



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE ESTATÍSTICA E MATEMÁTICA APLICADA
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM E
MÉTODOS QUANTITATIVOS

CARLOS HENRIQUE BARROSO SENA SOUSA

MODELO DE REGRESSÃO L-LOGÍSTICA COM PERSPECTIVA FREQUENTISTA

FORTALEZA

2018

CARLOS HENRIQUE BARROSO SENA SOUSA

MODELO DE REGRESSÃO L-LOGÍSTICA COM PERSPECTIVA FREQUENTISTA

Dissertação apresentada ao Programa de Pós-graduação em Modelagem e Métodos Quantitativos do Departamento de Estatística e Matemática Aplicada da Universidade Federal do Ceará, como parte dos requisitos necessários para a obtenção do título de Mestre em Modelagem e Métodos Quantitativos. Área de concentração: Interdisciplinar.

Orientador: Prof. Dr. Rafael Bráz Azevedo Farias
Coorientador: Prof. Dr. Juvêncio Santos Nobre.

FORTALEZA

2018

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

S696m Sousa, Carlos Henrique Barroso Sena.

Modelo de regressão L-logística com perspectiva frequentista / Carlos Henrique Barroso Sena Sousa. – 2018.

152 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Modelagem e Métodos Quantitativos, Fortaleza, 2018.

Orientação: Prof. Dr. Rafael Bráz Azevedo Farias.

Coorientação: Prof. Dr. Juvêncio Santos Nobre.

1. Modelo de Regressão L-logística. 2. Máxima Verossimilhança. 3. Metaheurísticas. 4. Reamostragem.
I. Título.

CDD 510

CARLOS HENRIQUE BARROSO SENA SOUSA

MODELO DE REGRESSÃO L-LOGÍSTICA COM PERSPECTIVA FREQUENTISTA

Dissertação apresentada ao Programa de Pós-graduação em Modelagem e Métodos Quantitativos do Departamento de Estatística e Matemática Aplicada da Universidade Federal do Ceará, como parte dos requisitos necessários para a obtenção do título de Mestre em Modelagem e Métodos Quantitativos. Área de concentração: Interdisciplinar.

Aprovada em: 29/10/2018.

BANCA EXAMINADORA

Prof. Dr. Rafael Bráz Azevedo Farias (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Juvêncio Santos Nobre (Coorientador)
Universidade Federal do Ceará (UFC)

Prof^ª Dra. Rosineide Fernando da Paz
Universidade Federal do Ceará (UFC)

Prof. Dr. Guilherme de Alencar Barreto
Universidade Federal do Ceará (UFC)

Dedico este trabalho meu irmão Fly que faleceu no ano de 2017 e minha tia avó madrinha Maria do Carmo (Tutu) que faleceu em 12-09-2018.

AGRADECIMENTOS

A lista de agradecimentos é bastante longa e envolve pessoas que estão na (ou passaram pela) minha vida desde 1990. Assim, inicio a lista de agradecimentos pela minha avó, Maria V. Sena, que me deu o suporte para que eu chegasse até este ponto. Em segundo lugar, agradeço a minha mãe, Abraídes Sena, por tudo que me proporcionou e principalmente pela oportunidade de viver (obs: a vó é mais velha, então deve ser citada primeira). Acredito que estas duas mulheres são a base da minha vida e palavras não podem descrever a importância delas.

Dando continuação a lista dos familiares, devo agradecer a Luna Sena (Cokinha), Eliete Sena (tia), a Camila Sena (prima), ao David Barroso (primo), a Maria do Carmo Barroso (minha madrinha) e os meus tios Francisco Sena, Tecla Sena e Fernando Sena.

Agora visando agradecer as instituições, devo destacar que o presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Este apoio foi por meio da bolsa de estudos que a Capes fornece aos alunos de Mestrado. Ainda pensando nas instituições, também quero agradecer à Universidade Federal do Ceará - UFC por apoiar e fornecer suporte durante a graduação e o mestrado.

Continuando a lista, mas agora na seção de amigos que surgiram no período de 1994 a 2008, quero agradecer inicialmente à família Barbosa: João Neto, Maria José, Leonillson e Joannillson. Em seguida, quero agradecer aos amigos (e seus familiares) que surgiram posteriormente a família Barbosa. Estes estimados amigos são: Levine Santiago, Victor Reis, Thiago Nepomuceno, George V. Mesquita, Allyson Fernandes, Diego Araujo, Hemerson Braga, Luís de Alcantra (Lee-Jr), Magno Barbosa, Marcio Barbosa, Romualdo Cavalcante, Felipe Moura, João Virino Neto, Josué Oliveira, Cavalcante Filho, Márcia Rodrigues e Sophia Rodrigues.

Seguindo na linha do tempo, agora no período de 2009 a 2015. Quero agradecer aos amigos que me ajudaram bastante durante todo período relacionado com a graduação. Assim, meus agradecimentos vão para Marina Sales, Cristina Guedes, Widemberg Nobre, Hildemar Calixto, Thiago Oliveira, Antony T. Dias e Erione Melo.

Ainda no mesmo período, agora visando os professores que tive a oportunidade de conhecer na graduação. Quero agradecer, por ordem das disciplinas e orientações, aos professores: Ana Maria Araújo, Neuman de Souza, César Teixeira, Alton Andrade, Juvêncio Nobre (meu ídolo), Maurício Mota, Sílvia Freitas, André Jalles e Rafael Farias.

Avançando para o mestrado, agora no período de 2016 a 2018. Quero agradecer, por ordem de disciplinas, aos professores e funcionários: Manoel Campêlo Neto, Ricardo Coelho, Guilherme Barreto e Wendell Matos.

Continuando no mesmo período, agora visando os amigos do programa de mestrado. Quero agradecer aos super alunos: Chagas Junior, Daniel Holanda, Francílio Araújo, Francinilton Silva, Judecir Júnior, Lívia Oliveira e Ronaldo Pessoa.

Agora na seção referente ao período de 2007, 2008 e 2016, quero agradecer a três garotas muito especiais na minha vida. Elas são Myla Miranda (Vampira), Karen Jeyne Lima (Gurupi, o grande amor que perdi) e Vanessa Garcia (Pintadinha querida).

Seguindo para a seção de tempo indeterminado, quero agradecer a você que está lendo este trabalho.

Por fim, dizem que o melhor sempre fica por último, então finalizo a lista agradecendo a Deus por todas as oportunidades e graças que me foram fornecidas até o presente momento.

“Fraco é aquele que não quer se superar.”

Player Tauz Ft. Wendel Bezerra.

RESUMO

Neste trabalho estudamos a distribuição L-logística proposta em Paz et al. (2016) e seu respectivo modelo de regressão para a modelagem de variáveis em um intervalo limitado, em geral $(0, 1)$, tais como taxas e proporções. No que concerne à modelagem deste tipo de variáveis, encontra-se na literatura um percentual significativo voltado para os convencionais modelos Beta e Simplex. Com o objetivo de diversificar a literatura, este trabalho realiza suas inferências por meio do método de estimação por máxima verossimilhança (ML) para contribuir com a vertente frequentista, tendo em vista que o trabalho realizado em Paz et al. (2016) apresenta somente a vertente bayesiana. O método de inferência é realizado através da combinação de algoritmos evolucionários (Algoritmo Genético e Evolução Diferencial) para a obtenção das estimativas pontuais (dispensando o uso de derivadas no processo) e métodos de reamostragem (Jackknife e Bootstrap) para a obtenção das estimativas do erro-padrão e intervalos de confiança. Simulações são realizadas para ilustrar a combinação das técnicas e comprovar sua eficiência. Logo após, o potencial do modelo é demonstrado através de duas aplicações em dados reais, no qual constata-se um melhor desempenho do modelo de regressão L-logística com relação aos modelos de regressão Beta e Simplex (em termos de medidas de qualidade do ajuste e diagnósticos de pontos influentes).

Palavras-chave: Modelo de Regressão L-logística. Máxima Verossimilhança. Metaheurística. Reamostragem.

ABSTRACT

In this work we study the L-logistic distribution proposed in Paz et al. (2016) and its respective regression model for modeling variables in a limited range, generally $(0, 1)$, such as rates and proportions. As far as the modeling of this type of variables is concerned, a significant percentage is found in the literature for Beta and Simplex models. With the objective of diversifying the literature, this work makes its inferences through of the estimation method by maximum likelihood (ML) to contribute to the frequentist method, considering that the work done in Paz et al. (2016) presents only the Bayesian method. The method of inference is accomplished through the combination of evolutionary algorithms (Genetic Algorithm and Differential Evolution) to obtain the point estimates (dispensing the use of derivatives in the process) and resampling methods (Jackknife and Bootstrap) to obtain the estimates of the standard error and confidence intervals. Simulations are performed to illustrate the combination of techniques and to prove their efficiency. Then, the potential of the model is demonstrated through two applications in real data, in which a better performance of the L-logistic regression model is verified in relation to the Beta and Simplex regression models (in terms of goodness of fit measures and a measure to influential points).

Keywords: L-logistic Regression Model. Maximum Likelihood. Metaheuristics. Resampling Methods.

LISTA DE FIGURAS

Figura 1 – Formas da distribuição L-logística sobre a perspectiva de m	20
Figura 2 – Formas da distribuição L-logística sobre a perspectiva de b	20
Figura 3 – Comportamento do Intervalo Interquartilico para alguns valores de m fixado.	21
Figura 4 – Comportamento do Intervalo Interquartilico para alguns valores de b fixado.	21
Figura 5 – Comportamento da Média e Desvio Padrão, segundo alguns valores dos parâmetros da distribuição L-logística.	23
Figura 6 – Comportamento da Moda e das medidas γ_M , $\gamma_{0.125}$ e γ_I , segundo alguns valores dos parâmetros da distribuição L-logística.	24
Figura 7 – Comportamento da medida γ_N , segundo alguns valores dos parâmetros da distribuição L-logística.	25
Figura 8 – Comportamento da medida k_q , segundo alguns valores dos parâmetros da distribuição L-logística.	25
Figura 9 – Comportamento da medida K_N , segundo alguns valores dos parâmetros da distribuição L-logística.	26
Figura 10 – Ilustração gráfica do comportamento de todos os cenários usados na simulação. Histogramas produzidos com amostras de tamanho 5×10^5 da distribuição L-logística.	44
Figura 11 – Gráfico de caixa da taxa de doenças cardíacas segundo a variável sexo.	87
Figura 12 – Gráfico de caixa da taxa de doenças cardíacas segundo a variável etnia.	87
Figura 13 – Gráfico da distância do tipo Cook para os modelo de regressão com função de ligação Logit. Perspectiva do teste da razão de verossimilhanças generalizada.	109
Figura 14 – Gráfico da distância do tipo Cook para o modelo de regressão com função de ligação Probit. Perspectiva do teste da razão de verossimilhanças generalizada.	109
Figura 15 – Gráfico da distância do tipo Cook para o modelo de regressão com função de ligação Complemento Log-log. Perspectiva do teste da razão de verossimilhanças generalizada.	110
Figura 16 – Gráfico da distância do tipo Cook para o modelo de regressão com função de ligação Logit. Perspectiva do teste do tipo Wald.	110
Figura 17 – Gráfico da distância do tipo Cook para o modelo de regressão com função de ligação Probit. Perspectiva do teste do tipo Wald.	111
Figura 18 – Gráfico da distância do tipo Cook para o modelo de regressão com função de ligação Complemento Log-log. Perspectiva do teste do tipo Wald.	111
Figura 19 – Gráfico de dispersão da variáveis.	113
Figura 20 – Gráfico de dispersão da variáveis após a transformação.	114
Figura 21 – Gráfico da distância do tipo Cook para os modelo de regressão com função de ligação Logit.	144
Figura 22 – Gráfico da distância do tipo Cook para o modelo de regressão com função de ligação Probit.	144
Figura 23 – Gráfico da distância do tipo Cook para o modelo de regressão com função de ligação Complemento Log-log.	145

