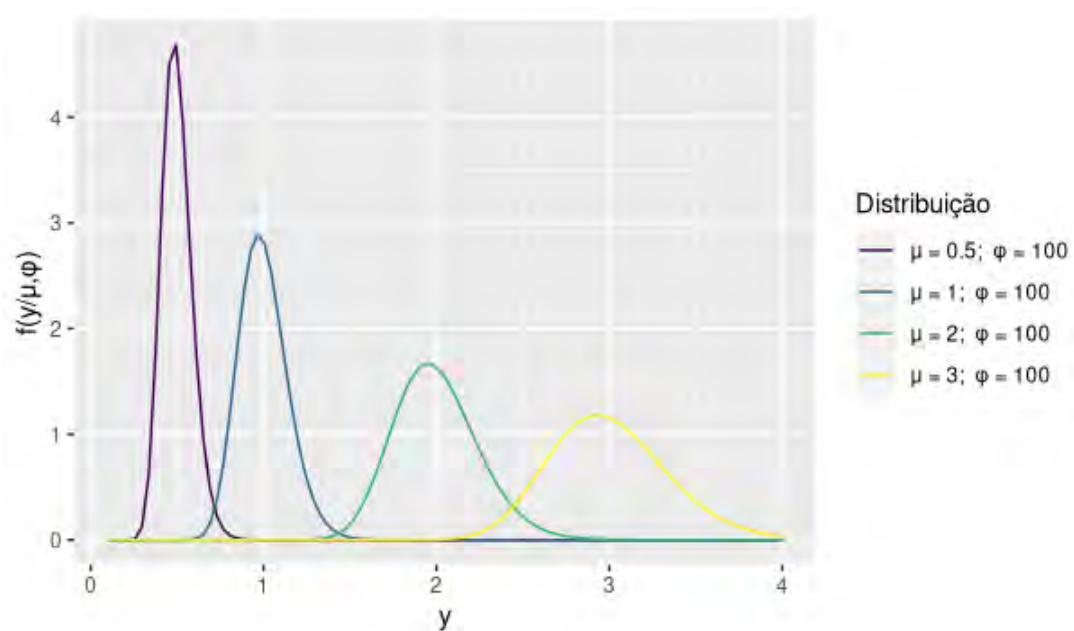


Figura 3 – Gráfico da função densidade de probabilidade da distribuição *BP* considerando $\phi = 100$ e distintos valores para μ .

3 Modelo de Regressão Beta Prime

Como descrito no capítulo anterior, modelos de regressão **BP** para variáveis positivas foram utilizados pelos autores MCDONALD; BUTLER (1990) e TULUPYEV et al. (2013) considerando a parametrização usual da distribuição **BP** (2.1). Entretanto, neste estudo será considerado o modelo de regressão **BP** utilizado por BOURGUIGNON; SANTOS-NETO; CASTRO (2018) considerando a nova parametrização da distribuição **BP** (2.2).

Sejam Y_1, \dots, Y_n variáveis aleatórias independentes, em que cada $Y_i, i = 1, \dots, n$ tem fdp expressa em (2.2) com média μ_i e parâmetro de precisão ϕ_i . Considerando que a média e o parâmetro de precisão satisfazem as seguintes relações

$$g_1(\mu_i) = \eta_i = \mathbf{x}_i^\top \boldsymbol{\beta} \quad \text{e} \quad g_2(\phi_i) = \kappa_i = \mathbf{z}_i^\top \boldsymbol{\lambda}, \quad i = 1, \dots, n. \quad (3.1)$$

em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_m)^\top$ e $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_M)^\top$ são vetores de parâmetros de regressão desconhecidos, $\boldsymbol{\beta} \in \mathbb{R}^m$ e $\boldsymbol{\lambda} \in \mathbb{R}^M$ com $m + M < n$, η_i e κ_i são os preditores lineares, $\mathbf{x}_i = (x_{i1}, \dots, x_{im})^\top$ e $\mathbf{z}_i = (z_{i1}, \dots, z_{iM})^\top$ são observações de m e M regressores conhecidos, para $i = 1, \dots, n$. As funções de ligação $g_1 : \mathbb{R} \rightarrow \mathbb{R}^+$ e $g_2 : \mathbb{R} \rightarrow \mathbb{R}^+$ em (6.1) devem ser estritamente monótonas, positivas e ao menos duas vezes diferenciáveis, de forma que $\mu_i = g_1^{-1}(\mathbf{x}_i^\top \boldsymbol{\beta})$ e $\phi_i = g_2^{-1}(\mathbf{z}_i^\top \boldsymbol{\lambda})$. As funções de ligação que podem ser empregadas nesse modelo são a raiz quadrada e logarítmica para μ e ϕ . Este estudo consideramos a função de ligação logarítmica, em que $\mu_i = g_1^{-1}(\eta_i) = \exp(\eta_i)$ e $\phi_i = g_2^{-1}(\kappa_i) = \exp(\kappa_i)$.

A função de log-verossimilhança de 2.2 é da forma

$$\ell(\boldsymbol{\beta}, \boldsymbol{\lambda}) = \sum_{i=1}^n \ell(\mu_i, \phi_i), \quad i = 1, \dots, n, \quad (3.2)$$

em que

$$\begin{aligned} \ell(\mu_i, \phi_i) = & [\mu_i(1 + \phi_i) - 1] \log(y_i) - [\mu_i(1 + \phi_i) + \phi_i + 2] \log(1 + y_i) - \log[\Gamma(\mu_i(1 + \phi_i))] \\ & - \log[\Gamma(\phi_i + 2)] + \log[\Gamma(\mu_i(1 + \phi_i) + \phi_i + 2)]. \end{aligned}$$

Segundo BOURGUIGNON; SANTOS-NETO; CASTRO (2018), o vetor escore, obtido diferenciando a função de log-verossimilhança (3.2), pode ser escrito na forma matricial para $\boldsymbol{\beta}$ e $\boldsymbol{\lambda}$, respectivamente, como

$$\begin{aligned} U_{\boldsymbol{\beta}} &= \mathbf{X}^\top \boldsymbol{\Phi} \mathbf{D}_1(\mathbf{y}^* - \boldsymbol{\mu}^*), \\ U_{\boldsymbol{\lambda}} &= \mathbf{Z}^\top \mathbf{D}_2(\mathbf{y}^* - \boldsymbol{\mu}^*) \end{aligned}$$

em que $\mathbf{X} = (x_1, \dots, x_n)^\top$ e $\mathbf{Z} = (z_1, \dots, z_n)^\top$ são matrizes de covariáveis com posto m e M , respectivamente. $\mathbf{D}_1 = \text{diag}\left(\frac{\partial \mu_1}{\partial \eta_{11}}, \dots, \frac{\partial \mu_n}{\partial \eta_{1n}}\right)$, $\mathbf{D}_2 = \text{diag}\left(\frac{\partial \phi_1}{\partial \eta_{21}}, \dots, \frac{\partial \phi_n}{\partial \eta_{2n}}\right)$, $\Phi = \text{diag}[(1 + \phi_1), \dots, (1 + \phi_n)]$, $\mathbf{y}^* = (y_1^*, \dots, y_n^*)^\top$ em que $y_i^* = \log\left(\frac{y_i}{1+y_i}\right)$, $\mathbf{y}^* = (y_1^*, \dots, y_n^*)^\top$ em que $y_i^* = \log\left(\frac{y_i^{\mu_i}}{(1+y_i)^{1+\mu_i}}\right)$, $\boldsymbol{\mu}^* = (\mu_1^*, \dots, \mu_n^*)^\top$ em que $\mu_i^* = [\Psi^{(0)}(\mu_i(1 + \phi_i)) - \Psi^{(0)}(\mu_i(1 + \phi_i) + \phi_i + 2)]$ e $\boldsymbol{\mu}^* = (\mu_1^*, \dots, \mu_n^*)^\top$ em que $\mu_i^* = \mu_i\left(\mu_i^* - \frac{\gamma_i}{\mu_i}\right)$ com $\gamma_i = \Psi^{(0)}(\mu_i(1 + \phi_i) + \phi_i + 2) - \Psi^{(0)}(\phi_i + 2)$. Para mais detalhes, ver (A.1).

A matriz de informação de Fisher (BOURGUIGNON; SANTOS-NETO; CASTRO, 2018) pode ser expressa da seguinte forma

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_{\beta\beta} & \mathbf{K}_{\beta\mathbf{v}} \\ \mathbf{K}_{\mathbf{v}\beta} & \mathbf{K}_{\mathbf{v}\mathbf{v}} \end{bmatrix}, \quad (3.3)$$

em que $\mathbf{K}_{\beta\beta} = \mathbf{X}^\top E_3 \mathbf{X}$, $\mathbf{K}_{\beta\mathbf{v}} = (\mathbf{K}_{\mathbf{v}\beta})^\top = \mathbf{Z}^\top E_5 \mathbf{X}$ e $\mathbf{K}_{\mathbf{v}\mathbf{v}} = \mathbf{Z}^\top E_4 \mathbf{Z}$. O i -ésimo elemento de E_3 , E_4 e E_5 são dados, respectivamente, por

$$e_{3i} = (1 + \phi_i)^2 [\Psi^{(1)}(\mu_i(1 + \phi_i)) - \Psi^{(1)}(\mu_i(1 + \phi_i) + \phi_i + 2)] \left(\frac{\partial \mu_i}{\partial \eta_{1i}}\right)^2,$$

$$e_{4i} = [\mu_i^2 \Psi^{(1)}(\mu_i(1 + \phi_i)) - (1 + \mu_i)^2 \Psi^{(1)}(\mu_i(1 + \phi_i) + \phi_i + 2) + \Psi^{(1)}(\phi_i + 2)] \left(\frac{\partial \phi_i}{\partial \eta_{2i}}\right)^2,$$

e

$$e_{5i} = -(1 + \phi_i) \{ \Psi^{(1)}(\mu_i(1 + \phi_i) + \phi_i + 2) + \mu_i [\Psi^{(1)}(\mu_i(1 + \phi_i) + \phi_i + 2) - \Psi^{(1)}(\mu_i(1 + \phi_i))] \} \left(\frac{\partial \mu_i}{\partial \eta_{1i}}\right) \left(\frac{\partial \phi_i}{\partial \eta_{2i}}\right).$$

em que, $\psi(\cdot) = \frac{d \log[\Gamma(\cdot)]}{d}$ é a função digamma.

Para mais detalhes, ver (A.2).

4 Testes de hipóteses

4.1 Introdução

Um dos principais usos de teste de hipótese consiste em comparar um modelo de regressão sob a hipótese nula (modelo restrito) com o modelo sob a hipótese alternativa (irrestrito). Em outras palavras, realiza-se uma comparação entre um modelo considerando todos os parâmetros de interesse e um modelo obtido por meio de restrições do espaço paramétrico (OLIVEIRA et al., 2016).