LISTA DE TABELAS

Tabela 1 – Configuração dos Algoritmos.	44
Tabela 2 – Resultados da simulação para o cenário 1. $\theta = (m; b)^\top = (0, 2; 2)^\top$	45
Tabela 3 – Resultados da simulação para o cenário 2. $\theta = (m; b)^\top = (0, 5; 2)^\top$	45
Tabela 4 – Resultados da simulação para o cenário 3. $\theta = (m; b)^\top = (0, 9; 2)^\top$	46
Tabela 5 – Resultados da simulação para o cenário 4. $\theta = (m; b)^\top = (0, 2; 1)^\top$	46
Tabela 6 – Resultados da simulação para o cenário 5. $\theta = (m; b)^\top = (0, 5; 1)^\top$	47
Tabela 7 – Resultados da simulação para o cenário 6. $\theta = (m; b)^\top = (0, 9; 1)^\top$	47
Tabela 8 – Resultados da simulação para o cenário 7. $\theta = (m; b)^\top = (0, 2; 0, 5)^\top$	48
Tabela 9 – Resultados da simulação para o cenário 8. $\theta = (m; b)^\top = (0, 5; 0, 5)^\top$	48
Tabela 10 – Resultados da simulação para o cenário 9. $\theta = (m; b)^\top = (0, 9; 0, 5)^\top$	49
Tabela 11 – Probabilidade de cobertura com Algoritmo Genético para $m = 0, 2$ no cenário 1.	50
Tabela 12 – Probabilidade de cobertura com Evolução Diferencial para $m = 0, 2$ no cenário 1.	51
Tabela 13 – Probabilidade de cobertura com Algoritmo Genético para $b = 2$ no cenário 1.	52
Tabela 14 – Probabilidade de cobertura com Evolução Diferencial para $b = 2$ no cenário 1.	53
Tabela 15 – Probabilidade de cobertura com Algoritmo Genético para $m = 0, 5$ no cenário 2.	54
Tabela 16 – Probabilidade de cobertura com Evolução Diferencial para $m = 0, 5$ no cenário 2.	55
Tabela 17 – Probabilidade de cobertura com Algoritmo Genético para $b = 2$ no cenário 2.	56
Tabela 18 – Probabilidade de cobertura com Evolução Diferencial para $b = 2$ no cenário 2.	57
Tabela 19 – Probabilidade de cobertura com Algoritmo Genético para $m = 0, 9$ no cenário 3.	58
Tabela 20 – Probabilidade de cobertura com Evolução Diferencial para $m = 0, 9$ no cenário 3.	59
Tabela 21 – Probabilidade de cobertura com Algoritmo Genético para $b = 2$ no cenário 3.	60
Tabela 22 – Probabilidade de cobertura com Evolução Diferencial para $b = 2$ no cenário 3.	61
Tabela 23 – Probabilidade de cobertura com Algoritmo Genético para $m = 0, 2$ no cenário 4.	62
Tabela 24 – Probabilidade de cobertura com Evolução Diferencial para $m = 0, 2$ no cenário 4.	63
Tabela 25 – Probabilidade de cobertura com Algoritmo Genético para $b = 1$ no cenário 4.	64
Tabela 26 – Probabilidade de cobertura com Evolução Diferencial para $b = 1$ no cenário 4.	65
Tabela 27 – Probabilidade de cobertura com Algoritmo Genético para $m = 0, 5$ no cenário 5.	66
Tabela 28 – Probabilidade de cobertura com Evolução Diferencial para $m = 0, 5$ no cenário 5.	67
Tabela 29 – Probabilidade de cobertura com Algoritmo Genético para $b = 1$ no cenário 5.	68
Tabela 30 – Probabilidade de cobertura com Evolução Diferencial para $b = 1$ no cenário 5.	69
Tabela 31 – Probabilidade de cobertura com Algoritmo Genético para $m = 0, 9$ no cenário 6.	70
Tabela 32 – Probabilidade de cobertura com Evolução Diferencial para $m = 0, 9$ no cenário 6.	71
Tabela 33 – Probabilidade de cobertura com Algoritmo Genético para $b = 1$ no cenário 6.	72
Tabela 34 – Probabilidade de cobertura com Evolução Diferencial para $b = 1$ no cenário 6.	73
Tabela 35 – Probabilidade de cobertura com Algoritmo Genético para $m = 0, 2$ no cenário 7.	74
Tabela 36 – Probabilidade de cobertura com Evolução Diferencial para $m = 0, 2$ no cenário 7.	75
Tabela 37 – Probabilidade de cobertura com Algoritmo Genético para $b = 0, 5$ no cenário 7.	76
Tabela 38 – Probabilidade de cobertura com Evolução Diferencial para $b = 0, 5$ no cenário 7.	77
Tabela 39 – Probabilidade de cobertura com Algoritmo Genético para $m = 0, 5$ no cenário 8.	78
Tabela 40 – Probabilidade de cobertura com Evolução Diferencial para $m = 0, 5$ no cenário 8.	79
Tabela 41 – Probabilidade de cobertura com Algoritmo Genético para $b = 0, 5$ no cenário 8.	80
Tabela 42 – Probabilidade de cobertura com Evolução Diferencial para $b = 0, 5$ no cenário 8.	81
Tabela 43 – Probabilidade de cobertura com Algoritmo Genético para $m = 0, 9$ no cenário 9.	82
Tabela 44 – Probabilidade de cobertura com Evolução Diferencial para $m = 0, 9$ no cenário 9.	83

Tabela 45 – Probabilidade de cobertura com Algoritmo Genético para $b = 0,5$ no cenário 9.	84
Tabela 46 – Probabilidade de cobertura com Evolução Diferencial para $b = 0,5$ no cenário 9.	85
Tabela 47 – Configuração dos Algoritmos.	86
Tabela 48 – Notação das covariáveis nos modelos.	88
Tabela 49 – Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Logit.	90
Tabela 50 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Logit.	90
Tabela 51 – Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Logit.	91
Tabela 52 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Logit.	91
Tabela 53 – Teste da Razão de Verossimilhanças Generalizada para a função de ligação Logit.	92
Tabela 54 – Teste tipo Wald com combinação do Algoritmo Genético com método Bootstrap para a função de ligação Logit e remoção da categoria Caucasianos.	92
Tabela 55 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Bootstrap para a função de ligação Logit e remoção da categoria Caucasianos.	92
Tabela 56 – Teste tipo Wald com combinação do Algoritmo Genético com método Jackknife para a função de ligação Logit e remoção da categoria Caucasianos.	93
Tabela 57 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Jackknife para a função de ligação Logit e remoção da categoria Caucasianos.	93
Tabela 58 – Medidas de qualidade do ajuste com todas a covariáveis e função de ligação Logit.	93
Tabela 59 – Medidas de qualidade do ajuste para a função de ligação Logit e a remoção da categoria Caucasianos dos modelos de regressão L-logística e Simplex.	94
Tabela 60 – Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Logit.	94
Tabela 61 – Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Logit e remoção da categoria Caucasianos.	95
Tabela 62 – Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Probit.	96
Tabela 63 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Probit.	96
Tabela 64 – Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Probit.	97
Tabela 65 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Probit.	97
Tabela 66 – Teste da Razão de Verossimilhanças Generalizada para a função de ligação Probit.	98
Tabela 67 – Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Probit e remoção da categoria Caucasianos.	98
Tabela 68 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Probit e remoção da categoria Caucasianos.	98
Tabela 69 – Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Probit e remoção da categoria Caucasianos.	99
Tabela 70 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Probit e remoção da categoria Caucasianos.	99
Tabela 71 – Medidas de qualidade do ajuste com todas a covariáveis e função de ligação Probit.	99

Tabela 72 – Medidas de qualidade do ajuste para a função de ligação Probit e a remoção da categoria Caucasianos dos modelos de regressão L-logística e Simplex.	100
Tabela 73 – Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Probit.	100
Tabela 74 – Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Probit e remoção da categoria Caucasianos.	101
Tabela 75 – Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Complemento Log-log.	102
Tabela 76 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Complemento Log-log.	102
Tabela 77 – Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Complemento Log-log.	103
Tabela 78 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Complemento Log-log.	103
Tabela 79 – Teste da Razão de Verossimilhanças Generalizada para a função de ligação Complemento Log-log.	104
Tabela 80 – Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Complemento Log-log e remoção da categoria Caucasianos.	104
Tabela 81 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Complemento Log-log e remoção da categoria Caucasianos.	104
Tabela 82 – Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Complemento Log-log e remoção da categoria Caucasianos.	105
Tabela 83 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Complemento Log-log e remoção da categoria Caucasianos.	105
Tabela 84 – Medidas de qualidade do ajuste com todas a covariáveis e função de ligação Complemento Log-log.	105
Tabela 85 – Medidas de qualidade do ajuste para a função de ligação Complemento Log-log e a remoção da categoria Caucasianos dos modelos de regressão L-logística e Simplex.	106
Tabela 86 – Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Complemento Log-log.	106
Tabela 87 – Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Complemento Log-log e remoção da categoria Caucasianos.	107
Tabela 88 – Estatísticas da distância do tipo Cook para todos os modelos e todas as funções de ligação. Perspectiva do teste da razão de verossimilhanças generalizadas.	108
Tabela 89 – Estatísticas da distância do tipo Cook para todos os modelos e todas as funções de ligação. Perspectiva do teste do tipo Wald.	108
Tabela 90 – Informações das variáveis.	112
Tabela 91 – Tabela de correlação das informações.	113
Tabela 92 – Transformação das covariáveis.	113
Tabela 93 – Tabela de correlação das informações transformadas.	114
Tabela 94 – Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Logit.	116
Tabela 95 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Logit.	117
Tabela 96 – Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Logit.	118

Tabela 97 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Logit.	119
Tabela 98 – Teste da Razão de Verossimilhanças Generalizada para a função de ligação Logit.	120
Tabela 99 – Teste tipo Wald com combinação do Algoritmo Genético com método Bootstrap para a função de ligação Logit e remoção das covariáveis x_4 , x_5 e x_6	120
Tabela 100 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Bootstrap para a função de ligação Logit e remoção das covariáveis x_4 , x_5 e x_6	121
Tabela 101 – Teste tipo Wald com combinação do Algoritmo Genético com método Jackknife para a função de ligação Logit e remoção das covariáveis x_4 , x_5 e x_6	121
Tabela 102 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Jackknife para a função de ligação Logit e remoção das covariáveis x_4 , x_5 e x_6	122
Tabela 103 – Medidas de qualidade do ajuste para a função de ligação Logit.	122
Tabela 104 – Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Logit.	123
Tabela 105 – Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Logit e remoção das covariáveis x_4 , x_5 e x_6	124
Tabela 106 – Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Probit.	125
Tabela 107 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Probit.	126
Tabela 108 – Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Probit.	127
Tabela 109 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Probit.	128
Tabela 110 – Teste da Razão de Verossimilhanças Generalizada para a função de ligação Probit.	129
Tabela 111 – Teste tipo Wald com combinação do Algoritmo Genético com método Bootstrap para a função de ligação Probit e remoção das covariáveis x_4 , x_5 e x_6	129
Tabela 112 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Bootstrap para a função de ligação Probit e remoção das covariáveis x_4 , x_5 e x_6	130
Tabela 113 – Teste tipo Wald com combinação do Algoritmo Genético com método Jackknife para a função de ligação Probit e remoção das covariáveis x_4 , x_5 e x_6	130
Tabela 114 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Jackknife para a função de ligação Probit e remoção das covariáveis x_4 , x_5 e x_6	131
Tabela 115 – Medidas de qualidade do ajuste para a função de ligação Probit.	131
Tabela 116 – Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Probit.	132
Tabela 117 – Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Probit e remoção das covariáveis x_4 , x_5 e x_6	133
Tabela 118 – Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Complemento Log-Log.	134
Tabela 119 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Complemento Log-Log.	135
Tabela 120 – Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Complemento Log-Log.	136
Tabela 121 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Complemento Log-Log.	137

Tabela 122 – Teste da Razão de Verossimilhanças Generalizada para a função de ligação Complemento Log-Log.	138
Tabela 123 – Teste tipo Wald com combinação do Algoritmo Genético com método Bootstrap para a função de ligação Complemento Log-Log e remoção das covariáveis x_4 , x_5 e x_6	138
Tabela 124 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Bootstrap para a função de ligação Complemento Log-Log e remoção das covariáveis x_4 , x_5 e x_6	139
Tabela 125 – Teste tipo Wald com combinação do Algoritmo Genético com método Jackknife para a função de ligação Complemento Log-Log e remoção das covariáveis x_4 , x_5 e x_6	139
Tabela 126 – Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Jackknife para a função de ligação Complemento Log-Log e remoção das covariáveis x_4 , x_5 e x_6	140
Tabela 127 – Medidas de qualidade do ajuste para a função de ligação Complemento Log-Log.	140
Tabela 128 – Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Complemento Log-Log.	141
Tabela 129 – Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Complemento Log-Log e remoção das covariáveis x_4 , x_5 e x_6	142
Tabela 130 – Estatísticas da distância do tipo Cook para todos os modelos e todas as funções de ligação.	143

SUMÁRIO

1	INTRODUÇÃO	17
1.1	Organização da Dissertação	17
2	MODELO DE REGRESSÃO L-LOGÍSTICO	19
2.1	Distribuição L-logística	19
2.2	Propriedades da Distribuição L-logística	21
2.2.1	Relação com outras distribuições	21
2.2.2	Moda	22
2.2.3	Momentos	23
2.2.4	Assimetria e Curtose	23
2.3	Estrutura de Regressão	26
3	INFERÊNCIA	28
3.1	Estimação por Máxima Verossimilhança	28
3.2	Ideia sobre Metaheurísticas	29
3.3	Algoritmo Genético	29
3.3.1	Operador de Seleção	31
3.3.2	Operador de Recombinação	32
3.3.3	Operador de Mutação	33
3.4	Evolução Diferencial	34
3.4.1	Operador de Mutação	35
3.4.2	Operador de Recombinação	35
3.4.3	Operador de Seleção	36
3.4.4	Outras variações	36
3.5	Erro de Estimação	36
3.6	Método Jackknife	37
3.6.1	Apresentação do método	37
3.6.2	Intervalo de Confiança	38
3.7	Método Bootstrap	38
3.7.1	Apresentação do método	38
3.7.2	Intervalos de Confiança	40
3.8	Testes de Hipóteses	41
3.8.1	Teste da Razão de Verossimilhanças Generalizada	41
3.8.2	Teste do tipo Wald com Bootstrap	41
3.9	Medidas de Diagnóstico	42
4	SIMULAÇÃO	43
4.1	Resultados: Estatísticas da Simulação	44
4.2	Resultados: Métodos de Reamostragem	49
5	APLICAÇÃO EM DADOS REAIS	86
5.1	Introdução	86
5.2	Aplicação I	86
5.2.1	Modelos de Regressão: Resultados e Conclusões	87
5.3	Aplicação II	112
5.3.1	Modelos de Regressão: Resultados e Conclusões	114
6	CONCLUSÕES	146
6.1	Considerações Finais	146

6.2	Estudos futuros	146
	REFERÊNCIAS	147
	APÊNDICE A – DERIVADAS PARCIAIS	152

1 INTRODUÇÃO

A análise de informação é fundamental quando deseja-se explicar algum fenômeno de interesse, concluir acerca de um determinado questionamento ou (principalmente) efetuar a melhor decisão.

Hoje uma boa avaliação dos dados ocupa um lugar significativo nas grandes áreas do conhecimento. Em especial, decorrente do avanço da estatística, a análise de dados numéricos foi simplificada para muitas dessas áreas, acarretando uma maior interdisciplinaridade. Dentre a expansiva quantidade de dados no mundo, há uma parcela crescente de estudiosos interessados em dados que desenvolvem suas características ao longo de um determinado intervalo (real) de interesse, regularmente o intervalo $[0, 1]$. Dados que expressam proporções e taxas, são geralmente encontrados na literatura, como ilustrações de dados no intervalo $[0, 1]$. No que diz respeito as análises desse tipo de fenômeno (dados no intervalo unitário), as metodologias variam entre análise descritiva dos dados e modelagem preditiva.

No que remete à modelagem, encontra-se na literatura um conjunto de modelos consolidados para trabalhar fenômenos (variáveis) com suporte limitado ou $[0, 1]$, dos quais cita-se as distribuições Simplex, Beta, Kumaraswamy e seus respectivos formatos (modelos) de regressão que podem ser vistos por exemplo em Song e Tan (2000), Ferrari e Cribari-Neto (2004) e Mitnik e Baek (2013). Contudo, embora plausíveis de uso, estas modelagens podem apresentar desvantagens tais como o uso de funções especiais, a falta de forma analítica para medidas de posição (ou dispersão), a falta de flexibilidade ou a complexidade de implementação computacional. Mesmo com algumas destas desvantagens, vale destacar o modelo Beta para a modelagem de dados no intervalo unitário (com ou sem estrutura de regressão), por apresentar boas propriedades e ser muito flexível. Além disso, por sua popularidade, o modelo Beta também recebeu muitas generalizações e expansões, como por exemplo, o modelo Beta Generalizado apresentada em McDonald (1984), o modelo Beta Retangular apresentado em Hahn (2008) e o modelo Beta Invertido (*McDonald Inverted Beta Distribution*) apresentado em Cordeiro e Lemonte (2012).

As pesquisas que visam variáveis com o suporte $[0, 1]$ ou limitado, estão em ritmo crescente. No entanto, um percentual significativo destas pesquisas está concentrada nas distribuições convencionais, com destaque na distribuição Beta, principalmente quando existe estrutura de regressão. Desta forma, torna-se importante proporcionar novas alternativas para as metodologias clássicas, afim de enriquecer a literatura.

Recentemente Paz et al. (2018) propuseram a distribuição L-logística, assim como sua estrutura de regressão, como modelagem alternativa robusta ao modelo Beta. Todavia, a apresentação da metodologia é feita sobre o enfoque bayesiano. Deste modo, neste trabalho apresenta-se o enfoque frequentista para a proposta de Paz et al. (2016), no qual realiza-se as inferências através do método de máxima verossimilhança proporcionado pela combinação de algoritmos evolutivos (Algoritmo Genético e Evolução Diferencial) para a obtenção das estimativas pontuais e métodos de reamostragem (Jackknife e Bootstrap) para a obtenção do erro padrão das estimativas e intervalos de confiança.

1.1 Organização da Dissertação

O presente trabalho é constituído por 6 capítulos. No Capítulo 2 apresentamos a definição da distribuição L-logística e suas propriedades, assim como a estrutura de regressão e os métodos de diagnóstico. No Capítulo 3 apresentamos o método de estimação por máxima verossimilhança, baseado na combinação de metaheurísticas de otimização contínua e métodos de reamostragem (Jackknife e Bootstrap), além de apresentar cada metaheurística e método de reamostragem de forma isolada. No Capítulo 4 apresentamos o estudo de simulação para a distribuição L-logística, afim de mostrar a eficiência da estimação por máxima verossimilhança através da combinação das técnicas. No Capítulo 5 apresentamos a modelagem de dois conjuntos de dados reais. O primeiro conjunto trata de informações referentes

à problemas cardíacos em adultos maiores de 35 nos estados dos Estados Unidos. Nesta aplicação, visa-se modelar a taxa de problemas cardíacos com o auxílio das covariáveis etnia e sexo. Já o segundo conjunto de dados refere-se à modelagem do índice da percepção da corrupção de 2016 com o auxílio dos seguintes indicadores como covariáveis: índice liberdade econômica de 2016, produto interno bruto per capita de 2016, índice de democracia de 2016, índice de desenvolvimento humano de 2015, densidade populacional de 2015 e índice de inflação de 2016. No Capítulo 6 apresentamos os comentários finais. No Apêndice A apresentamos as derivadas parciais da distribuição L-logística.

2 MODELO DE REGRESSÃO L-LOGÍSTICO

2.1 Distribuição L-logística

A distribuição L-logística, apresentada em Paz et al. (2018), consiste em um modelo para variáveis contínuas com suporte $[0, 1]$ que surge como reparametrização da distribuição L_b proposta em Tadikamalla e Johnson (1989). Vale apontar que embora o conteúdo sobre a distribuição L-logística esteja formalizada em Paz et al. (2018), cronologicamente a pesquisa foi inicialmente apresentada em Paz et al. (2016) e este fato forneceu a bibliografia necessária (base) para esta dissertação.

Uma variável aleatória Y é dita possuir distribuição L-logística, denotada por $Y \sim LL(m, b)$, se sua função de densidade de probabilidade é dada por

$$f_Y(y; m, b) = \frac{b(1-m)^b m^b y^{b-1} (1-y)^{b-1}}{[(1-m)^b y^b + m^b (1-y)^b]^2} \mathbf{1}_{(0,1)}(y), \quad 0 < m < 1, \quad b > 0. \quad (1)$$

Os parâmetros m e b são classificados como parâmetros de posição e dispersão, respectivamente, no qual o parâmetro de posição denota a mediana da distribuição. Além disso, estes parâmetros proporcionam uma variedade de formas para a função de densidade, demonstrando flexibilidade e características de unimodalidade, crescimento, decrescimento e constância, como ilustrado nas Figuras 1 e 2.

A função de distribuição acumulada de uma variável aleatória $LL(m, b)$ é dada por

$$F_Y(y; m, b) = \left(1 + \left(\frac{m(1-y)}{(1-m)y} \right)^b \right)^{-1}, \quad (2)$$

de forma que a respectiva função quantílica é dada por

$$Q_Y(\alpha; m, b) = F_Y^{-1}(\alpha; m, b) = \frac{\alpha^{1/b} m}{(1-\alpha)^{1/b} (1-m) + \alpha^{1/b} m}, \quad (3)$$

no qual α denota a probabilidade associada ao α -ésimo quantil da distribuição L-logística, $\alpha \in (0, 1)$.

A Equação 3 fornece o desvio interquartil (IQR) expresso expresso por

$$IQR = Q_Y(0,75; m, b) - Q_Y(0,25; m, b) = \frac{3^{1/b} m}{(1-m) + 3^{1/b} m} - \frac{m}{m + 3^{1/b} (1-m)}, \quad (4)$$

além de fornecer um método para gerar variáveis aleatórias com distribuição $LL(m, b)$, no qual se $U \sim \text{Uniforme}(0, 1)$, então

$$Y = Q(U; m, b) = \frac{U^{1/b} m}{(1-U)^{1/b} (1-m) + U^{1/b} m} \sim LL(m, b). \quad (5)$$

A seguir, nas Figuras 3 e 4, o comportamento do IQR para alguns valores dos parâmetros de posição e dispersão.

Figura 1: Formas da distribuição L-logística sobre a perspectiva de m .

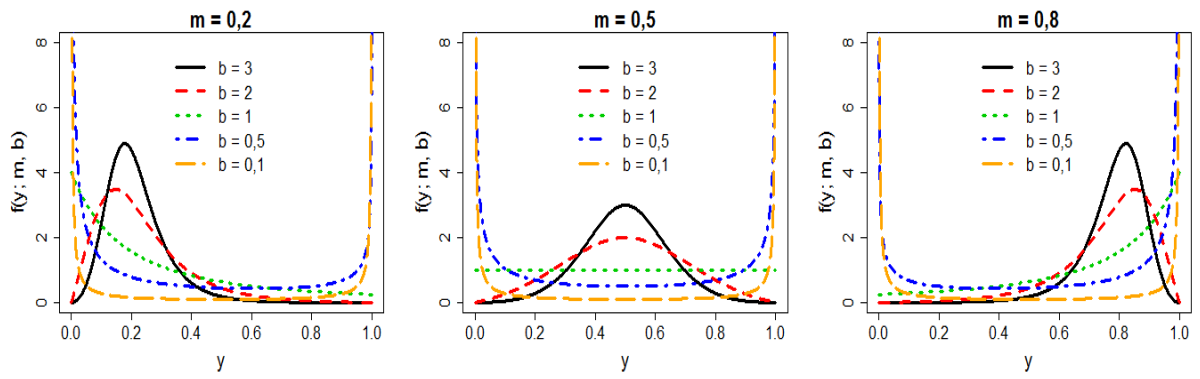


Figura 2: Formas da distribuição L-logística sobre a perspectiva de b .