A estatística de teste é uma variável aleatória obtida a partir de dados amostrais e utilizada em testes de hipótese para determinar se a hipótese nula deve ser ou não rejeitada, ou seja, se há evidência estatística de acordo com os dados amostrais de que as hipóteses feitas sobre os parâmetros populacionais são verdadeiras.

Neste estudo, o interesse é testar hipóteses sobre os parâmetros do modelo de regressão beta prime. Como regra de decisão, se o valor da estatística de teste for maior que o valor crítico da distribuição da estatística de teste sob a hipótese nula, rejeita-se a hipótese nula, caso contrário não rejeita-se tal hipótese.

Considerando que o vetor de parâmetros θ pode ser escrito por $\theta = (\theta_1^\top, \theta_2^\top)^\top$, sendo θ_1 o vetor de parâmetros de interesse de dimensão r e θ_2^\top o vetor de parâmetros de perturbação de dimensão $p - r$. Considere as hipóteses

$$H_0 : \theta_1 = \theta_1^{(0)} \quad \text{versus} \quad H_1 : \theta_1 \neq \theta_1^{(0)} \quad (4.1)$$

em que $\theta_1^{(0)} \in \Theta_0$, o espaço paramétrico restrito $\Theta_0 \subset \Theta$.

Neste capítulo, são apresentados os seguintes testes assintóticos: razão de verossimilhanças (RV) (NEYMAN; PEARSON, 1928), escore (S) (RAO, 1948), Wald (W) (WALD, 1943), gradiente (G) (TERRELL, 2002) e razão de verossimilhanças sinalizada (R) (SEVERINI, 2000).

4.2 Teste da Razão de Verossimilhanças

Testes baseados na função de verossimilhança são amplamente empregados dentre as diversas técnicas para avaliação de modelos estatísticos. Este fato ocorre devido a sua simplicidade e eficiência para realizar inferência acerca de um vetor de parâmetros de interesse. Segundo CORDEIRO (1999), a função de verossimilhança "informa a ordem natural de preferência entre diversas possibilidades de θ . Generalizando, entre os possíveis

candidatos para estimar o parâmetro verdadeiro a partir dos dados, o vetor de parâmetros mais plausível é aquele de maior verossimilhança”.

O teste da razão de verossimilhanças está naturalmente associado aos estimadores de máxima verossimilhança. Tal teste é usualmente empregado para testar a hipótese nula de que um determinado vetor de parâmetros satisfaz ou não certa restrição, ou ainda para testar a hipótese nula de se um modelo restrito é ou não apropriado. Neste segundo caso, poderíamos dizer que o teste também cumpre a função de comparar modelos.

Considerando as variáveis aleatórias Y_1, \dots, Y_n com função de densidade dada por $f_n(y_1, \dots, y_n | \theta)$ em que θ representa um vetor de parâmetros, podendo ser escalar ou vetorial, ou seja, pertence ao espaço paramétrico Θ . Sendo $y = (y_1, \dots, y_n)$ os valores observados da variável aleatória $Y = (Y_1, \dots, Y_n)^T$ e considerando que as variáveis são i.i.d., a função de verossimilhança é definida como:

$$L(\theta) = f_n(y_1, \dots, y_n | \theta) = \prod_{i=1}^n f(y_i | \theta). \quad (4.2)$$

O estimador de máxima verossimilhança (EMV) de θ é o valor que maximiza $L(\theta)$.

Definição 1. Assumindo y_1, \dots, y_n uma realização de Y_1, \dots, Y_n , um teste de razão de verossimilhança (TRV) para $H_0 : \theta \in \Theta_0$ contra $H_1 : \theta \in \Theta_0^c$ é qualquer teste que tem uma região de forma $\{y \in \text{Im}Y : \lambda(y) \leq c\}$. Sendo que a estatística do teste é

$$\lambda(Y) = \frac{\sup_{\Theta_0} L(\theta | Y)}{\sup_{\Theta} L(\theta | Y)}. \quad (4.3)$$

O entendimento do TRV pode ser melhor entendido quando $f(y | \theta)$ é a função de probabilidade de uma variável aleatória discreta. Neste caso, o numerador de $\lambda(y)$ representa a probabilidade máxima de Y_1, \dots, Y_n assumirem os valores da amostra observada y_1, \dots, y_n sendo válida a hipótese nula. O denominador de $\lambda(Y)$ representa a probabilidade máxima de Y_1, \dots, Y_n assumirem os valores da amostra observada y_1, \dots, y_n sendo a hipótese nula verdadeira ou não.

Observe que $0 \leq \lambda(y) \leq 1$ pois o numerador é o supremo com relação a $\theta \in \Theta_0$ (subconjunto de Θ), enquanto que o denominador é o supremo com relação a todo o conjunto Θ . Assim, deve estar próximo de 1 se H_0 for verdadeira e, caso contrário, o denominador deve ser grande em relação ao numerador e $\lambda(Y)$ deve ser próximo de zero.

Existe uma relação entre TRV's e EMV's. Suponha que existam $\hat{\theta}$ e $\hat{\theta}_0$, respectivamente estimadores de θ em Θ [Estimador Irrestrito] e em Θ_0 [Estimador Restrito]. A

estatística TRV é obtida por:

$$\lambda(Y) = \frac{L(\hat{\theta}_0|Y)}{L(\hat{\theta}|Y)}. \quad (4.4)$$

A equação acima pode apresentar a seguinte notação:

$$\lambda(Y) = \frac{L(\tilde{\theta})}{L(\hat{\theta})}. \quad (4.5)$$

Esta estatística necessita de transformações especiais no caso de amostras finitas. A distribuição de $\lambda(Y)$ pode ser complicada e não ter distribuição conhecida. Para grandes amostras a estatística de teste:

$$RV = -2\log(\lambda(Y)) = 2[\ell(\hat{\theta}) - \ell(\tilde{\theta})]. \quad (4.6)$$

tem distribuição assintótica qui-quadrado com r graus de liberdade, χ_r^2 , sendo r o número de restrições impostas em H_0 . logo, rejeita-se H_0 quando $RV > \chi_{r,(1-\alpha)}^2$, ao nível de significância α . A equação (4.6) pode ser utilizada quando o parâmetro de interesse é um vetor ou escalar.

4.3 Teste da Razão de Verossimilhanças Sinalizada

Com o objetivo de melhorar as aproximações obtidas para os valores críticos em amostras finitas, ao longo dos anos foram propostas diversas correções para estatísticas de testes assintóticos. Nesse contexto, SEVERINI (2000) apresentou a estatística da razão de verossimilhança sinalizada dada por

$$R = \text{sin}(\hat{\theta}_1 - \theta_1^{(0)})\sqrt{RV}, \quad (4.7)$$

em que $\hat{\theta}_1$ é o estimador de máxima verossimilhança irrestrito, $\theta_1^{(0)}$ é o valor do parâmetro sob a hipótese nula e RV é a estatística da razão de verossimilhanças. Logo, rejeita-se H_0 quando $R > \Phi^{-1}(1 - \alpha)$, ao nível de significância α , em que Φ^{-1} é o quantil da distribuição normal padrão. A equação (4.7) pode ser utilizada para os casos em que o parâmetros de interesse é escalar.

4.4 Teste Escore

Outro teste comumente usado para testar hipóteses sobre conjuntos de parâmetros é o teste escore proposto por RAO (1948). Para o caso em que o parâmetro de interesse

é um escalar, considerando $H_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_1^{(0)}$ e $H_1 : \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_1^{(0)}$, a estatística de teste escore é dada por:

$$S = \left[U(\boldsymbol{\theta}_1^{(0)}) \right]^2 K^{-1}(\boldsymbol{\theta}_1^{(0)}), \quad (4.8)$$

em que $U(\boldsymbol{\theta})$ é a função escore e $K^{-1}(\boldsymbol{\theta})$ é a inversa da matriz de informação de Fisher K avaliadas em $\boldsymbol{\theta}_1^{(0)}$. Considerando certas condições de regularidade e sob H_0 , S tem distribuição assintótica χ_1^2 . Logo, rejeita-se H_0 ao nível de significância δ se $S > \chi_{1,(1-\alpha)}^2$.

No caso em que $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$ e deseja-se testar $H_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_1^{(0)}$ contra $H_1 : \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_1^{(0)}$, sendo $\boldsymbol{\theta}_1$ um vetor $r \times 1$ de parâmetros de interesse e $\boldsymbol{\theta}_2$ um vetor de $(p - r) \times 1$ de parâmetros de perturbação, a estatística escore é

$$S = \tilde{U}_r^\top \tilde{K}_{rr}^{\theta\theta} \tilde{U}_r, \quad (4.9)$$

em que $U_r(\cdot)$ denota o vetor $r \times 1$ que contém r elementos da função escore e $K_{rr}(\cdot)$ é a matriz $r \times r$ que contém os elementos da inversa da informação de Fisher referentes aos r parâmetros testados, ambos avaliados em $\tilde{\boldsymbol{\theta}}_1 = (\boldsymbol{\theta}_1^{(0)\top}, \tilde{\boldsymbol{\theta}}_2^\top)$, sendo $\boldsymbol{\theta}_1^{(0)}$ dado e $\tilde{\boldsymbol{\theta}}_2$ o vetor de parâmetros estimador restrito de $\boldsymbol{\theta}_2$.

Nesse contexto, sob H_0 e quando são atendidas certas condições de regularidade, em que r é o número de restrições em teste, S tem distribuição assintótica χ_r^2 . Rejeita-se H_0 , se $S > \chi_{r,(1-\alpha)}^2$, ao nível de significância α .

4.5 Teste Wald

O teste Wald também pode ser utilizado para testar restrições sobre $\boldsymbol{\theta}$ baseado na distribuição assintótica de $\hat{\boldsymbol{\theta}}$. Este teste é uma generalização do teste t de Student e recebeu esse nome em homenagem ao seu desenvolvedor, o estatístico WALD (1943). A estatística de teste para $H_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_1^{(0)}$ contra $H_1 : \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_1^{(0)}$ pode ser expressa como

$$W = (\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^{(0)})^2 K(\hat{\boldsymbol{\theta}}_1), \quad (4.10)$$

quando $\boldsymbol{\theta}_1$ é escalar, em que $\hat{\boldsymbol{\theta}}_1$ é estimador de máxima verossimilhança de $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_1^{(0)}$ é dado e $K(\hat{\boldsymbol{\theta}}_1)$ é a informação de Fisher avaliada no EMV irrestrito. Nesse contexto, W tem distribuição assintótica χ_1^2 . Rejeita-se H_0 , se $W > \chi_{1,(1-\alpha)}^2$, ao nível de significância α .