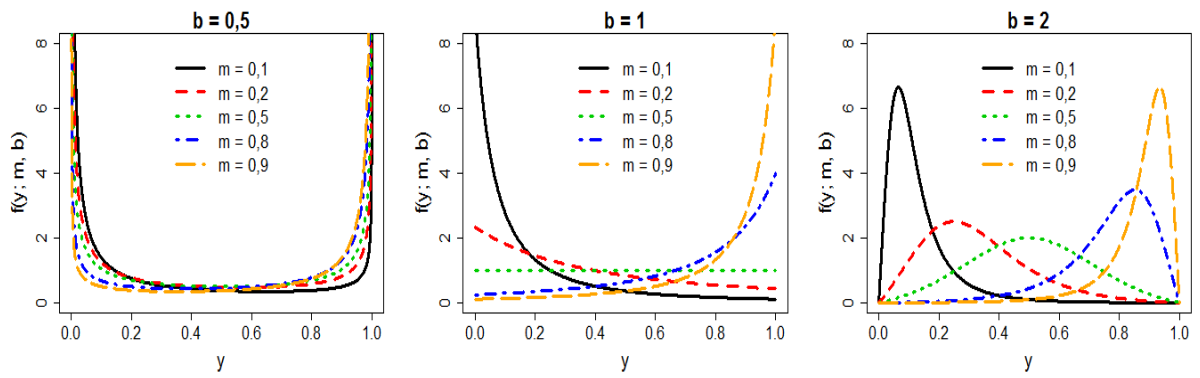


Figura 3: Comportamento do Intervalo Interquartílico para alguns valores de m fixado.

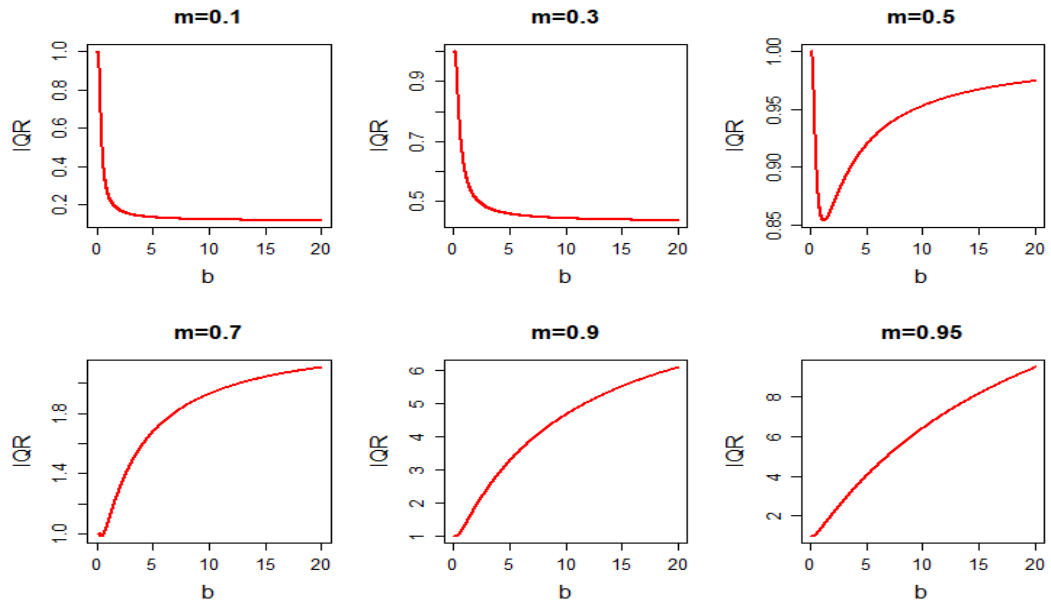
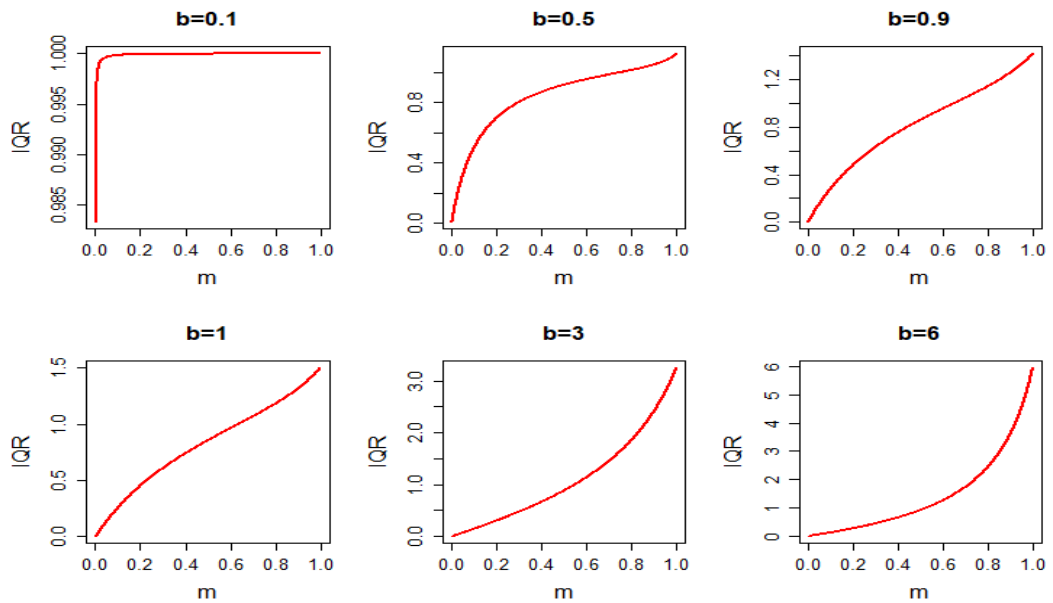


Figura 4: Comportamento do Intervalo Interquartílico para alguns valores de b fixado.



2.2 Propriedades da Distribuição L-logística

Nesta seção apresenta-se algumas propriedades da distribuição L-logística, tais como a relação com algumas distribuições, moda, momentos e medidas de assimetria e curtose, com todas as demonstrações presentes em Paz et al. (2016).

2.2.1 Relação com outras distribuições

A propriedade a seguir, demonstra a relação entre a distribuição L-logística e a distribuição Logística.

Propriedade 1. Se $Y \sim LL(m, b)$, então $Z = \ln\left(\frac{Y(1-m)}{(1-Y)m}\right)$ tem distribuição Logística Padrão.

As duas próximas propriedades apresentam parametrizações alternativas da distribuição L-logística.

Propriedade 2. Se $Y \sim LL(m, b)$, então a reparametrização $m = e^{-\frac{\delta}{b}}\left(1 + e^{-\frac{\delta}{b}}\right)^{-1}$, com $\delta > 0$ e $b > 0$, acarreta a distribuição L_b na qual sua função de densidade é dada por

$$f_Y(y; \delta, b) = \frac{b e^{\delta} y^{b-1} (1-y)^{b-1}}{(e^{\delta} y^b + (1-y)^b)^2} \mathbb{1}_{(0,1)}(y), \quad (6)$$

em que δ e b denotam parâmetros de dispersão e sua função de distribuição acumulada é definida por

$$F_Y(y; \delta, b) = \left(1 + e^{-\delta} \left(\frac{1-y}{y}\right)^b\right)^{-1}. \quad (7)$$

Propriedade 3. Se $Y \sim LL(m, b)$, então para $b > 0$ e $\kappa = \left(1 + \left(\frac{m}{1-m}\right)^b\right)^{-1}$, com $\kappa \in (0, 1)$, obtém-se a forma alternativa da distribuição L-logística, $Y \sim LL(\kappa, b)$, dada por

$$f_Y(y; \kappa, b) = \frac{b \kappa (1-\kappa) y^{b-1} (1-y)^{b-1}}{[\kappa y^b + (1-\kappa)(1-y)^b]^2} \mathbb{1}_{(0,1)}(y), \quad (8)$$

em que κ e b denotam parâmetros de dispersão e sua função de distribuição é definida por

$$F_Y(y; \delta, b) = \left[1 + \left(\frac{1-\kappa}{\kappa}\right) \left(\frac{1-y}{y}\right)^b\right]^{-1}. \quad (9)$$

Como visto em Paz et al. (2016), embora as distribuições apresentadas nas propriedades 2 e 3 possuam expressões simples, a distribuição L-logística possui a mediana como parâmetro natural, facilitando a interpretação do modelo e proporcionando o uso da análise de regressão. Além disso, como visto em Tadikamalla e Johnson (1982) e Wang e Rennolls (2005), a distribuição L_b surge através da transformação da distribuição logística padrão e pode ser estendida para o suporte limitado. A seguir, a próxima propriedade apresenta a extensão da distribuição L-logística para casos com o suporte limitado.

Propriedade 4. A extensão da distribuição L-logística para uma variável com suporte limitado (c, d) , é obtida através da transformação $Y = (d-c)X + c$, com $X \sim LL(m, b)$, e possui função de densidade dada por

$$f_Y(y; m, b, c, d) = \frac{(d-c)b(1-m)^b m^b (y-c)^{b-1} (d-y)^{b-1}}{[(1-m)^b (y-c)^b + m^b (d-y)^b]^2} \mathbb{1}_{(c,d)}(y), \quad (10)$$

em que $m \in (0, 1)$, $b > 0$, $c < d$, $c, d \in \mathbb{R}$ e sua função de distribuição é definida por

$$F_Y(y; m, b, c, d) = \frac{[(y-c) - m(y-c)]^b}{[(y-c) - m(y-c)]^b + [m(d-y)]^b}. \quad (11)$$

2.2.2 Moda

Como apresentado em Paz et al. (2016), a distribuição L-logística possui moda y_0 somente para $b > 1$, onde pode ser obtida através da solução da equação a seguir

$$\left(\frac{1-m}{m}\right)^b = \left(\frac{1-y_0}{y_0}\right)^b \frac{b+2y_0-1}{b-2y_0+1}. \quad (12)$$

Os autores também sugerem assumir $\varphi = -b \log\left(\frac{m}{1-m}\right)$ e reescrever a Equação 12 da forma

$$\varphi = \log\left(\left(\frac{1-y_0}{y_0}\right)^b \frac{b+2y_0-1}{b-2y_0+1}\right), \quad (13)$$

para facilitar o processo e proporcionar o uso do teste da primeira derivada para estudar o comportamento da moda.

2.2.3 Momentos

O t -ésimo momento não central da distribuição $LL(m, b)$ é dado por

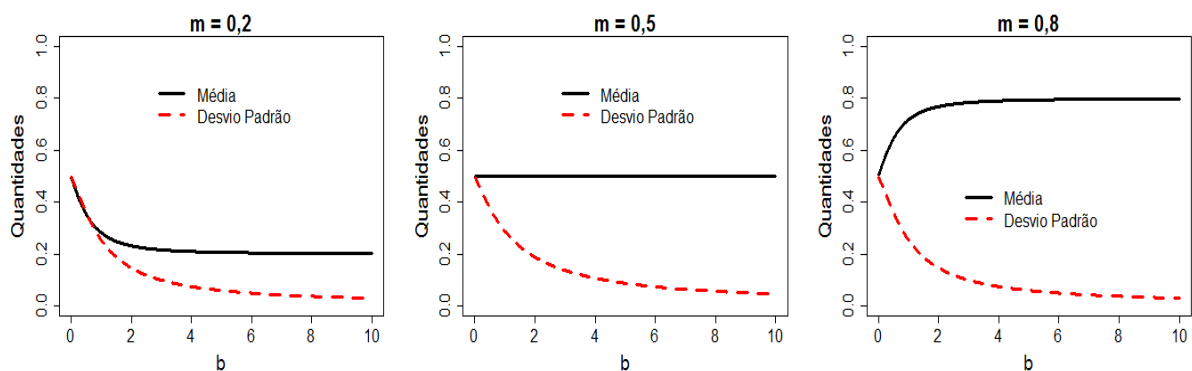
$$E(Y^t) = \int_0^1 y^t \frac{b(1-m)^b m^b y^{b-1} (1-y)^{b-1}}{[(1-m)^b y^b + m^b (1-y)^b]^2} dy \quad (14)$$

e vale apontar que o processo de obtenção da solução analítica para 14 pode conter algum empecilho. Assim, tendo em vista possíveis dificuldades, Paz et al. (2016) sugerem o uso de métodos numéricos e recomendam o uso da expressão

$$\int_0^1 \left[1 + \left(\frac{1-v}{v}\right)^{1/b} \left(\frac{1-m}{m}\right)\right]^t dv, \quad (15)$$

para obter o t -ésimo momento de interesse. A seguir, na Figura 5, apresenta-se o gráfico da média e do desvio padrão como função do parâmetro de dispersão b , para alguns valores do parâmetro de posição m . Vale ressaltar o uso da Equação 15 e da função `integrate` (configuração padrão) alocada no pacote ‘stats’ do R.

Figura 5: Comportamento da Média e Desvio Padrão, segundo alguns valores dos parâmetros da distribuição L-logística.



2.2.4 Assimetria e Curtose

Como visto na Seção 2.2.3, as medidas de assimetria e curtose podem ser calculadas numericamente por meio da Equação 15. No entanto, em Paz et al. (2016), este procedimento não é apresentado, mas comenta-se que a distribuição será simétrica, sempre que $m = 0,5 \forall b \in (0, \infty)$.

Para valores de m diferentes de 0,5, os autores indicam as medidas de assimetria γ_M e γ_p apresentadas em Arnold e Groeneveld (1995) e Hinkley (1975).

A medida γ_M é definida por

$$\gamma_M = 1 - 2F_X(x_0), \quad (16)$$

no qual F_X e x_0 denotam a função de distribuição e a moda da distribuição de interesse, com $\gamma_M \in [-1, 1]$, no qual valores próximos de 1 e -1 indicam forte assimetria a direita e esquerda, respectivamente.

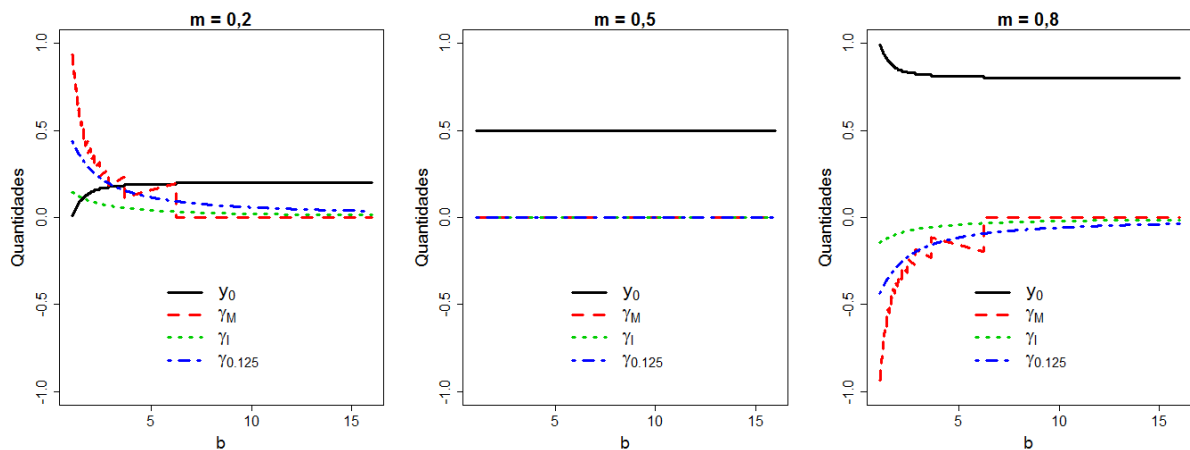
A medida γ_p é definida por

$$\gamma_p = \frac{F_X^{-1}(1-p) + F_X^{-1}(p) - 2F_X^{-1}(0,5)}{F_X^{-1}(1-p) - F_X^{-1}(p)}, \quad (17)$$

no qual F_X^{-1} denota a função quantílica, $p \in (0, \frac{1}{2})$ e $\gamma_p \in [-1, 1]$, no qual valores próximos de 1 e -1 também indicam forte assimetria à direita e à esquerda, respectivamente. A medida também apresenta simetria para $\gamma_p = 0$ e a dependência sobre p pode ser removida através de integração, como visto em Groeneveld e Meeden (1984). Ainda em Paz et al. (2016), também menciona-se a possibilidade de eliminar a influência de p (por meio de integração), mas os autores optam por não realizar este procedimento, adotando $p = 0,125$, decorrente da indicação apresentada em Brys et al. (2003).

A seguir, na Figura 6, apresenta-se o comportamento da moda e das medidas γ_M e $\gamma_{0,125}$, com inclusão da medida γ_I que denota medida γ_p após a remoção da influência de p . Pode-se perceber que para valores para $m = 0,5$ todas as medidas concordam com a simetria da distribuição $LL(m, b)$ e para $m \neq 0,5$ todas as medidas demonstram a mesma tendência. Em seguida, na Figura 7, apresenta-se o comportamento da medida de assimetria baseada nos momentos numéricos, denotada por γ_N . Vale lembrar que todas as medidas são apresentadas para valores do parâmetro de dispersão $b \leq 15$ e alguns valores do parâmetro de posição m . Além disso, note que a medida $\gamma_N \in (-4, 4)$ (para $b \leq 15$) e também fornece indicativo de assimetria que assemelha-se às medidas γ_M , $\gamma_{0,125}$ e γ_I . Também vale lembrar que os valores da moda foram obtidos para $b > 1$.

Figura 6: Comportamento da Moda e das medidas γ_M , $\gamma_{0,125}$ e γ_I , segundo alguns valores dos parâmetros da distribuição L-logística.

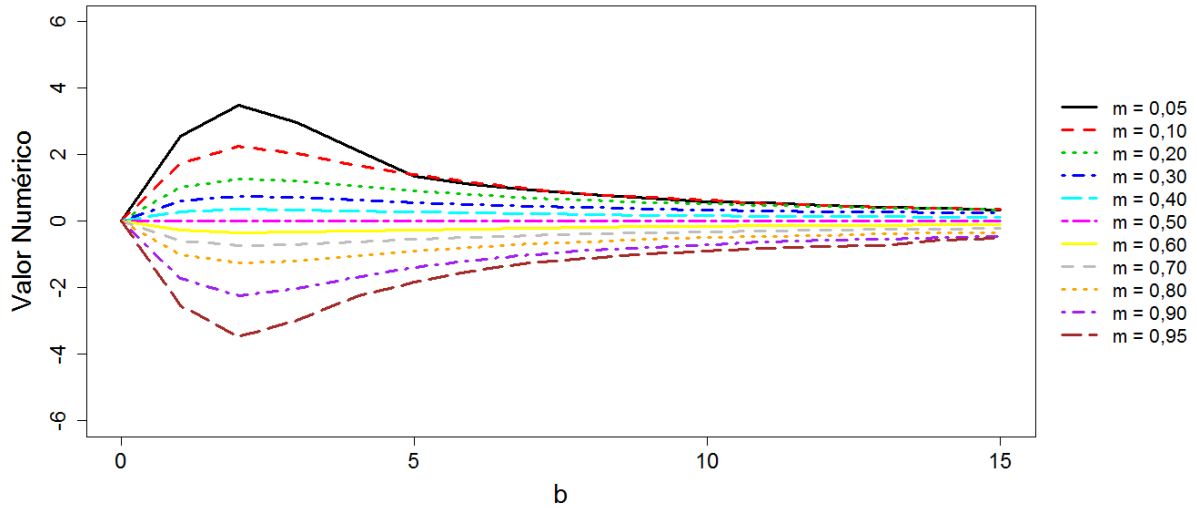


Com relação às medidas de curtose, Paz et al. (2016) sugerem uma medida baseada nos quantis, introduzida inicialmente em Moors (1988), dada por

$$k_q = \frac{F_X^{-1}(7/8) - F_X^{-1}(5/8) + F_X^{-1}(3/8) - F_X^{-1}(1/8)}{F_X^{-1}(6/8) - F_X^{-1}(2/8)}, \quad (18)$$

em que F^{-1} denota a função quantílica. Além disso k_q é uma medida positiva que fornece o grau da concentração dos dados nas caudas da distribuição $LL(m, b)$, em que quanto maior o valor de k_q , maior a

Figura 7: Comportamento da medida γ_N , segundo alguns valores dos parâmetros da distribuição L-logística.



concentração de dados nas caudas. A seguir, na Figura 8, o comportamento da medida k_q . Em seguida, na Figura 9, o comportamento da medida de curtose baseada nos momentos numéricos, denotada por K_N . Vale ressaltar que as medidas são obtidas para valores do parâmetro de dispersão $b \leq 15$ e alguns valores do parâmetro de posição m . Também note que há uma divergência entre as medidas e somente a medida k_q aparenta estar de acordo com o comportamento real da distribuição L-logística, como pode ser visto nas Figuras 1 e 2.

Figura 8: Comportamento da medida k_q , segundo alguns valores dos parâmetros da distribuição L-logística.

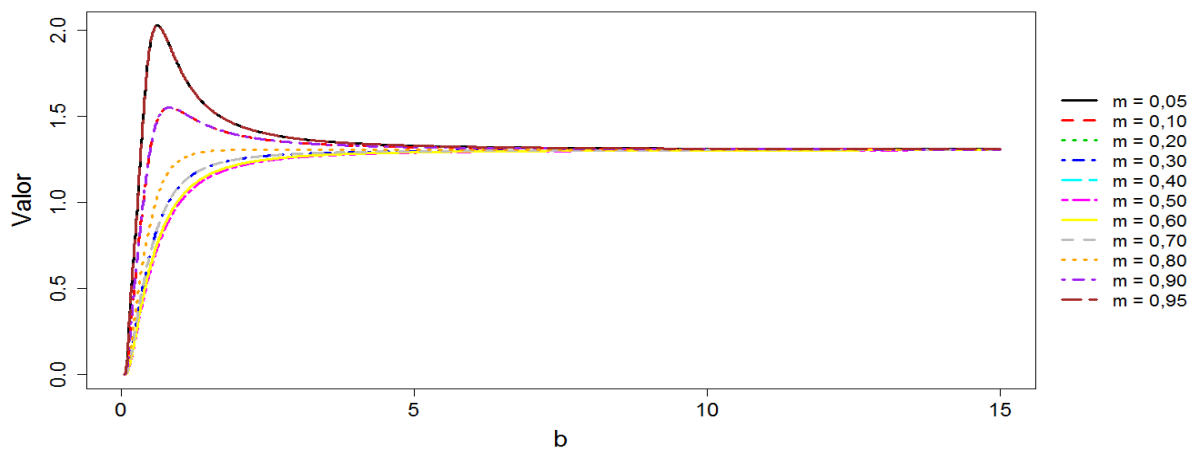
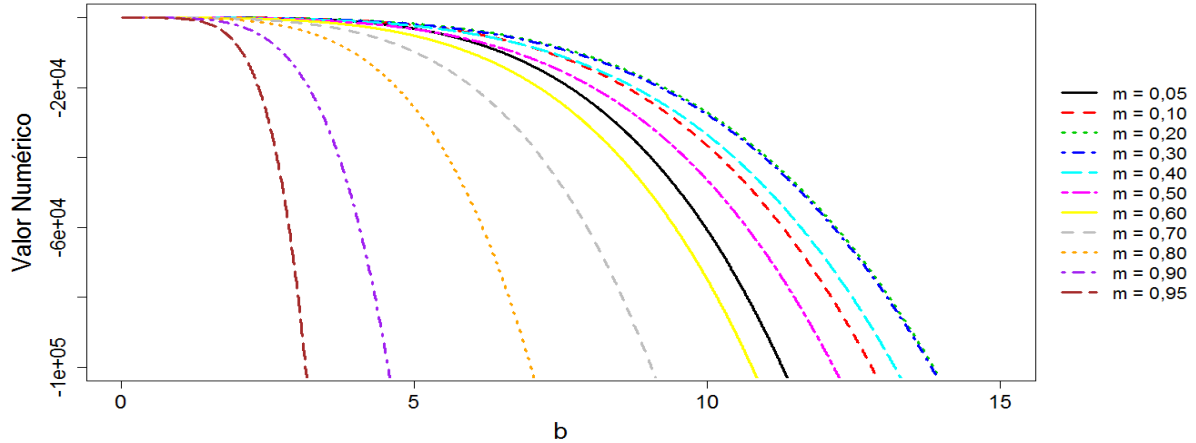


Figura 9: Comportamento da medida K_N , segundo alguns valores dos parâmetros da distribuição L-logística.



2.3 Estrutura de Regressão

A abordagem tradicional empregada no estudo de variáveis com o suporte limitado, segundo um conjunto de variáveis explicativas, é dada pelo modelo de regressão Simplex proposto em Song e Tan (2000) e o popular modelo de regressão Beta proposto em Ferrari e Cribari-Neto (2004). Contudo, como alternativa aos métodos tradicionais, Paz et al. (2016) apresentam o modelo de regressão L-logística. Este modelo diferencia-se da metodologia tradicional dos modelos lineares generalizados (MLG) apresentada em Nelder e Wedderburn (1972), por utilizar uma função da mediana e não utilizar uma distribuição regular da família exponencial linear. O uso da mediana como função de regressão surge como alternativa robusta ao tradicional modelo de regressão linear no que concerne a presença de observações atípicas, como visto em Blattberg e Sargent (1971) e Bassett e Koenker (1978).