Da mesma forma, quando deseja-se testar a hipótese nula, sendo $\boldsymbol{\theta}_1$ um vetor de interesse levando em consideração $\boldsymbol{\theta}_2$ um vetor $(p - r) \times 1$ de parâmetros de perturbação,

a estatística de teste Wald é dada por

$$W = (\hat{\theta}_1 - \theta_1^{(0)})^\top [K^{rr}(\hat{\theta})]^{-1} (\hat{\theta}_1 - \theta_1^{(0)}), \quad (4.11)$$

em que $K^{rr}(\hat{\theta})$ é o bloco $r \times r$ da matriz de informação de Fisher referente a θ avaliado no EMV irrestrito.

Assim como nos testes anteriores, considerando as condições gerais de regularidade e sob a hipótese nula, W tem distribuição assintótica χ_r^2 . Rejeita-se H_0 , se $W > \chi_{r,(1-\alpha)}^2$, ao nível de significância α .

4.6 Teste Gradiente

O teste gradiente foi proposto por TERRELL (2002). Derivada a partir da estatística escore e Wald modificada (HAYAKAWA; PURI, 1985), a estatística gradiente apresenta um cálculo mais simples, sem envolver produto e inversa de matrizes, por exemplo. Esta estatística foi citada por RAO (RAO) em um artigo em que diz: "A sugestão de Terrel é atraente por ser mais simples de calcular. Seria de interesse investigar o desempenho da estatística [gradiente]."

De acordo com a proposta de TERRELL (2002) quando θ_1 é escalar, temos

$$G = U(\theta_1^{(0)})(\hat{\theta}_1 - \theta_1^{(0)}) \quad (4.12)$$

em que $U(\theta)$ é a função escore, $\theta_1^{(0)}$ é o valor do parâmetro sob a hipótese nula e $\hat{\theta}_1$ é estimador do parâmetro de interesse.

No caso em que θ_1 é um vetor de parâmetros de interesse, a estatística G é dada por

$$G = \tilde{U}_r^\top (\hat{\theta}_1 - \theta_1^{(0)}) \quad (4.13)$$

em que U_r é o vetor $r \times 1$ que contém r elementos da função escore avaliado no EMV restrito, $\theta_1^{(0)}$ é o vetor de parâmetros sob a hipótese nula e $\hat{\theta}_1$ é vetor de estimadores irrestritos do parâmetro de interesse.

Assim como nos testes escore e Wald, considerando as condições gerais de regularidade e sob a hipótese nula, G têm distribuição assintótica χ_r^2 . Rejeita-se H_0 , se $G > \chi_{r,(1-\alpha)}^2$, ao nível de significância α .

5 Testes de hipóteses para o modelo de regressão Beta Prime

Nesse trabalho, serão abordados testes de hipóteses sobre os vetores de parâmetros β e λ do modelo de regressão beta prime baseados em 4.1. As hipóteses são apresentadas a seguir.

$$H_0 : \beta_1 = \beta_1^{(0)} \quad \text{versus} \quad H_1 : \beta_1 \neq \beta_1^{(0)} \quad (5.1)$$

e

$$H_0 : \lambda_1 = \lambda_1^{(0)} \quad \text{versus} \quad H_1 : \lambda_1 \neq \lambda_1^{(0)} \quad (5.2)$$

com $\beta = (\beta_1^\top, \beta_2^\top)^\top$, sendo β_1^\top o vetor de parâmetros de interesse de dimensão r e β_2^\top o vetor de parâmetros de perturbação de dimensão $p - r$ para submodelo da média, e $\lambda = (\lambda_1^\top, \lambda_2^\top)^\top$, sendo λ_1^\top o vetor de parâmetros de interesse de dimensão r e λ_2^\top o vetor de parâmetros de perturbação de dimensão $p - r$ para submodelo da precisão.

Dessa forma, sejam Y_1, Y_2, \dots, Y_n variáveis aleatórias independentes, em que cada Y_i , $i = 1, \dots, n$, segue a função de densidade do modelo **BP** (2.2). As estatísticas para testar H_0 (5.1) podem ser escritas da seguinte forma

$$RV = 2[\ell(\widehat{\theta}) - \ell(\widetilde{\theta})], \quad (5.3)$$

$$R = \text{sin}(\widehat{\beta}_1 - \beta_1^{(0)})\sqrt{RV}, \quad (5.4)$$

$$S = \widetilde{U}_{r\beta}^\top \widetilde{K}_{rr}^{\beta\beta} \widetilde{U}_{r\beta}, \quad (5.5)$$

$$W = (\widehat{\beta}_1 - \beta_1^{(0)})\widehat{K}_{rr}^{\beta\beta}(\widehat{\beta}_1 - \beta_1^{(0)}), \quad (5.6)$$

$$G = \widetilde{U}_{r\beta}^\top(\widehat{\beta}_1 - \beta_1^{(0)}), \quad (5.7)$$

em que $\widehat{\theta} = (\widehat{\beta}^\top, \widehat{\lambda}^\top)^\top$ é o estimador de máxima verossimilhança irrestrito de θ e $\widetilde{\theta} = (\widetilde{\beta}^\top, \widetilde{\lambda}^\top)^\top$ é o estimador de máxima verossimilhança restrito de θ , obtido pela imposição da hipótese nula, $\beta_1^{(0)}$ é o vetor de valores sob a hipótese nula, $\widetilde{U}_{r\beta}^\top$ é um vetor coluna r -dimensional contendo os r primeiros elementos da função escore de β avaliados no EMV

restrito, $\widehat{K}_{rr}^{\beta\beta}$ e $\widetilde{K}_{rr}^{\beta\beta}$ são as matrizes $r \times r$ formadas pelas r primeiras linhas e r primeiras colunas da inversa da matriz $K_{\beta\beta}$ avaliadas nos estimadores de máxima verossimilhança irrestrito e restrito, respectivamente.

Já as estatísticas para testar H_0 (5.2) podem ser escritas da seguinte forma

$$RV = 2[\ell(\widehat{\theta}) - \ell(\widetilde{\theta})], \quad (5.8)$$

$$R = \text{sin}l(\widehat{\lambda}_1 - \lambda_1^{(0)})\sqrt{RV}, \quad (5.9)$$

$$S = \widetilde{U}_{r\lambda}^\top \widetilde{K}_{rr}^{\lambda\lambda} \widetilde{U}_{r\lambda}, \quad (5.10)$$

$$W = (\widehat{\lambda}_1 - \lambda_1^{(0)}) \widehat{K}_{rr}^{\lambda\lambda} (\widehat{\lambda}_1 - \lambda_1^{(0)}), \quad (5.11)$$

$$G = \widetilde{U}_{r\lambda}^\top (\widehat{\lambda}_1 - \lambda_1^{(0)}), \quad (5.12)$$

em que $\lambda_1^{(0)}$ é seu vetor de valores sob a hipótese nula, $\widetilde{U}_{r\lambda}^\top$ é um vetor coluna r -dimensional contendo os r primeiros elementos da função escore de λ avaliados no EMV restrito, $\widehat{K}_{rr}^{\lambda\lambda}$ e $\widetilde{K}_{rr}^{\lambda\lambda}$ são as matrizes $r \times r$ formadas pelas r primeiras linhas e r primeiras colunas da inversa da matriz $K_{\lambda\lambda}$ avaliadas nos estimadores de máxima verossimilhança irrestrito e restrito, respectivamente.

Sob condições usuais de regularidade e sob a hipótese nula, a distribuição das estatísticas RV , W , S e G são aproximadamente qui-quadrado r graus de liberdade, χ_r^2 , enquanto R apresenta distribuição aproximadamente normal.

6 Estudo de Simulação

Neste capítulo, através de simulação de Monte Carlo comparamos os desempenhos, em amostras finitas, dos testes de razão de verossimilhança RV , razão de verossimilhança sinalizada R , escore S , Wald W e gradiente G para os parâmetros que indexam a estrutura da média e da precisão do modelo de regressão BP . Todas as simulações foram realizadas utilizando a linguagem de programação R (R Core Team, 2021), 10000 réplicas de Monte Carlo e considerando os níveis nominais de 1%, 5% e 10%. Os critérios de avaliação são o tamanho empírico (probabilidade de rejeitar a hipótese nula quando ela é verdadeira) e o poder empírico (probabilidade de rejeitar a hipótese nula quando ela é falsa).

Consideramos o modelo de regressão BP

$$g_1(\mu_i) = \eta_i = x_i^\top \beta \quad \text{e} \quad g_2(\phi_i) = \kappa_i = z_i^\top \lambda, \quad i = 1, \dots, n. \quad (6.1)$$

em que $g_1(\cdot)$ e $g_2(\cdot)$ são funções de ligação, $\beta = (\beta_0, \beta_1)^\top$ e $\lambda = (\lambda_0, \lambda_1)^\top$ os vetores de coeficientes de regressão desconhecidos e x_i e z_i são vetores de covariáveis. Para cada covariável foram geradas 20 observações $\mathcal{U}(0, 1)$ e replicadas 2, 4, 5 e 10 vezes obtendo tamanhos de amostra 40, 80, 100 e 200. Em cada réplica geramos uma amostra aleatória da variável dependente $Y = (Y_1, \dots, Y_n)^\top$ com $Y_i \sim BP(\mu_i, \phi_i)$, em que $\mu_i = g_1^{-1}(\eta_i) = \exp(\eta_i)$ e $\phi_i = g_2^{-1}(\kappa_i) = \exp(\kappa_i)$.

As estimativas dos parâmetros do modelo de regressão BP (6.1) podem ser obtidas por meio do pacote `gamlss`, contido no R (R Core Team, 2021). Para cada amostra aleatória gerada da variável de interesse $Y_i \sim BP(\mu_i)$ utilizamos a função de geração de números aleatórios `rBP` para a distribuição BP implementada no pacote `glmBP`. O pacote `glmBP` fornece funções para ajustar modelos de regressão BP usando a função `gamlss` e contém uma coleção de rotinas para análise de dados da distribuição de BP . Este pacote está disponível em sua versão atual através do GitHub em `devtools::install_github("santosneto/glmBP")` ((BOURGUIGNON; SANTOS-NETO; CASTRO, 2018)).

Como observado no Capítulo 2, a distribuição BP pode ser bastante flexível a depender dos valores dos parâmetros. Diante disso, neste trabalho foram considerados quatro cenários (Tabela 1): média baixa e precisão alta (cenário I), média baixa e precisão baixa (cenário II), média alta e precisão alta (cenário III) e média alta e precisão baixa (cenário IV).

Adicionalmente, para o submodelo da precisão consideramos o número de restrições, $r = 2$. É importante citar que isso equivale a testar a hipótese de precisão fixa, e para tal situação, removemos a variável explicativa na estrutura do ϕ e consideramos $\lambda_1 = \lambda_2 = 0$.

Tabela 1 – Cenários considerados na simulação.

Cenários	Valores iniciais				Intervalo	
	β_0	β_1	λ_0	λ_1	μ	ϕ
I	1,3	-2,2	4,2	-0,2	(0,58; 3,39)	(54,85; 66,60)
II	1,3	-2,2	1,6	-0,3	(0,58; 3,39)	(3,7; 4,94)
III	3,8	-2,7	4,2	-0,2	(4,67; 40,55)	(54,85; 66,60)
IV	3,8	-2,7	1,6	-0,3	(4,67; 40,55)	(3,7; 4,94)

6.1 Simulação de Tamanho

As simulações de tamanho são realizadas sob a hipótese nula considerando os parâmetros apresentados na Tabela 1, com função de ligação logarítmica para os submodelos da média, μ , e da precisão, ϕ . Nesse cenário, consideramos $\beta_2 = \lambda_2 = 0$ para realizar os testes de hipótese consideramos as estatística de teste RV , S , W , G e R . Diante disso, baseado-se em 5.1 e 5.2, testamos as seguintes hipóteses

$$H_0 : \beta_2 = 0 \quad \textit{versus} \quad H_1 : \beta_2 \neq 0 \quad (6.2)$$

e

$$H_0 : \lambda_2 = 0 \quad \textit{versus} \quad H_1 : \lambda_2 \neq 0 \quad (6.3)$$

Nas simulações de tamanho, as Tabelas 2 a 5 apresentam os resultados dos testes avaliados sobre o vetor de parâmetros do submodelo da média, enquanto as Tabelas 6 a 9 apresentam os resultados dos testes avaliados sobre o vetor de parâmetros do submodelo da precisão considerando $r = 1$ e as Tabelas 10 a 13 considerando $r = 2$.

É esperado que à medida que o tamanho da amostra aumenta, os percentuais de rejeição converjam para os níveis nominais especificados, 1%, 5% e 10%. Este fato é confirmado observando os resultados considerando todas estatísticas de teste para todos os cenários.

Nas simulações de tamanho dos testes para o submodelo da média (μ), alguns testes convergiram mais rapidamente que outros. Por exemplo, na Tabela 2 as estatísticas S , G e R tiveram taxas de rejeição mais próximas dos níveis nominais para tamanhos de amostras pequenos, como $n = 20$ e $n = 40$. Isto também é possível observar para os demais cenários.

Contudo, nas simulações de tamanho para o submodelo de ϕ com $r = 1$ e $r = 2$ podemos observar que os tamanhos do teste Wald não convergiram para os níveis nominais à medida que o tamanho da amostra aumentava, como por exemplo na Tabela 6 em que as taxas de rejeição para tamanhos de amostras grandes como $n = 100$ e $n = 200$ foram 92,52% e 92,71%, respectivamente. Isto também ocorre para todos os cenários.

Dentre os testes para as simulações de tamanho do submodelo de ϕ , temos que as estatísticas S , G e R para $r = 1$ convergiram mais rapidamente para os níveis nominais, como por exemplo na Tabela 8 em que as estatísticas tiveram taxas de rejeição mais próximas dos níveis nominais para tamanhos de amostras pequenos, como $n = 20$ e $n = 40$. Enquanto que para $r = 2$ a estatística S apresentou melhor desempenho comparado as demais para tamanhos de amostras pequenos em todos os cenários, como é possível observar na Tabela 10.

Por sua simplicidade, o teste da razão de verossimilhança original é bastante utilizado na literatura. Contudo, sendo o mesmo notavelmente liberal, é importante citar o ganho pelo emprego do teste na versão sinalizada.

Em geral, nas simulações de tamanho para o submodelo de μ os testes considerando a estatística RV e W foram os que apresentaram o pior desempenho cujos tamanhos convergiram mais lentamente para os níveis nominais adotados. Por exemplo, na Tabela 3 para $n = 20$, o teste baseado na estatística W apresentou taxa de rejeição de 22,38%, 16,13% e 7,72% para os níveis nominais de 10%, 5% e 1%, respectivamente. Enquanto que o teste baseado na estatística RV apresentou taxa de rejeição de 17,85%, 10,72% e 3,72%.

Tabela 2 – Taxa de rejeição da hipótese nula (%) para a o cenário I, submodelo de μ .

$\alpha = 10\%$					
Estatística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	16,8	12,08	11,42	11,21	10,25
S	12,36	10,85	10,77	10,81	9,95
W	21,12	13,57	12,23	11,79	10,58
G	14,62	11,42	11,06	10,91	10,08
R	13,75	11,28	10,87	10,57	10,14
$\alpha = 5\%$					
Estatística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	10,18	6,61	6,04	5,91	5,08
S	6,23	5,34	5,39	5,32	4,85
W	14,24	7,93	6,7	6,41	5,22
G	7,86	5,8	5,63	5,49	4,89
R	8,46	5,95	5,67	5,67	5,32
$\alpha = 1\%$					
Estatística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	3,32	1,52	1,21	1,13	0,99
S	1	0,92	0,88	0,91	0,88
W	6,79	2,46	1,56	1,44	1,17
G	1,4	1,07	0,93	0,95	0,88
R	2,85	1,5	1,34	1,27	1,05

Tabela 3 – Taxa de rejeição da hipótese nula (%) para a o cenário II, submodelo de μ .

$\alpha = 10\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	17,85	12,36	10,95	11,31	10,34
S	11,69	10,72	10,11	10,61	10,02
W	22,38	14,5	12,31	12,51	11,3
G	14,84	11,62	10,49	11,07	10,17
R	14,33	11,46	11,23	10,95	10,2
$\alpha = 5\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	10,72	6,68	5,37	5,67	5,15
S	5,59	5,3	4,66	5,07	4,95
W	16,13	8,78	6,49	6,7	5,72
G	7,56	5,79	5,0	5,34	4,98
R	8,95	6,14	5,9	5,92	5,13
$\alpha = 1\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	3,72	1,46	1,15	1,22	1,15
S	0,76	0,9	0,92	1,05	1,04
W	7,72	2,62	1,54	1,71	1,47
G	1,13	1,03	0,98	1,1	1,07
R	3,03	1,57	1,11	1,34	1,26

Tabela 4 – Taxa de rejeição da hipótese nula (%) para a o cenário III, submodelo de μ .

$\alpha = 10\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	16,22	12,73	11,32	10,84	10,53
S	11,41	11,02	10,59	10,32	10,34
W	20,41	14,06	12,09	11,45	10,93
G	14,13	11,88	10,91	10,45	10,48
R	13,88	11,35	10,54	10,78	10,32
$\alpha = 5\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	9,74	7,16	5,93	5,71	5,43
S	6,33	6	5,32	5,29	5,25
W	13,88	8,63	6,58	6,09	5,65
G	7,68	6,38	5,56	5,38	5,3
R	8,47	6,38	5,64	5,59	5,36
$\alpha = 1\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	2,98	1,87	1,34	1,23	1,07
S	1,07	1,25	1,04	1,02	0,98
W	6,31	2,86	1,67	1,48	1,28
G	1,43	1,35	1,09	1,04	1,02
R	2,52	1,86	1,27	1,21	1,12

Tabela 5 – Taxa de rejeição da hipótese nula (%) para a o cenário IV, submodelo de μ .

$\alpha = 10\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	18,18	12,73	11,49	10,83	10,81
S	12,08	10,93	10,59	10,2	10,41
W	23,18	15,06	12,91	11,95	11,55
G	14,9	11,96	11,12	10,47	10,7
R	15,24	12	11,18	10,8	10,61
$\alpha = 5\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	11,32	6,68	5,96	5,88	5,37
S	6,12	5,28	5,32	5,44	5,08
W	16,42	9,13	6,85	6,57	5,98
G	7,79	5,82	5,55	5,66	5,17
R	9,4	6,56	6,07	5,6	5,79
$\alpha = 1\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	4,54	1,6	1,41	1,29	1,21
S	0,95	0,87	1,06	1,12	1,15
W	8,34	2,71	2	1,77	1,42
G	1,37	1,05	1,18	1,12	1,12
R	3,19	1,57	1,43	1,35	1,3

Tabela 6 – Taxa de rejeição da hipótese nula (%) para a o cenário I, submodelo de ϕ com $r = 1$.

$\alpha = 10\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	16.99	12.57	10.76	11.17	10.15
S	8.04	9.15	9.55	9.87	9.44
W	95.65	93.74	92.85	92.52	92.71
G	16.05	12.34	10.68	11.1	10.12
R	13.89	11.21	10.91	11.12	10.78
$\alpha = 5\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	10.08	6.87	5.6	5.65	5.1
S	3.5	4.33	4.63	4.52	4.82
W	94.74	92.45	91.63	91.2	91.17
G	8.96	6.54	5.41	5.53	5.07
R	8.34	6.24	5.58	5.8	5.39
$\alpha = 1\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	2.96	1.72	1.36	1.06	1.21
S	0.47	0.65	0.92	0.72	1.03
W	92.82	90.13	88.89	88.56	88.39
G	2.22	1.43	1.3	1.04	1.2
R	2.58	1.66	1.17	1.1	1.05

Tabela 7 – Taxa de rejeição da hipótese nula (%) para a o cenário II, submodelo de ϕ com $r = 1$.

$\alpha = 10\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	24.22	13.07	11.04	10.82	10.5
S	8.08	9.11	9.39	9.46	9.78
W	85.6	78.17	75.62	74.39	74.19
G	22.67	12.75	10.88	10.74	10.41
RS	16.46	11.78	10.72	10.23	10.37
$\alpha = 5\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	17.97	7.31	5.91	5.69	5.24
S	3.76	4.45	4.75	4.83	4.74
W	82.8	74.12	71.1	69.35	68.98
G	16.29	7.16	5.84	5.55	5.21
RS	11.18	6.57	5.63	5.37	5.21
$\alpha = 1\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	11.38	2.1	1.22	1.26	0.99
S	0.57	0.85	0.8	0.92	0.85
W	77.49	66.51	62.42	60.88	60.5
G	9.68	1.99	1.17	1.2	0.96
RS	5.09	1.72	1.16	1.18	1.02

Tabela 8 – Taxa de rejeição da hipótese nula (%) para a o cenário III, submodelo de ϕ com $r = 1$.

$\alpha = 10\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	17.19	12.46	10.72	10.88	10.26
S	8.17	9.16	9.25	9.68	9.68
W	96.9	95.21	94.19	94.14	94.25
G	16.22	12.1	10.63	10.81	10.23
R	14.33	11.32	10.84	10.76	10.32
$\alpha = 5\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	10.03	6.86	5.41	5.55	5.17
S	3.91	4.57	4.41	4.72	4.73
W	96.29	94.34	92.92	93.17	93.05
G	8.85	6.53	5.35	5.48	5.12
R	8.37	6.09	5.36	5.62	5.25
$\alpha = 1\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	3.2	1.63	1.38	1.13	1.13
S	0.56	0.91	1.11	0.91	0.98
W	94.8	92.46	90.89	91.22	90.88
G	2.57	1.5	1.32	1.12	1.11
R	2.27	1.52	1.16	1.27	1.21

Tabela 9 – Taxa de rejeição da hipótese nula (%) para a o cenário IV, submodelo de ϕ com $r = 1$.