Assim como nos MLGs, o modelo de regressão L-logística apresenta as seguintes características:

1. O estudo de uma variável de interesse (variável resposta) Y com o auxílio de um conjunto de variáveis explicativas (covariáveis) x_1, x_2, \dots, x_{p-1} .
2. A realização de um processo de amostragem aleatória no qual obtém-se o vetor resposta observado denotado por $\mathbf{y} = (y_1, y_2, \dots, y_n)^\top$ e a matriz de covariáveis observadas de dimensão $n \times (p-1)$, denotada por $X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{p-1})$.
3. A estruturação do modelo através de três componentes:
 - i. **Componente Aleatório:** representa um vetor aleatório $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)^\top$ que apresenta mediana condicional da forma

$$\text{Md}(\mathbf{Y}|X) = \mathbf{m} = (m_1, m_2, \dots, m_n)^\top; \quad (19)$$

- ii. **Componente Sistemático:** representa a estruturação da matriz de covariáveis observadas X que produz o preditor linear $\boldsymbol{\eta} = (\eta_1, \eta_2, \dots, \eta_n)^\top$, no qual o i -ésimo componente é definido por

$$\eta_i = \mathbf{x}_i \boldsymbol{\beta}, \quad (20)$$

em que \mathbf{x}_i denota a i -ésima linha da matriz X e $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_{p-1})^\top$ denota o vetor de parâmetros desconhecidos do modelo. Caso haja a inclusão do intercepto no modelo, temos $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \dots, \beta_{p-1})^\top$;

- iii. **Função de ligação:** representa qualquer função monótona duplamente diferenciável $g(\cdot)$ que fornece a relação, $\eta_i = g(\mu_i)$, da mediana condicional do componente aleatória com o preditor linear, assim gerando o modelo da forma

$$\boldsymbol{\eta} = g(\mathbf{m}). \quad (21)$$

3 INFERÊNCIA

Este trabalho possui enfoque no paradigma frequentista e baseia suas inferências no método de estimação por máxima verossimilhança. Caso haja interesse no paradigma bayesiano, o leitor pode consultar Paz et al. (2016).

3.1 Estimação por Máxima Verossimilhança

Seja Y_i , com $i = 1, 2, \dots, n$, uma amostra aleatória de tamanho n da distribuição L-logística com vetor de parâmetros desconhecidos $\boldsymbol{\theta} = (m, b)^\top$. Denotando por \mathbf{y} o vetor de observações, a função de verossimilhança é dada por

$$L(\boldsymbol{\theta}; \mathbf{y}) = \prod_{i=1}^n f_Y(y_i; m, b) = \prod_{i=1}^n \frac{b(1-m)^b m^b y_i^{b-1} (1-y_i)^{b-1}}{[(1-m)^b y_i^b + m^b (1-y_i)^b]^2}, \quad (22)$$

com respectiva função de log-verossimilhança

$$\begin{aligned} l(\boldsymbol{\theta}; \mathbf{y}) &= \log(L(\boldsymbol{\theta}; \mathbf{y})) \\ &= \sum_{i=1}^n (\log(b) + b \log(m(1-m)) + (b-1) \log(y_i(1-y_i)) \\ &\quad - 2 \log((1-m)^b y_i^b + m^b (1-y_i)^b)). \end{aligned} \quad (23)$$

Para obter as estimativas de máxima verossimilhança, faz-se necessário resolver o sistema de equações gerado por

$$\nabla l(\boldsymbol{\theta}; \mathbf{y}) = \mathbf{0}, \quad (24)$$

no qual $\nabla l(m, b; \mathbf{y})$ denota o vetor gradiente e suas componentes estão descritas no Apêndice A. Vale apontar que o sistema em 24 é formado por equações não-lineares e sua solução analítica pode não ser de fácil obtenção. No entanto, como argumento para a existência do estimador de máxima verossimilhança para a distribuição L-logística, pode-se apontar a propriedade da equivariância (invariância) e a propriedade que atesta que uma variável aleatória com esta distribuição pode ser obtida através da transformação de uma variável aleatória com distribuição Logística Padrão.

No que diz respeito a forma de obtenção das estimativas, encontra-se na estatística algoritmos tradicionais que que fornecem bons resultados para o método de máxima verossimilhança. Dentre tais algoritmos pode-se citar os algoritmos BFGS, Escore de Fisher e Newton-Raphson. O leitor interessado pode ler sobre estes algoritmos em Nocedal e Wirght (2006) e McCullagh e Nelder (1989), por exemplo. Além disso, convém destacar as seguintes características que estes algoritmos compartilham:

- i) Baseiam-se no uso de derivadas;
- ii) Avaliam uma única solução cadidata por vez do processo de otimização;
- iii) Forma determinística de trabalho (solução inicial z_0^* sempre gera solução final z_n^*);
- iv) Realização de buscas locais (em razão do uso de derivadas).

A popularidade destes algoritmos deve-se a muitos problemas que envolvem funções de log-verossimilhança contínuas e unimodais. Contudo, existem situações que afetam o desempenho destes algoritmos tais como a multimodalidade ou não existência de derivadas da função de log-verossimilhança. Assim, pesquisadores com estas problemáticas podem optar pelo uso de métodos numéricos de derivação, algoritmos livres de derivadas ou metaheurísticas aplicadas à otimização. Este trabalho opta pelo de metaheurísticas, por ser uma abordagem relativamente incomum na área, tendo portando um amplo

campo ainda a ser explorado. Como exemplos de estimação por máxima verossimilhança através de metaheurísticas, cita-se Akman et al. (2011), Ari et al. (2012) e Zhang et al. (2013).

3.2 Ideia sobre Metaheurísticas

Em otimização, metaheurísticas são algoritmos que aplicam estratégias que combinando mecanismos de aprendizagem e adaptação, para resolver problemas de otimização. Elas são construídas para fornecer soluções boas o suficiente, quando não atingirem a solução ótima. A ideia por trás das metaheurísticas, consiste em dispor de um algoritmo prático e eficiente que trabalhe bem em muitos problemas, fornecendo soluções com boa qualidade em curto espaço de tempo. Como visto em Zäpfel et al. (2010, p. 72), o termo metaheurística não possui definição formal, mas descreve uma combinação de conceitos que estrutura heurísticas em um processo iterativo mestre, no qual sofre pequenas e relativas modificações para abranger uma variedade de problemas.

As metaheurísticas também podem ser classificadas de muitas formas. Dentre elas, os algoritmos podem ser considerados como baseado em trajetória ou em população. Como descrito em Yang (2010), as metaheurísticas baseadas em trajetórias são algoritmos de busca que atualizam sua única solução ao longo das iterações e seu mecanismo de busca forma uma trajetória ao longo do espaço de soluções. Em contra partida, as metaheurísticas baseadas em populações são algoritmos de busca que operam com múltiplas soluções ao longo das iterações, apresentando a melhor solução encontrada até o momento. Como descrito em Yang (2014), as metaheurísticas baseadas em populações geralmente apresentam características como: intensificação e diversificação. A diversificação descreve a capacidade de gerar soluções variadas para explorar o espaço de busca em uma escala global (evitando ótimos locais), enquanto que a intensificação descreve a capacidade de focar a busca em determinadas regiões, explorando a informação da existência de boas soluções locais. Os termos intensificação e diversificação também podem ser encontrados na literatura como exploração e exploração respectivamente.

Este trabalho aborda a estimação por máxima verossimilhança, através das metaheurísticas Algoritmo Genético (em inglês: *Genetic Algorithm* - GA) e Evolução Diferencial (em inglês: *Differential Evolution* - DE). Estas metaheurísticas são algoritmos estocásticos baseados em populações e fazem parte da classe de algoritmos evolucionários (*Evolutionary Algorithms*), compreendidos na área de metaheurísticas inspiradas na natureza (*nature-inspired metaheuristics*).

3.3 Algoritmo Genético

O Algoritmo Genético (*Genetic Algorithm*) é uma metodologia que baseia sua estrutura na abstração da teoria da evolução apresentada em Darwin (1859). Foi desenvolvido por John Holland e seus contribuintes na Universidade de Michigan, com propósito de entender o processo adaptativo dos sistemas naturais. Sua formalização foi apresentada em Holland (1975) e sua visibilidade foi ampliada após o trabalho de Goldberg (1989).

Hoje existem vários estudos, expansões e variações da metodologia original, constituindo uma classe própria denotada por Algoritmos Genéticos (*Genetic Algorithms*). O método original é conhecido como método clássico ou canônico, mas também pode ser encontrado com a notação Algoritmo Genético Simples (*Simple Genetic Algorithm*), segundo Srinivas e Patnaik (1995). Algumas nomenclaturas são comuns na classe de Algoritmos Genéticos, sendo elas: indivíduo, gene, população, geração, função objetivo e aptidão. Contudo, geralmente estes termos são encontrados de forma ambígua na literatura. Assim, por simplicidade, os termos serão usados com as seguintes descrições:

- **indivíduo**, descreve qualquer vetor solução (no formato codificado) do problema de interesse;
- **gene**, descreve qualquer variável codificada (elemento do vetor) que compõe o indivíduo;

- **população**, descreve o conjunto de indivíduos dentro do algoritmo;
- **geração**, descreve uma população após modificações evolutivas em uma determinada iteração do algoritmo;
- **função objetivo**, descreve a função que deve ser otimizada no problema;
- **aptidão**, descreve uma medida de qualidade (geralmente baseada na função objetivo) para um particular indivíduo.

Embora existam várias versões da metodologia, como visto em Michalewicz (1996), geralmente os algoritmos apresentam as seguintes características:

1. Um tipo de codificação (representação) para as soluções do problema;
2. Uma forma de criar um conjunto inicial de soluções, onde geralmente denota-se este conjunto por população inicial;
3. Um mecanismo, com base na função a ser otimizada, que avalia as soluções e gera notas (aptidões);
4. Hiperparâmetros para o funcionamento do algoritmo, geralmente denotados por parâmetros genéticos;
5. Heurísticas secundárias, denotadas por operadores genéticos (*Genetic Operators*), que modificam a população durante o decorrer do processamento.

Ao longo de seu processamento, o GA estrutura estas características e trabalha fazendo modificações na população inicial (emulando o processo evolutivo) até atingir a convergência global ou proporcionar boas soluções. A população inicial tem importância fundamental no processo de obtenção do ótimo e normalmente é construída de forma aleatória para diversificar ao máximo os indivíduos e proporcionar a presença de características que podem não surgir durante o processamento do algoritmo, assim prejudicando a busca pela solução ótima. Como apontado em Silva (2008), a diversidade da população inicial é de grande importância, pois características não existentes na população inicial, dificilmente aparecem nas próximas gerações durante o processo evolutivo. Desse modo, o tamanho da população inicial (denotado por N_p) é tido como um hiperparâmetro fundamental, pois o mesmo influencia fortemente o desempenho e o comportamento dos Algoritmos Genéticos, sendo necessário calibrá-lo para que haja uma boa representação do espaço de busca e reduza a possibilidade de convergência prematura em ótimos locais nas próximas gerações. Como visto em Gen e Cheng (2000), após a construção da população inicial, os indivíduos são avaliados e recebem sua aptidão, no qual os mais aptos (melhores soluções) possuem maiores chances de gerarem novos indivíduos denotados por filhos ou proles (*Offspring*). Há dois tipos de mecanismos (operadores genéticos) que geram diversidade na população, são eles a mutação (*Mutation Operator*) e a recombinação (*Crossover Operator*). A mutação promove modificações nos genes de um único indivíduo. A recombinação cria novos indivíduos combinando as informações dos indivíduos denotados por pais (*Parents*). Em seguida, os filhos recebem sua aptidão e uma nova população é formada pelos indivíduos mais aptos selecionados entre pais e filhos. Por fim, após muitas gerações, o algoritmo converge para boas soluções (bons indivíduos), onde espera-se que seja a solução ótima do problema ou a melhor solução possível. Vale lembrar que o mecanismo de escolha dos indivíduos mais aptos é fornecido pelo operador genético de seleção (*Selection Operator*). O leitor interessado, pode ler o pseudocódigo da metaheurística em Srinivas e Patnaik (1994, p. 19). Além disso, como apontado em Haupt e Haupt (2004), os Algoritmos Genéticos apresentam bom desempenho em problemas complexos de otimização discreta e contínua; não necessitam da informação gerada pela derivada da função objetivo; realizam buscas simultâneas fornecendo uma lista de boas soluções (ao contrário dos algoritmos BFGS, Escala de Fisher e Newton-Raphson); lidam de maneira satisfatória com muitas variáveis e trabalham bem com computação paralela.