$\alpha = 10\%$					
Estatística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	-	12.86	11.27	11.15	10.49
S	-	8.36	9.25	9.59	9.56
W	-	83.29	81.27	81.03	80.03
G	-	12.74	11.06	11.06	10.43
R	-	11.79	10.98	10.94	10.68
$\alpha = 5\%$					
Estatística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	-	7.17	5.76	5.97	5.06
S	-	3.59	4.29	4.68	4.8
W	-	80.28	77.76	77.77	76.57
G	-	7.06	5.58	5.9	5.03
R	-	6.71	5.61	5.67	5.33
$\alpha = 1\%$					
Estatística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	-	1.96	1.44	1.17	1.09
S	-	0.66	0.91	0.86	0.87
W	-	74.48	71.44	70.75	69.69
G	-	1.8	1.27	1.07	1.1
R	-	1.51	1.24	1.18	1

Tabela 10 – Taxa de rejeição da hipótese nula (%) para a o cenário I, submodelo de ϕ com $r = 2$.

$\alpha = 10\%$					
Estatística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	18.51	13.84	11.45	11.23	10.08
S	7.6	9.1	9.47	9.75	9.46
W	68.97	70.23	70.85	71.19	70.02
G	17.27	13.22	11.15	11.03	9.96
$\alpha = 5\%$					
Estatística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	10.89	7.54	6.1	5.88	4.91
S	3.67	4.72	5.02	5.11	4.48
W	68.9	70.11	70.72	71.07	69.87
G	9.78	7.05	6.03	5.8	4.9
$\alpha = 1\%$					
Estatística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	3.53	1.84	1.49	1.37	1.02
S	0.79	0.99	1.27	0.96	0.92
W	68.75	69.95	70.3	70.75	69.52
G	2.9	1.56	1.38	1.26	0.96

Tabela 11 – Taxa de rejeição da hipótese nula (%) para a o cenário II, submodelo de ϕ com $r = 2$.

$\alpha = 10\%$					
Estatística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	-	14.15	12	11.37	10.71
S	-	9.41	9.98	9.63	9.93
W	-	67.06	66.82	66.22	66.25
G	-	13.91	11.79	11.18	10.6
$\alpha = 5\%$					
Estatística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	-	8.1	6.54	5.76	5.52
S	-	4.76	4.98	4.63	5.03
W	-	66.26	65.89	65.13	65.2
G	-	7.9	6.31	5.71	5.38
$\alpha = 1\%$					
Estatística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	-	2.17	1.48	1.3	1.07
S	-	0.93	1.1	1.02	1.01
W	-	64.57	63.94	63.09	62.75
G	-	2.21	1.38	1.25	1.01

Tabela 12 – Taxa de rejeição da hipótese nula (%) para a o cenário III, submodelo de ϕ com $r = 2$.

$\alpha = 10\%$					
Estatística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	18.99	13.35	11.31	10.89	11.02
S	7.3	8.82	9.52	9.51	10.09
W	68.99	69.81	69.76	69.71	69.99
G	17.38	12.71	11.11	10.75	10.86
$\alpha = 5\%$					
Estatística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	11.76	7.36	6.2	5.74	5.78
S	3.68	4.43	4.85	4.74	5.3
W	68.93	69.76	69.66	69.57	69.88
G	10.2	6.93	5.88	5.59	5.69
$\alpha = 1\%$					
Estatística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	3.81	1.93	1.35	1.24	1.15
S	0.6	0.91	0.9	1.01	1.08
W	68.78	69.59	69.43	69.35	69.66
G	2.94	1.64	1.25	1.15	1.11

Tabela 13 – Taxa de rejeição da hipótese nula (%) para a o cenário IV, submodelo de ϕ com $r = 2$.

$\alpha = 10\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	-	13.94	11.81	11.3	10.76
S	-	9.26	9.4	9.41	10.05
W	-	67.41	67.81	67.43	67.31
G	-	13.7	11.59	11	10.7
$\alpha = 5\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	-	7.86	5.95	5.93	5.77
S	-	4.62	4.64	4.65	5.23
W	-	67.01	67.15	66.79	66.5
G	-	7.53	5.85	5.72	5.71
$\alpha = 1\%$					
Estadística	n = 20	n = 40	n = 80	n = 100	n = 200
RV	-	2.03	1.34	1.29	1.24
S	-	1.05	0.97	1.01	1.08
W	-	65.87	65.86	65.51	65.17
G	-	1.96	1.28	1.28	1.22

6.2 Simulação de Poder

As simulações de poder foram realizadas para os testes bilaterais razão de verossimilhança, escore, Wald e gradiente e o teste unilateral da razão de verossimilhança sinalizada sob as hipóteses alternativas apresentadas em 5.1 e 5.2 para os submodelos da média com $r = 1$ e da precisão com $r = 1, 2$.

Para encontrar os resultados para as simulações de poder no submodelo da média quando $r = 1$, β_0 e β_1 são os valores iniciais dos parâmetros do modelo *BP* apresentados na Tabela 1 e β_2 têm os seguintes valores -1.0, -0.5, 0.5 e 1.0 para as estatísticas *RV*, *S*, *W* e *G*. Já no submodelo da precisão, λ_0 e λ_1 são os valores iniciais dos parâmetros do modelo *BP* apresentados na Tabela 1 e λ_2 têm valores -1.0, -0.5, 0.5 e 1.0 quando $r = 1$ e $r = 2$ para as estatísticas *RV*, *S* e *G*. Para a estatística *R*, quando o teste passa a ser unilateral, temos $\beta_2 = 0.5, 1.0$ e $\lambda_2 = 0.5, 1.0$.

Os valores críticos utilizados para realizar as simulações de poder foram obtidos nas simulações de tamanho com o objetivo de realizar comparações entre testes que possuem o mesmo tamanho. Dessa forma, é importante ressaltar que os poderes dos testes Wald não serão apresentados pois os mesmos apresentaram tamanhos muito distorcidos no que se refere ao submodelo de ϕ .

As Tabelas 14 a 17 apresentam os resultados dos testes avaliados sobre o vetor de parâmetros do submodelo da média com n igual a 20, 40 e 80, enquanto as Tabelas 18 a 21 apresentam os resultados dos testes avaliados sobre o vetor de parâmetros do submodelo

da precisão considerando $r = 1$ e as Tabelas 22 a 25 considerando $r = 2$ com n igual a 100 e 200.

Como era de se esperar, verificamos que os poderes aumentam conforme os tamanhos amostrais aumentam. Adicionalmente, observamos que os poderes são maiores quanto maior o erro que cometemos na geração dos dados. Além disso, podemos concluir que os testes apresentaram melhores resultados nos cenários em que foram considerados a precisão alta (cenários I e III). Dentre esses dois cenários, o cenário III, considerando valores altos para a média, foi o que apresentou melhores resultados. É importante citar também que, de forma geral, os poderes foram maiores para os testes realizados sobre os parâmetros da estrutura da média.

Para $r = 1$, observamos que o teste da razão de verossimilhanças sinalizada apresentou melhores desempenhos no que se refere ao poder seguido do teste gradiente. Por exemplo, na Tabela 18 em que os poderes para $\alpha = 10\%$ para $\lambda_2 = 1.0$ foram 74.52% e 92,82% para o teste R e 61.42% e 87.67% para o teste G com tamanhos de amostras grandes como $n = 100$ e $n = 200$, respectivamente. Isto também ocorre para todos os cenários. Esse desempenho superior do teste baseado na estatística R fica mais evidente nos cenários com precisão menor. Para $r = 2$, verificamos que os testes são igualmente poderosos, com poderes levemente superiores para os testes baseados nas estatísticas RV e G .

Dessa forma, baseado nos resultados encontrados nas simulações de tamanho e poder, indicamos o uso do teste da razão de verossimilhanças sinalizada para testar parâmetros na estrutura da média ou da precisão, quando $r = 1$. Na situação em que $r = 2$, quando se deseja testar se a precisão é constante, por exemplo, recomendamos o uso do teste gradiente.

Tabela 14 – Poder dos testes (%), submodelo de μ para a o cenário I.

Teste Bilateral										
β_2	Estatística	n=20			n=40			n=80		
		10%	5%	1%	10%	5%	1%	10%	5%	1%
-1.0	RV	99.94	99.83	98.59	100	100	100	100	100	100
	S	99.14	98.41	95.79	100	100	99.96	100	100	100
	W	99.98	99.93	97.33	100	100	100	100	100	100
	G	99.9	99.81	98.89	100	100	100	100	100	100
-0.5	RV	89.35	80.37	52.77	99.65	99.36	95.96	100	100	99.99
	S	86.83	78.07	55.41	99.61	98.99	94.94	100	99.99	99.99
	W	89.16	78.95	39.84	99.72	99.41	96.57	100	100	99.99
	G	89.43	80.93	56.12	99.67	99.4	96.15	100	100	99.99
0.5	RV	94.1	87.69	62.96	99.96	99.82	98.56	100	100	100
	S	92.4	85.61	64.32	99.92	99.79	98.04	100	100	100
	W	94.08	85.84	49.08	99.97	99.85	98.68	100	100	100
	G	94.32	88.4	66.91	99.96	99.82	98.61	100	100	100
1.0	RV	99.99	99.97	99.76	100	100	100	100	100	100
	S	99.69	99.45	98.58	100	100	100	100	100	100
	W	100	100	99.8	100	100	100	100	100	100
	G	99.98	99.96	99.85	100	100	100	100	100	100
Teste Unilateral										
β_2	Estatística	n=20			n=40			n=80		
		10%	5%	1%	10%	5%	1%	10%	5%	1%
0.5	R	97.63	94.14	73.56	100	99.96	99.36	100	100	100
1.0		100	99.99	99.89	100	100	100	100	100	100

Tabela 15 – Poder dos testes (%), submodelo de μ para a o cenário II.

Teste Bilateral										
β_2	Estatística	n=20			n=40			n=80		
		10%	5%	1%	10%	5%	1%	10%	5%	1%
-1.0	RV	52.91	38.5	8.24	85.6	76.64	53.96	98.97	97.5	90.49
	S	54.62	41.58	21.34	84.94	75.83	53.07	98.83	97.53	90.06
	W	51.43	34.48	1.44	85.8	76.32	52.37	98.96	97.47	90.7
	G	54.89	43.18	21.08	85.81	76.83	54.8	98.99	97.58	90.88
-0.5	RV	23.36	14.34	2.46	41.81	29.81	12.18	67.78	56.3	30.84
	S	24.84	15.95	5.76	42.24	30.34	12.77	67.7	57.2	31.92
	W	23.14	12.71	1.34	41.84	28.91	11.56	67.49	55.43	30.74
	G	24.42	15.78	4.85	42.18	29.93	12.57	67.93	56.73	31.61
0.5	RV	26.64	16.6	2.41	46.8	34.59	14.89	74.51	63.31	37.22
	S	27.58	17.12	5.85	46.16	33.11	13.8	73.6	63.1	35.77
	W	25.99	14.79	1.16	46.88	33.94	14.51	74.58	62.69	37.61
	G	27.83	18.31	6.12	46.57	34.34	14.62	74.31	63.17	37.36
1.0	RV	67.01	51.87	13.59	94.94	90.13	73.49	99.95	99.77	98.31
	S	66.9	53.39	28.32	94.13	88.66	70.03	99.94	99.75	97.88
	W	65.74	47.58	2.01	95.1	90.03	73	99.93	99.79	98.41
	G	68.36	56.48	30.09	94.83	89.93	73.08	99.95	99.77	98.3
Teste Unilateral										
β_2	Estatística	n=20			n=40			n=80		
		10%	5%	1%	10%	5%	1%	10%	5%	1%
0.5	R	40.54	25.75	6.21	63.2	46.41	20.99	83.86	73.54	48.6
1.0		80.16	66.5	28.79	97.88	94.84	80.6	99.99	99.95	99.2

Tabela 17 – Poder dos testes (%), submodelo de μ para a o cenário IV.

Teste Bilateral										
		n=20			n=40			n=80		
β_2	Estatística	10%	5%	1%	10%	5%	1%	10%	5%	1%
-1.0	RV	60.21	46.83	2.57	96.59	93.43	81.02	99.98	99.9	99.21
	S	60.07	49.9	29.69	95.86	92.34	79.51	99.98	99.85	99.05
	W	59.52	43.32	1.3	96.86	94.05	80.28	99.97	99.91	99.37
	G	61.84	52.07	28.16	96.55	93.59	80.74	99.98	99.9	99.23
-0.5	RV	29.88	17.66	1.22	56.43	43.62	22.06	84.24	74.21	49.36
	S	32.45	21.11	7.84	56.51	43.62	22.73	84.16	74.35	50.59
	W	29.01	15.75	1.09	56.15	43.89	20.21	84.15	74.42	50.53
	G	31.65	20.85	6.38	56.5	43.78	21.65	84.28	74.11	49.41
0.5	RV	32.28	19.5	1.13	58.88	47.04	24.22	86.12	77.49	52.51
	S	33.6	22.37	7.46	58.72	45.85	23.6	85.65	76.56	52.29
	W	30.82	16.18	1.16	58.43	46.58	22	86.01	77.18	53.04
	G	34.59	23.86	8.02	58.95	47.32	24.23	86.03	77.38	52.88
1.0	RV	76.48	60.92	2.63	97.97	95.82	86.4	100	100	99.72
	S	76.14	63.67	35.27	97.43	94.9	83.9	100	100	99.68
	W	74.83	54.29	1.31	98.09	96.04	84.88	100	99.99	99.73
	G	78.46	66.9	37.81	97.96	95.85	86.06	100	100	99.72
Teste Unilateral										
		n=20			n=40			n=80		
β_2	Estatística	10%	5%	1%	10%	5%	1%	10%	5%	1%
0.5	R	4.8	1	0	72.85	57.81	31	92.44	85.66	62.93
1.0		25.43	3.57	0.02	99.31	97.82	90.48	100	100	99.92

Tabela 18 – Poder dos testes (%), submodelo de ϕ com $r = 1$ para a o cenário I.

Teste Bilateral							
		n=100			n=200		
λ_2	Estatística	10%	5%	1%	10%	5%	1%
-1.0	RV	60.01	48.22	26.56	87.25	79.21	56.14
	S	59.07	47.72	26.86	86.94	78.21	55.18
	G	60.1	48.28	26.48	87.25	79.17	56.3
-0.5	RV	25.3	16.55	6.23	41.57	29.82	11.73
	S	25.41	17.06	6.64	42.17	29.48	11.91
	G	25.33	16.63	6.16	41.62	29.8	11.8
0.5	RV	24.55	15.96	6.05	41.01	28.62	11.28
	S	24.24	16.06	6.07	41.05	28.08	11.33
	G	24.5	15.93	5.9	41.01	28.47	11.26
1.0	RV	61.42	49.8	27.51	87.68	79.7	57.71
	S	60.92	48.96	26.95	87.49	78.7	56.67
	G	61.42	49.69	27.26	87.67	79.64	57.69
Teste Unilateral							
		n=100			n=200		
λ_2	Estatística	10%	5%	1%	10%	5%	1%
0.5	R	36.57	23.76	9.29	53.14	39.85	17.59
1.0		73.82	60.94	35.7	93.28	87.08	68.54

Tabela 19 – Poder dos testes (%), submodelo de ϕ com $r = 1$ para a o cenário II.

Teste Bilateral							
		n=100			n=200		
λ_2	Estatística	10%	5%	1%	10%	5%	1%
-1.0	RV	45.31	33	13.88	70.03	58.58	35.77
	S	44.73	32.08	14.1	69.31	58.13	35.46
	G	45.43	33.09	13.74	70.07	58.54	35.62
-0.5	RV	21.06	12.31	3.75	31.27	20.76	8.35
	S	21.31	12.59	3.87	31.07	21.16	8.66
	G	21.1	12.41	3.76	31.26	20.76	8.24
0.5	RV	21.48	12.93	3.92	34.37	23.79	9.74
	S	21.5	12.95	3.87	34.17	23.58	9.83
	G	21.45	13.04	3.92	34.28	23.71	9.62
1.0	RV	54.55	41.43	20.46	81.09	71.87	49.22
	S	54.36	40.85	20.11	80.73	71.41	48.73
	G	54.53	41.56	20.4	81.05	71.73	48.87
Teste Unilateral							
		n=100			n=200		
λ_2	Estatística	10%	5%	1%	10%	5%	1%
0.5	R	33.9	21.16	6.87	48.76	34.32	14.32
1.0		69.18	54.7	28.39	89.24	81.15	58.86

Tabela 20 – Poder dos testes (%), submodelo de ϕ para a o cenário III.

Teste Bilateral							
		n=100			n=200		
λ_2	Estatística	10%	5%	1%	10%	5%	1%
-1.0	RV	61.35	49.1	26.69	87.35	80.05	57.77
	S	60.47	48.61	26.63	86.98	78.93	57.9
	G	61.46	49.24	26.98	87.37	80.03	58.11
-0.5	RV	25.2	16.15	5.13	40.4	29.05	11.2
	S	24.89	16.5	5.24	40.57	28.79	11.8
	G	25.3	16.25	5.22	40.38	29.02	11.31
0.5	RV	24.92	16.05	5.64	40.07	28.8	11.54
	S	24.61	15.62	5.66	40.21	28.47	11.71
	G	24.89	16.01	5.71	40.01	28.69	11.61
1.0	RV	62.47	48.97	26.47	86.92	79.55	58.31
	S	61.25	48.37	25.62	86.66	78.43	57.87
	G	62.39	48.86	26.55	86.9	79.4	58.38
Teste Unilateral							
		n=100			n=200		
λ_2	Estatística	10%	5%	1%	10%	5%	1%
0.5	R	36.58	24.31	8.44	54.33	39.14	16.38
1.0		74.52	62.15	34.05	92.82	86.52	66.29

Tabela 21 – Poder dos testes (%), submodelo de ϕ com $r = 1$ para a o cenário IV.

Teste Bilateral							
		n=100			n=200		
λ_2	Estatística	10%	5%	1%	10%	5%	1%
-1.0	RV	-	-	-	69.05	57.83	33.04
	S	-	-	-	68.68	56.51	33.86
	G	-	-	-	69.04	57.61	32.99
-0.5	RV	20.62	12.57	3.83	30.97	21.43	7.29
	S	20.83	13.11	3.87	31.31	21.03	8.19
	G	20.61	12.57	3.8	30.97	21.33	7.36
0.5	RV	21.56	13.11	4.39	34.37	24.32	8.7
	S	21.51	13.46	4.3	34.52	23.61	9.33
	G	21.53	13.1	4.35	34.34	24.21	8.7
1.0	RV	53.58	40.87	21.16	81.21	72.19	47.57
	S	53.05	40.95	20.34	80.59	70.93	48.41
	G	53.59	40.76	20.92	81.2	71.88	47.57
Teste Unilateral							
		n=100			n=200		
λ_2	Estatística	10%	5%	1%	10%	5%	1%
0.5	R	32.92	20.83	6.86	47.13	34.01	14.33
1.0		67.25	53.28	28.19	88.96	81.01	59.43

Tabela 22 – Poder dos testes (%), submodelo de ϕ com $r = 2$ para a o cenário I.

Teste Bilateral							
		n=100			n=200		
$\lambda_{1,2}$	Estatística	10%	5%	1%	10%	5%	1%
-1.0	RV	59.82	47.25	24.27	88.02	81.01	60.01
	S	58.05	45.09	23.79	86.65	79.38	57.27
	G	59.87	47.16	24.55	87.99	80.96	60.05
-0.5	RV	23.48	14.7	4.5	39.89	28.57	11.31
	S	23.3	14.16	4.79	39.04	28.03	10.9
	G	23.46	14.55	4.52	39.83	28.39	11.39
0.5	RV	24.39	15.11	4.99	40.65	29.28	12.47
	S	24.34	15.08	5.48	40.13	29.94	12.67
	G	24.45	15.12	5.14	40.75	29.31	12.72
1.0	RV	63.12	51.33	28.04	89.91	83.26	64.25
	S	62.72	49.94	29.24	89.25	82.84	63.62
	G	63.2	51.35	28.65	89.97	83.3	64.81

Tabela 23 – Poder dos testes (%), submodelo de ϕ com $r = 2$ para a o cenário II.

		Teste Bilateral					
		n=100			n=200		
$\lambda_{1,2}$	Estatística	10%	5%	1%	10%	5%	1%
-1.0	RV	-	-	-	65.68	53.07	30.51
	S	-	-	-	63.71	51.06	27.1
	G	-	-	-	65.06	52.51	29.84
-0.5	RV	18.86	11.13	2.78	29.57	18.97	6.73
	S	18.96	11.13	3.19	29.3	18.45	6.07
	G	18.55	11.11	2.75	29.24	19.05	6.64
0.5	RV	22.33	13.83	4.29	35.16	24.03	9.49
	S	22.41	14.41	4.75	35.66	23.96	9.66
	G	21.97	14.03	4.42	34.99	24.17	9.63
1.0	RV	57.27	45.53	22.94	85.53	77.29	57.67
	S	57.37	45.15	23.94	84.75	76.85	56.17
	G	57.23	45.83	23.29	85.45	77.55	58.02

Tabela 24 – Poder dos testes (%), submodelo de ϕ com $r = 2$ para a o cenário III.