A ampla aplicabilidade do GA em diferentes tipos de problemas de otimização, também deve-se ao tipo de codificação empregada no algoritmo. Por muito tempo a codificação das soluções foi alvo de debate na classe dos GAs e várias formas de codificações surgiram na literatura para proporcionar implementações eficientes, em que o tipo de problema abordado determina o tipo de codificação empregada e o tipo de estrutura do algoritmo. Como visto em Gen e Cheng (2000), existem algumas classificações para os métodos de codificação empregados no GA, contudo para obtenção das estimativas de máxima verossimilhança, convém comentar somente sobre os métodos binário e real. O método binário foi inicialmente proposto em Holland (1975) e consiste na transformação das soluções em números binários, tornando o vetor de soluções uma cadeia binária semelhante ao DNA, fazendo alusão ao processo evolutivo. A codificação binária possui fácil implementação e geralmente encontra-se nos pacotes computacionais, sendo bastante utilizada ainda. No entanto, a precisão das soluções é limitada pela quantidade de bits empregados na codificação e em problemas contínuos este método é pouco atrativo, pois também pode não preservar a localidade das soluções durante o processo de decodificação, acarretando valores que não sejam ótimos no espaço real da função de interesse. Por sua vez, como comentado em Man et al. (1996), temos o método real que foi proposto para lidar com problemas de otimização com soluções contínuas. Este método codifica as soluções do problema utilizando formato digital dos números (ponto flutuante) usados nos computadores, o que torna o processo mais rápido e preciso, como apresentado em Janikow e Michalewicz (1991). Muitos estudos foram realizados para comprovar a eficiência da codificação real, onde pode-se destacar Goldberg (1991) que fundamenta a convergência dos algoritmos genéticos com codificação real (real-encoded genetic algoritmos) e Eshelman e Schaffer (1992) que comparam as duas codificações, demonstrando a vantagem da codificação real em problemas contínuos. Este trabalho adota o método real como codificação para o GA empregado na obtenção das estimativas de máxima verossimilhança.

3.3.1 Operador de Seleção

A seleção é um mecanismo que fornece às melhores soluções (indivíduos mais aptos) a maior probabilidade de contribuir na construção de uma ou mais novas soluções (filhos) para a próxima geração, sendo este o operador responsável pelo fato de que em média, as soluções irão sempre melhorar a cada geração. Como visto em Zäpfel et al. (2010), o operador seleção efetua suas decisões baseando-se na aptidão dos indivíduos e realizando comparações entre os indivíduos da população corrente, afim de selecionar os melhores indivíduos para gerar novos filhos. Segundo Goldberg (1989), o operador de seleção equilibra a busca pelo ótimo fornecendo uma troca entre a diversificação e a exploração. Essa troca decorre da ideia de que boas soluções possuem mais chance de contribuir para o encontro de soluções ainda melhores no espaço de busca, emulando de forma artificial o processo de seleção natural em que os melhores indivíduos terão chance de sobreviver e reproduzir em seu ambiente. Além disso, como comentado em Gen e Cheng (2000), este operador tem papel fundamental na convergência do algoritmo genético, pois a má construção desta heurística pode desequilibrar o processo de intensificação e diversificação, o que ocasiona convergência em ótimo local (tendência em intensificação) ou lentidão no processo de busca (tendência em diversificação), assim sendo recomendado que a seleção ajude o algoritmo a explorar mais no início do processo e intensificar mais no final do processo.

Depois de decidir a codificação, o segundo passo é decidir qual o operador de seleção será usado no algoritmo genético. Normalmente, para a codificação real, encontra-se na literatura os métodos de seleção por roleta e torneio. O método de seleção por roleta (*roulette wheel selection*), também conhecido como seleção proporcionada por aptidão (*fitness-proportional selection*), é um dos métodos mais conhecidos e foi proposto inicialmente em Holland (1975). Neste método, a probabilidade de seleção determina o número de vezes que um indivíduo pode ser selecionado para gerar filhos e é dada pela

aptidão do indivíduo dividida pela soma das aptidões da população corrente. O termo roleta surge da analogia com jogo de roleta, onde imagina-se a população corrente na forma de uma roleta cujo os espaços são formados pela aptidão de cada indivíduo. Assim, a chance de que um indivíduo seja selecionado é proporcional ao seu espaço na roleta. Contudo, como apontado em Bäck (1996), o método da roleta deve ser construído de modo que não venha a falhar se algum indivíduo apresentar aptidão negativa. Por sua vez, como comentado em Goldberg (1989), o método de seleção por torneio (*Tournament Selection*) consiste na seleção aleatória de r indivíduos da população corrente, geralmente $r = 2$, onde em seguida o indivíduo mais apto é selecionado. Este procedimento é efetuado até que haja um número suficiente de indivíduos para que se aplique os operadores de recombinação e mutação. O método de seleção por torneio é fácil de ser implementado e tem baixa complexidade, contudo, existe a possibilidade do método não selecionar os indivíduos mais aptos da população corrente, o que pode tornar o processo de convergência do algoritmo mais demorado. Vale ressaltar que o empecilho gerado pela seleção por torneio pode ser amenizado através de uma sub-heurística denotada por elitismo. O elitismo foi inicialmente proposto por De Jong (1975) e pode ser visto como um apêndice do método de seleção. O método consiste em conservar um pequeno grupo contendo os melhores indivíduos encontrados a cada geração, onde este grupo é utilizado para produzir novos filhos, compondo o grupo de indivíduos formado pelo processo de seleção e acelerando o processamento do GA tornando-o mais rápido. No entanto, como visto em Coley (1999), embora o elitismo melhore a performance do GA, ele deve ser implementado com precaução, pois o mesmo pode acarretar em convergência em ótimos locais. O leitor interessado em métodos alternativos de seleção pode consultar Brindle (1981) que examina a performance do GA para 6 métodos de seleção.

3.3.2 Operador de Recombinação

Dentro dos Algoritmos Genéticos, a recombinação é um mecanismo que tenta emular o processo de reprodução de organismos biológicos, combinando as informações dos indivíduos (pais), obtidos no processo de seleção, para formar novos indivíduos (filhos). Como visto em Bäck (1996), este mecanismo se baseia na ideia de que boas soluções (indivíduos aptos) podem ser combinadas para favorecer o surgimento de soluções ainda melhores (filhos mais aptos).

O operador de recombinação proporciona aos algoritmos genéticos uma boa parcela de sua eficiência. Ele é considerado como o operador genético predominante e é aplicado com uma determinada probabilidade aos indivíduos obtidos pelo operador de seleção. Esta probabilidade é um hiperparâmetro do algoritmo que geralmente é denotado por probabilidade (ou taxa) de recombinação (ou cruzamento) e como sugerido em Silva (2008), deve variar entre 0,6 e 0,9.

A recombinação pode ser implementada por meio da geração de números aleatórios no intervalo unitário, onde o operador será aplicado aos indivíduos se o número gerado for menor ou igual a probabilidade de recombinação. Não havendo recombinação, os filhos serão cópias dos pais, assim preservando algumas soluções. Contudo, como comentado em Silva (2008), embora a recombinação seja um mecanismo que favorece a busca e reduz as chances de convergência prematura em ótimos locais, altos valores da probabilidade de recombinação acarretarão em custo computacional devido à exploração de regiões não promissoras dentro do espaço de busca.

Diversos métodos de recombinação surgiram na literatura durante as últimas décadas, o que ocasionou a classificação do tipo de recombinação com base na quantidade de indivíduos envolvidos. Como visto em Engelbrecht (2007), as principais classes são descritas por:

- **Recombinação Assexual**, onde um novo indivíduo é gerado por somente um indivíduo;
- **Recombinação Sexual**, onde dois indivíduos contribuem na geração de um ou mais novos indivíduos;

- **Multi-Recombinação**, onde mais de dois indivíduos contribuem na geração de um ou mais novos indivíduos.

Os operadores iniciais de recombinação foram propostos na literatura com base na codificação binária das soluções e são operadores pertencentes à classe de recombinação Sexual, fazendo alusão a interação dos organismos biológicos. O primeiro operador de recombinação foi proposto em Holland (1975) e é denotado por operador de cruzamento de um único ponto (*one-point crossover operator*). Outros operadores populares e facilmente encontrados nos pacotes computacionais são o operador de cruzamento de dois pontos (*two-point crossover operator*) e o operador de cruzamento uniforme (*uniform crossover operator*), onde o leitor interessado na construção destes operadores, pode consultar Engelbrecht (2007, p. 145). Posteriormente, visando problemas de otimização com soluções contínuas, surgiram os métodos de recombinação com base na codificação real, dos quais cita-se os populares métodos de cruzamento aritmético (*arithmetical crossover operator*) e o método de cruzamento por mistura (*blend crossover operator*). O método de cruzamento aritmético consiste em gerar novos indivíduos através de uma combinação linear convexa entre 2 ou mais indivíduos, tendo como caso especial o método de cruzamento médio (*average crossover operator*) que consiste em gerar novos indivíduos por meio da média aritmética. Por sua vez, o método de cruzamento por mistura (também conhecido por BLX- α) foi proposto em Eshelman e Schaffer (1993) e consiste em gerar novos filhos de forma aleatória, dentro do espaço gerado por dois indivíduos envolvidos no processo de recombinação. O leitor interessado na construção de operadores de recombinação para codificação real, pode consultar Gen e Cheng (2000, p. 27-33).

3.3.3 Operador de Mutação

Na classe de Algoritmos Genéticos, o operador de mutação surge como mecanismo para gerar diversidade, assemelhando-se neste propósito ao operador de recombinação. A primeira proposta de mutação foi apresentada como operador de suporte em Holland (1975) e foi inspirada na mutação do DNA de um organismo vivo e seus impactos na diversidade evolutiva. Segundo Goldberg (1989), embora a seleção e a recombinação forneçam uma estrutura eficiente de busca pelo ótimo, ocasionalmente esta estrutura pode vir a perder (não explorar) alguns locais interessantes e com isso atrapalhar a convergência global do algoritmo. Assim a mutação tem a finalidade de prevenir tais perdas e realiza este procedimento diversificando (aleatoriamente) alguns indivíduos para que a busca não permaneça fixada em áreas particulares, ocasionando maior varredura e evitando possíveis perdas de informação durante a recombinação (balanço entre intensificação e diversificação). Portanto, utiliza-se a mutação para que qualquer ponto no espaço de busca não possua probabilidade nula, permitindo que novas soluções sejam exploradas e aumentando a probabilidade de obtenção do ótimo global.

No passado, como visto em Melanie (1996), a recombinação era tida como o maior instrumento de variação e inovação dentro dos algoritmos genéticos, enquanto que a mutação somente assegurava a população contra ótimos locais (ideia de suporte). Contudo, este pensamento foi deixado de lado à medida que estudos comparativos surgiam na literatura para verificar a força da mutação com relação a recombinação. Dentre estes estudos, cita-se Mülenbein (1992) que demonstra como o poder da mutação foi subestimado nos algoritmos genéticos tradicionais e Spears (1993) que discute os benefícios relativos da mutação e da recombinação. Ainda em Melanie (1996), aponta-se que os dois operadores devem ser incluídos da forma que favoreça a obtenção do ótimo.

Assim como na recombinação, a mutação ocorre no algoritmo com uma determinada frequência, resultante do hiperparâmetro denotado por probabilidade (taxa) de mutação. A probabilidade de mutação geralmente é ajustada para valores baixos tais como valores entre 0,00001 e 0,1 ou valores entre 0,005 e 0,01, como sugerido em Silva (2008) e Schaffer et al. (1989), respectivamente. Ainda em Schaffer et al. (1989), comenta-se que a probabilidade de mutação geralmente possui relação inversa com o tamanho

da população empregada no GA, o que reflete provavelmente o fato de que aumentar qualquer um dos hiperparâmetros promove o aumento da exploração. Portanto, estas quantidades devem ser calibradas de modo que o nível de busca mantenha-se mais ou menos constante, pois grandes (ou pequenas) populações combinadas com altas (ou pequenas) probabilidades de mutação atrapalham a performance do algoritmo, como comentado por Bäck (1996, p. 113). Além disso, uma probabilidade de mutação muito baixa poderá comprometer a diversidade das soluções ou tornar o algoritmo dependente somente do operador de recombinação, enquanto que uma probabilidade de mutação muito alta poderá comprometer a convergência do algoritmo por excesso de diversidade (perda de semelhança entre pais e filhos).