		Teste Bilateral					
		n=100			n=200		
$\lambda_{1,2}$	Estatística	10%	5%	1%	10%	5%	1%
-1.0	RV	60.9	48.46	25.31	87.38	79.49	59.03
	S	59.5	46.87	23.41	86.22	77.47	55.59
	G	60.81	48.54	24.86	87.33	79.29	58.51
-0.5	RV	23.77	14.84	4.53	38.48	26.48	10.06
	S	23.74	14.8	4.34	38.29	25.82	9.59
	G	23.72	14.97	4.43	38.47	26.25	9.78
0.5	RV	24.88	15.61	5.02	38.84	27.23	11.25
	S	24.59	15.73	5.24	39.1	27.16	11.73
	G	24.93	15.71	4.98	38.9	27.16	11.24
1.0	RV	63.28	51.29	27.51	88.57	81.5	62.2
	S	62.75	50.89	28.21	88.13	80.97	61.82
	G	63.35	51.65	27.46	88.62	81.41	62.35

Tabela 25 – Poder dos testes (%), submodelo de ϕ com $r = 2$ para a o cenário IV.

		Teste Bilateral					
		n=100			n=200		
$\lambda_{1,2}$	Estatística	10%	5%	1%	10%	5%	1%
-1.0	RV	-	-	-	-	-	-
	S	-	-	-	-	-	-
	G	-	-	-	-	-	-
-0.5	RV	19.63	11.58	3.65	28.45	18.35	5.78
	S	20.11	12.02	3.41	28.12	18.02	5.95
	G	19.54	11.51	3.47	28.38	18.31	5.57
0.5	RV	22.27	12.84	4	34.65	23.19	8.23
	S	22.34	13.96	4.09	34.69	23.44	8.8
	G	22.24	12.89	3.86	34.73	23.24	8.11
1.0	RV	57.7	44.26	22.77	84.57	75.89	54.35
	S	57.46	45.17	22.66	83.9	75.15	54.62
	G	57.91	44.23	22.39	84.65	76.18	54.23

7 Considerações Finais

7.1 Conclusão

É comum em inferência que após a estimação dos parâmetros do modelo de regressão sejam feitos testes de hipóteses sobre eles. Neste estudo, o interesse é avaliar os desempenhos dos testes da razão de verossimilhanças, escore e Wald usuais comparativamente aos testes da razão de verossimilhanças sinalizada e gradiente sobre os parâmetros do modelo de regressão beta prime. Os resultados apresentados foram obtidos por meio de simulação de Monte Carlo em amostras finitas. Considerou-se o modelo de regressão beta prime com dispersão variável abordando testes de hipóteses sobre os vetores de parâmetros que indexam os submodelos da média e da dispersão.

De forma geral, através das simulações, os resultados apresentados sugerem que, a medida que o tamanho da amostra aumenta, o tamanho empírico dos testes abordados ficam muito próximos dos níveis de significância selecionados. Verifica-se que os testes baseados nas estatísticas escore, gradiente e razão de verossimilhanças sinalizada, apresentaram taxas de rejeição empíricas mais próximas dos níveis nominais do que os testes baseados nas estatísticas razão de verossimilhanças convencional e Wald.

Além disso, observamos que o teste da razão de verossimilhanças sinalizada apresentou melhores desempenhos no que se refere ao poder para $r = 1$, enquanto que para $r = 2$ os testes baseados nas estatísticas RV e G são levemente mais poderosos que os demais.

Portanto, baseado nos resultados encontrados nas simulações de tamanho e poder, indicamos o uso do teste da razão de verossimilhanças sinalizada para testar parâmetros na estrutura da média ou da precisão, quando $r = 1$. Na situação em que $r = 2$, quando se deseja testar se a precisão é constante, por exemplo, recomendamos o uso do teste gradiente.

Referências

- ABERNATHY, R. W.; SMITH, R. P. Applying series expansion to the inverse beta distribution to find percentiles of the f-distribution. *ACM Transactions on Mathematical Software (TOMS)*, ACM New York, NY, USA, v. 19, n. 4, p. 474–480, 1993. Citado na página 16.
- BOURGUIGNON, M.; SANTOS-NETO, M.; CASTRO, M. de. A new regression model for positive data. *arXiv preprint arXiv:1804.07734*, 2018. Citado 8 vezes nas páginas 7, 8, 14, 18, 21, 22, 30 e 51.
- CORDEIRO, G. M. *Introdução a teoria assintótica*. [S.l.]: IMPA, 1999. Citado na página 23.
- HAYAKAWA, T.; PURI, M. L. Asymptotic expansions of the distributions of some test statistics. *Ann. Inst. Statist. Math*, v. 37, n. Part A, p. 95–108, 1985. Citado na página 27.
- KEEPING, E. *Introduction to statistical inference*, vol. 26. DoverPublications. com, 1962. Citado na página 16.
- LEÃO, J. et al. The negative binomial beta prime regression model with cure rate: application with a melanoma dataset. *Journal of Statistical Theory and Practice*, Springer, v. 15, n. 3, p. 1–21, 2021. Citado na página 16.
- MCDONALD, J. B. Some generalized functions for the size distribution of income. *Econometrica*, v. 52, p. 647–663, 1984. Citado na página 16.
- MCDONALD, J. B. Model selection: some generalized distributions. *Communications in Statistics-Theory and Methods*, Taylor & Francis, v. 16, n. 4, p. 1049–1074, 1987. Citado na página 16.
- MCDONALD, J. B.; BUTLER, R. J. Regression models for positive random variables. *Journal of Econometrics*, Elsevier, v. 43, n. 1-2, p. 227–251, 1990. Citado 2 vezes nas páginas 16 e 21.
- MCDONALD, J. B.; XU, Y. J. A generalization of the beta distribution with applications. *Journal of Econometrics*, Elsevier, v. 66, n. 1-2, p. 133–152, 1995. Citado na página 16.
- NELDER, J. A.; WEDDERBURN, R. W. Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)*, Wiley Online Library, v. 135, n. 3, p. 370–384, 1972. Citado na página 14.
- NEYMAN, J.; PEARSON, E. S. On the use and interpretation of certain test criteria for purposes of statistical inference: Part i. *Biometrika*, JSTOR, p. 175–240, 1928. Citado na página 23.
- OLIVEIRA, D. d. S. et al. Teste da razão de verossimilhança via bootstrap em modelos não encaixados. 2016. Citado na página 23.

QUEIROZ, M. Pinto Ferreira de. Testes de hipóteses em regressão beta baseados em verossimilhança perfilada ajustada e em bootstrap. Dissertação (Mestrado) — Universidade Federal de Pernambuco, 2011. Citado na página 14.

R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria, 2021. Disponível em: <<https://www.R-project.org/>>. Citado na página 30.

RAO, C. In advances in ranking and selection, multiple comparisons and reliability. [S.l.]: Birkhauser, 2005. Citado na página 27.

RAO, C. R. Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. In: CAMBRIDGE UNIVERSITY PRESS. Mathematical Proceedings of the Cambridge Philosophical Society. [S.l.], 1948. v. 44, n. 1, p. 50–57. Citado 2 vezes nas páginas 23 e 25.

SEVERINI, T. A. Likelihood methods in statistics. [S.l.]: Oxford University Press, 2000. Citado 2 vezes nas páginas 23 e 25.

TERRELL, G. R. The gradient statistic. Computing Science and Statistics, v. 34, n. 34, p. 206–215, 2002. Citado 2 vezes nas páginas 23 e 27.

TULUPYEV, A. et al. Beta prime regression with application to risky behavior frequency screening. Statistics in medicine, Wiley Online Library, v. 32, n. 23, p. 4044–4056, 2013. Citado 2 vezes nas páginas 16 e 21.

WALD, A. Tests of statistical hypotheses concerning several parameters when the number of observations is large. Transactions of the American Mathematical society, JSTOR, v. 54, n. 3, p. 426–482, 1943. Citado 2 vezes nas páginas 23 e 26.

A Apêndice

Neste apêndice serão apresentados os detalhes necessários para obtenção das quantidades: a função escore U e matriz de informação de Fisher K . Estas quantidades foram obtidas inicialmente no artigo de BOURGUIGNON; SANTOS-NETO; CASTRO (2018).

A.1 Vetor Escore

Primeiramente mostraremos como foram obtidos os elementos do vetor escore para essa classe de modelos. A função escore é obtida pela diferenciação da função de log-verossimilhança (3.2) com respeito ao β_j e v_k e são dados por

$$\begin{aligned}
 U_j &= \frac{\partial \ell(\beta, \mathbf{v})}{\partial \beta_j} = \sum_{t=1}^n \frac{\partial \ell(\mu_t, \phi_t)}{\partial \mu_t} \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_j} \\
 &= \sum_{t=1}^n [(1 + \phi_t) \log y_t - (1 + \phi_t) \log(1 + y_t) - (1 + \phi_t) \psi(\mu_t(1 + \phi_t)) \\
 &\quad + (1 + \phi_t) \psi(\mu_t(1 + \phi_t) + \phi_t + 2)] \frac{d\mu_t}{d\eta_{1t}} x_{tj} \\
 &= \sum_{t=1}^n (1 + \phi_t) \left[\log\left(\frac{y_t}{1 + y_t}\right) - \psi(\mu_t(1 + \phi_t)) + \psi(\mu_t(1 + \phi_t) + \phi_t + 2) \right] \frac{d\mu_t}{d\eta_{1t}} x_{tj} \\
 &= \sum_{t=1}^n (1 + \phi_t) (y_t^* - \mu_t^*) \frac{d\mu_t}{d\eta_{1t}} x_{tj},
 \end{aligned}$$

e

$$\begin{aligned}
U_k &= \frac{\partial \ell(\boldsymbol{\beta}, \mathbf{v})}{\partial v_k} = \sum_{t=1}^n \frac{\partial \ell(\mu_t, \phi_t)}{\partial \phi_t} \frac{d\phi_t}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial v_k} \\
&= \sum_{t=1}^n [\mu_t \log y_t - (\mu_t + 1) \log(1 + y_t) - \mu_t \psi(\mu_t(1 + \phi_t)) - \psi(\phi_t + 2) \\
&\quad + (\mu_t + 1) \psi(\mu_t(1 + \phi_t) + \phi_t + 2)] \frac{d\phi_t}{d\eta_{2t}} z_{tk} \\
&= \sum_{t=1}^n [\log\left(\frac{y_t^{\mu_t}}{(1 + y_t)^{(\mu_t + 1)}}\right) - \mu_t(\psi(\mu_t(1 + \phi_t)) - \psi(\mu_t(1 + \phi_t) + \phi_t + 2)) - \psi(\phi_t + 2) \\
&\quad + \psi(\mu_t(1 + \phi_t) + \phi_t + 2)] \frac{d\phi_t}{d\eta_{2t}} z_{tk} \\
&= \sum_{t=1}^n [y_t^* - \mu_t(\mu_t^* - \frac{\gamma}{\mu_t})] \frac{d\phi_t}{d\eta_{2t}} z_{tk} \\
&= \sum_{t=1}^n (y_t^* - \mu_t^*) \frac{d\phi_t}{d\eta_{2t}} z_{tk},
\end{aligned}$$

em que, $\psi(\cdot) = \frac{d \log \Gamma(\cdot)}{d}$ é a função digamma, $y_t^* = \log\left(\frac{y_t}{(1 + y_t)}\right)$, $\mu_t^* = \psi(\mu_t(1 + \phi_t)) - \psi(\mu_t(1 + \phi_t) + \phi_t + 2)$, $y_t^* = \log\left(\frac{y_t^{\mu_t}}{(1 + y_t)^{(\mu_t + 1)}}\right)$ e $\mu_t^* = \mu_t(\mu_t^* - \frac{\gamma}{\mu_t})$ onde $\gamma = \psi(\phi_t + 2) - \psi(\mu_t(1 + \phi_t) + \phi_t + 2)$.

A função escore pode também ser escrita na forma matricial

$$\begin{aligned}
U_{\boldsymbol{\beta}} &= \mathbf{X}^{\top} \boldsymbol{\Phi} \mathbf{D}_1 (\mathbf{y}^* - \boldsymbol{\mu}^*), \\
U_{\mathbf{v}} &= \mathbf{Z}^{\top} \mathbf{D}_2 (\mathbf{y}^* - \boldsymbol{\mu}^*),
\end{aligned}$$

em que $\mathbf{D}_1 = \text{diag}\left[\frac{\partial \mu_1}{\partial \eta_{11}}, \dots, \frac{\partial \mu_n}{\partial \eta_{1n}}\right]$, $\mathbf{D}_2 = \text{diag}\left[\frac{\partial \phi_1}{\partial \eta_{21}}, \dots, \frac{\partial \phi_n}{\partial \eta_{2n}}\right]$, $\boldsymbol{\Phi} = \text{diag}[(1 + \phi_1), \dots, (1 + \phi_n)]$, $\mathbf{y}^* = (y_1^*, \dots, y_n^*)^{\top}$, $\mathbf{y}^* = (y_1^*, \dots, y_n^*)^{\top}$, $\boldsymbol{\mu}^* = (\mu_1^*, \dots, \mu_n^*)^{\top}$ e $\boldsymbol{\mu}^* = (\mu_1^*, \dots, \mu_n^*)^{\top}$.

A.2 Matriz de Informação de Fisher

A matriz de informação de Fisher é obtida a partir da segunda derivada da função log-verossimilhança (3.2), que são expressos da forma a seguir:

1. para β_j e β_l , com $j, l = 1, \dots, m$, temos

$$\begin{aligned}
U_{jl} &= \frac{\partial^2 \ell(\boldsymbol{\beta}, \mathbf{v})}{\partial \beta_j \partial \beta_l} = \sum_{t=1}^n \left\{ \frac{\partial^2 \ell(\mu_t, \phi_t)}{\partial \mu_t^2} \left(\frac{d\mu_t}{d\eta_{1t}}\right)^2 + \frac{\partial \ell(\mu_t, \phi_t)}{\partial \mu_t} \left(\frac{\partial}{\partial \mu_t} \frac{d\mu_t}{d\eta_{1t}}\right) \frac{d\mu_t}{d\eta_{1t}} \right\}_{x_{tj} x_{tl}} \\
&= \sum_{t=1}^n \left\{ -(1 + \phi_t)^2 [\psi^{(1)}(\mu_t(1 + \phi_t)) - \psi^{(1)}(\mu_t(1 + \phi_t) + \phi_t + 2)] \left(\frac{d\mu_t}{d\eta_{1t}}\right)^2 \right. \\
&\quad \left. + \frac{\partial \ell(\mu_t, \phi_t)}{\partial \mu_t} \left(\frac{\partial}{\partial \mu_t} \frac{d\mu_t}{d\eta_{1t}}\right) \frac{d\mu_t}{d\eta_{1t}} \right\}_{x_{tk} x_{tl}}.
\end{aligned}$$

Sob condições de regularidade, obtemos $E\left(\frac{\partial \ell(\mu_t, \phi_t)}{\partial \mu_t}\right) = 0$. Assim, temos que

$$\begin{aligned} K_{\beta\beta} &= -E(U_{jl}) \\ &= \sum_{t=1}^n (1 + \phi_t)^2 [\psi^{(1)}(\mu_t(1 + \phi_t)) - \psi^{(1)}(\mu_t(1 + \phi_t) + \phi_t + 2)] \left(\frac{d\mu_t}{d\eta_{1t}}\right)^2 x_{tj} x_{tl}. \end{aligned}$$

2. Para β_j e v_k , com $j = 1, \dots, m$ e $k = 1, \dots, M$, temos

$$\begin{aligned} U_{jk} &= \frac{\partial^2 \ell(\beta, v)}{\partial \beta_j \partial v_k} = \sum_{t=1}^n \frac{\partial^2 \ell(\mu_t, \phi_t)}{\partial \mu_t \partial \phi_t} \frac{d\mu_t}{d\eta_{1t}} \frac{d\phi_t}{d\eta_{2t}} x_{tj} z_{tk} \\ &= \sum_{t=1}^n \{ (y_t^* - \mu_t^*) + (1 + \phi_t) [\psi^{(1)}(\mu_t(1 + \phi_t) + \phi_t + 2) + \mu_t [\psi^{(1)}(\mu_t(1 + \phi_t) + \phi_t + 2) \\ &\quad - \psi^{(1)}(\mu_t(1 + \phi_t))]] \} \frac{d\mu_t}{d\eta_{1t}} \frac{d\phi_t}{d\eta_{2t}} x_{tj} z_{tk}. \end{aligned}$$

Tomando o valor esperado sobre a função, temos que

$$\begin{aligned} K_{\beta v} &= -E(U_{jk}) \\ &= - \sum_{t=1}^n (1 + \phi_t) [\psi^{(1)}(\mu_t(1 + \phi_t) + \phi_t + 2) + \mu_t [\psi^{(1)}(\mu_t(1 + \phi_t) + \phi_t + 2) - \psi^{(1)}(\mu_t(1 + \phi_t))]] \\ &\quad \frac{d\mu_t}{d\eta_{1t}} \frac{d\phi_t}{d\eta_{2t}} x_{tj} z_{tk}. \end{aligned}$$

3. Para v_k e v_p , com $k, p = 1, \dots, m$, temos

$$\begin{aligned} U_{kp} &= \frac{\partial^2 \ell(\beta, v)}{\partial v_k \partial v_p} = \sum_{t=1}^n \left\{ \frac{\partial^2 \ell(\mu_t, \phi_t)}{\partial \phi_t^2} \left(\frac{d\phi_t}{d\eta_{2t}}\right)^2 + \frac{\partial \ell(\mu_t, \phi_t)}{\partial \phi_t} \left(\frac{\partial}{\partial \phi_t} \frac{d\phi_t}{d\eta_{2t}}\right) \frac{d\phi_t}{d\eta_{2t}} \right\} z_{tk} z_{tp} \\ &= \sum_{t=1}^n \left\{ -[\mu_t^2 \psi^{(1)}(\mu_t(1 + \phi_t)) - (\mu_t + 1)^2 \psi^{(1)}(\mu_t(1 + \phi_t) + \phi_t + 2) + \psi^{(1)}(\phi_t + 2)] \left(\frac{d\phi_t}{d\eta_{2t}}\right)^2 \right. \\ &\quad \left. + \frac{\partial \ell(\mu_t, \phi_t)}{\partial \phi_t} \left(\frac{\partial}{\partial \phi_t} \frac{d\phi_t}{d\eta_{2t}}\right) \frac{d\phi_t}{d\eta_{2t}} \right\} z_{tk} z_{tp}. \end{aligned}$$

Sob condições de regularidade, obtemos $E\left(\frac{\partial \ell(\mu_t, \phi_t)}{\partial \phi_t}\right) = 0$. Assim, temos que

$$\begin{aligned} K_{vv} &= -E(U_{kp}) \\ &= \sum_{t=1}^n (\mu_t^2 \psi^{(1)}(\mu_t(1 + \phi_t)) - (\mu_t + 1)^2 \psi^{(1)}(\mu_t(1 + \phi_t) + \phi_t + 2) + \psi^{(1)}(\phi_t + 2)) \\ &\quad \left(\frac{d\phi_t}{d\eta_{2t}}\right)^2 z_{tk} z_{tp}. \end{aligned}$$

A matriz de informação de Fisher pode ser escrita na forma matricial

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_{\beta\beta} & \mathbf{K}_{\beta\nu} \\ \mathbf{K}_{\nu\beta} & \mathbf{K}_{\nu\nu} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^\top \mathbf{E}_3 \mathbf{X} & \mathbf{X}^\top \mathbf{E}_5 \mathbf{Z} \\ \mathbf{Z}^\top \mathbf{E}_5 \mathbf{X} & \mathbf{Z}^\top \mathbf{E}_4 \mathbf{Z} \end{bmatrix},$$

em que cada elemento de E_3 , E_4 e E_5 são, respectivamente, dado por

$$e_{3i} = (1 + \phi_i)^2 [\Psi^{(1)}(\mu_i(1 + \phi_i)) - \Psi^{(1)}(\mu_i(1 + \phi_i) + \phi_i + 2)] \left(\frac{\partial \mu_i}{\partial \eta_{1i}} \right)^2,$$

$$e_{4i} = [\mu_i^2 \Psi^{(1)}(\mu_i(1 + \phi_i)) - (1 + \mu_i)^2 \Psi^{(1)}(\mu_i(1 + \phi_i) + \phi_i + 2) + \Psi^{(1)}(\phi_i + 2)] \left(\frac{\partial \phi_i}{\partial \eta_{2i}} \right)^2,$$

e

$$e_{5i} = -(1 + \phi_i) \{ \Psi^{(1)}(\mu_i(1 + \phi_i) + \phi_i + 2) + \mu_i [\Psi^{(1)}(\mu_i(1 + \phi_i) + \phi_i + 2) - \Psi^{(1)}(\mu_i(1 + \phi_i))] \} \left(\frac{\partial \mu_i}{\partial \eta_{1i}} \right) \left(\frac{\partial \phi_i}{\partial \eta_{2i}} \right).$$