A implementação da mutação também pode ser realizada por meio da geração de números aleatórios no intervalo $[0, 1]$, no qual o operador será aplicado a um gene de um indivíduo se o número gerado for menor ou igual a probabilidade de mutação. Não havendo mutação, o indivíduo permanecerá o mesmo. Diferente da recombinação, o operador de mutação é aplicado em cada gene e a probabilidade de que um indivíduo venha a sofrer mutação é dada por $1 - (1 - p_m)^{n_s}$, onde p_m e n_s descrevem respectivamente a probabilidade de mutação e o número de genes que compõe o indivíduo no algoritmo. A metodologia mais antiga de mutação foi apresentada em Holland (1975) para a codificação binária e é conhecida por mutação aleatória (*random mutation*). Esta metodologia é bastante popular por sua facilidade de implementação, o que acarreta fácil acesso em pacotes computacionais. O leitor interessado na construção desta mutação, pode ler o pseudocódigo em Engelbrecht (2007, p. 155). Em contrapartida, para a codificação real, algumas mutações também são bastante populares, das quais cita-se a mutação uniforme (*uniform mutation*) e a mutação gaussiana (*gaussian mutation*) que foi proposta em Hinterding (1995). A mutação uniforme consiste em substituir cada gene selecionado por um ponto aleatório em seus respectivos limites no espaço de busca. Por sua vez, a mutação gaussiana consiste em adicionar aos genes selecionados, um ruído normal com média 0 e desvio padrão σ , onde o desvio padrão torna-se mais um hiperparâmetro do algoritmo. O leitor interessado em outros métodos de mutação, consultar Gen e Cheng (2000, p. 33-34).

3.4 Evolução Diferencial

O algoritmo de Evolução Diferencial (DE) foi formalizado em Storn e Price (1995) e surgiu na literatura como uma poderosa ferramenta construída para lidar com problemas de otimização contínua. Posteriormente, como comentado em Storn e Price (1996), a metaheurística demonstrou sua eficiência na primeira competição internacional sobre otimização evolucionária, obtendo o terceiro lugar. Embora situado na categoria de algoritmos evolucionários, o DE surgiu de argumentos naturalmente matemáticos (geométricos) e não possui forte inspiração na evolução natural dos indivíduos. O nome Evolução Diferencial deve-se ao fato de que o algoritmo evolui sua população por meio de operações de diferença entre vetores, tornando o processo evolutivo puramente artificial. Como visto em Storn e Price (1997), o DE foi desenvolvido para lidar com funções não diferenciáveis, não lineares e multimodais, também sendo projetado para ter boas propriedades de convergências, bom desempenho em computação paralela e estrutura simplificada com poucos hiperparâmetros.

A estratégia de busca do DE assemelha-se ao método empregado no GA, sendo estruturada por meio dos hiperparâmetros N_p e P_r que denotam respectivamente a quantidade de vetores usado na busca (alusão ao tamanho da população inicial) e a probabilidade de recombinação. Além disso, o algoritmo utiliza as mesmas nomenclaturas do GA. Entretanto, o termo gene deixa de ser usado e o termo indivíduo passa a denotar diretamente um vetor real, dispensando o uso de codificações. A implementação do DE é feita por meio de N_p vetores D -dimensionais, denotados por $\mathbf{x}_{i,G}$, no qual D descreve o número de variáveis que compõe um vetor solução (alusão a quantidade de genes em um indivíduo), G descreve a geração em que encontra-se o indivíduo e $i = 1, 2, \dots, N_p$. Os vetores são construídos inicialmente de

forma aleatória, por meio da distribuição uniforme, para abranger todo espaço de busca. Em casos em que haja uma solução já disponível, a população inicial pode ser gerada pela adição de ruídos normais com média 0 e desvio padrão σ , como visto em Storn e Price (1997). Em seguida, o algoritmo gera novos vetores (indivíduos) adicionando a diferença ponderada entre dois vetores da população a um terceiro vetor (vetor base), onde este processo recebe o nome de mutação e os novos vetores são denotados por indivíduos ou vetores mutantes (*mutated vectors*). Posteriormente, ocorre o processo de recombinação no qual os vetores mutantes são combinados com outros vetores pré-determinados, chamados de vetores alvo (target vectors), para construir os vetores de teste (*trial vectors*). Por fim, ocorre o processo de seleção em que se um vetor de teste apresentar aptidão superior ao do seu respectivo vetor alvo, este irá substituí-lo na população corrente, assim formando a nova geração no final do processo. Vale ressaltar que cada vetor na população corrente deve servir como vetor alvo de maneira que ocorra N_p comparações para compor a nova geração. O leitor interessado, pode ver o pseudocódigo do algoritmo em Engelbrecht (2007, p. 241).

3.4.1 Operador de Mutação

Diferentemente do GA, o DE não necessita da probabilidade de mutação e utiliza este mecanismo como principal fonte de busca pelo ótimo. O operador de mutação é geralmente denotado por mutação diferencial e sua versão mais comum, apresentada inicialmente em Storn e Price (1995), consiste na construção de um vetor mutante $\mathbf{v}_{i,G}$ para cada vetor alvo $\mathbf{x}_{i,G}$, da forma

$$\mathbf{v}_{i,G} = \mathbf{x}_{r_1,G} + F(\mathbf{x}_{r_2,G} - \mathbf{x}_{r_3,G}), \quad (25)$$

em que $r_1, r_2, r_3 \in \{1, 2, \dots, i-1, i+1, \dots, N_p\}$ são índices aleatórios mutuamente diferentes. O vetor $\mathbf{x}_{r_1,G}$, embora selecionado aleatoriamente, recebe a nomenclatura especial de vetor base (ou vetor a receber a mutação) e a quantidade $F(\mathbf{x}_{r_2,G} - \mathbf{x}_{r_3,G})$ define a direção e o tamanho do passo que será efetuado na busca. A constante real (não negativa) F denota um hiperparâmetro de controle que gerencia a troca entre a exploração e a exploração no espaço de busca. Para maiores detalhes sobre esta estrutura de mutação, ver Feoktistov (2006, pags 17-20). O leitor interessado em valores interessantes para o hiperparâmetro, pode adotar $F \in [0, 2]$ como apresentado em Storn e Price (1997) ou ler mais sobre os efeitos do ajuste de F em Price et al. (2005, pags. 75-79).

3.4.2 Operador de Recombinação

A recombinação surge no DE como forma de aumentar a diversificação das soluções, assemelhando-se ao papel da mutação nos algoritmos genéticos. Este operador tem sua versão mais famosa denotada por recombinação binomial e apresentada em Storn e Price (1995). A recombinação Binomial pode ser implementado por meio da geração de números aleatórios no intervalo $[0, 1]$, no qual o operador será aplicado elemento a elemento do vetor alvo $\mathbf{x}_{i,G}$, para produzir o vetor de teste $\mathbf{u}_{i,G}$. Assim, o mecanismo de construção do vetor de teste é dado por

$$u_{ij,G} = \begin{cases} v_{ij,G}, & \text{se } R_j \leq P_r \text{ ou } j = I_i \\ x_{ij,G}, & \text{se } R_j > P_r \text{ e } j \neq I_i \end{cases}, \quad (26)$$

em que $P_r \in [0, 1]$ denota a probabilidade de recombinação, $R_j \sim U(0, 1)$, $j = 1, 2, \dots, D$ e $u_{ij,G}$ denota o j -ésimo elemento do vetor de teste $\mathbf{u}_{i,G}$. Por sua vez, a quantidade I_i denota um índice escolhido aleatoriamente no conjunto $\{1, 2, \dots, D\}$ de modo a garantir que o vetor de teste $\mathbf{u}_{i,G}$ receba pelo menos uma componente do vetor mutante $\mathbf{v}_{i,G}$, prevenindo a ocorrência de vetores de teste iguais aos vetores alvos.

3.4.3 Operador de Seleção

O operador de seleção surge como última etapa do DE e tem o objetivo de fornecer os melhores indivíduos para compor a nova geração. A primeira proposta de seleção, formulada em Storn e Price (1995), foi desenvolvida segundo o critério ganancioso/guloso (*greeding criterion*) que consiste em comparar cada vetor de teste com seu respectivo vetor alvo e selecionar aquele que apresentar melhor aptidão. Este tipo de seleção também é conhecida como seleção determinística e tem a vantagem de garantir que a aptidão média da população sempre seja melhorada, como visto em Engelbrecht (2007).

3.4.4 Outras variações

A estrutura original do DE (algoritmo canônico) recebeu variações ao longo do tempo, como comentado em Storn e Price (1997). Assim, afim de classificar as muitas variações de DE, surgiu a notação DE/ $x/y/z$, em que x , y , z especificam respectivamente o método de seleção do vetor base $\mathbf{x}_{r_1,G}$, o número de vetores de diferença usados na mutação e o tipo de estratégia de recombinação utilizada. Assim, configurando o DE canônico da forma DE/rand/1/bin, pelo método de seleção do vetor base ser canônico, por usar somente um vetor de diferença e aplicar a mutação binomial. Das outras versões de DE encontrados na literatura, cita-se

- **DE/best/2/bin**: Nesta configuração, sugerida em Price (1996), o DE apresenta operador de mutação da forma

$$\mathbf{v}_{i,G} = \mathbf{x}_{best,G} + F(\mathbf{x}_{r_1,G} + \mathbf{x}_{r_2,G} - \mathbf{x}_{r_3,G} - \mathbf{x}_{r_4,G}), \quad (27)$$

no qual o vetor base $\mathbf{x}_{best,G}$ denota a melhor solução encontrada na população corrente e os demais vetores possuem seus índices aleatórios mutuamente diferentes, com $r_1, r_2, r_3, r_4 \in \{1, 2, \dots, i-1, i+1, \dots, N_p\}$;

- **DE/rand-to-best/1/bin**: Nesta configuração, sugerida em Storn (1996), o DE apresenta operador de mutação da forma

$$\mathbf{v}_{i,G} = \mathbf{x}_{i,G} + L(\mathbf{x}_{best,G} - \mathbf{x}_{i,G}) + F(\mathbf{x}_{r_1,G} - \mathbf{x}_{r_2,G}), \quad (28)$$

em que vetor base $\mathbf{x}_{i,G}$ denota a o próprio vetor alvo, o vetor $\mathbf{x}_{best,G}$ denota a melhor solução encontrada na população corrente, o hiperparâmetro L é uma constante real (não negativa) que controla o passo em que o vetor base dá em direção ao vetor $\mathbf{x}_{best,G}$ e os demais vetores possuem seus índices aleatórios mutuamente diferentes, com $r_1, r_2 \in \{1, 2, \dots, i-1, i+1, \dots, N_p\}$.

O leitor interessado em mais variações, pode consultar Storn (1996) e Engelbrecht (2007).

3.5 Erro de Estimação

As metodologias científicas utilizadas para descrever problemas do mundo real são construídas sobre um número de suposições e técnicas. O fato de que determinados problemas não possam ser medidos com absoluta precisão, faz com que haja a necessidade de avaliar a variabilidade associada ao processo, de modo a garantir a eficiência da técnica ou suposição envolvida na descrição do problema.

Esta variação não explicada pela metodologia utilizada surge de maneira aleatória na natureza e faz com que a repetição de um determinado problema acarrete valores diferentes de medição, impossibilitando a obtenção direta do verdadeiro valor da característica de interesse (parâmetro). Na estatística, quando estamos interessados em modelar um determinado conjunto de dados, esta variação surge com a nomenclatura de erro de estimação. Este erro é um componente fundamental do processo

de inferência estatística e ocorre como estimativa da incerteza associada ao processo de estimação, possibilitando com que um pesquisador avalie a qualidade da estimativa com respeito a discrepâncias com estimativas ou previsões anteriores.

Os métodos tradicionais de estimação possuem sua própria forma de fornecer o erro associado à estimativa. Contudo, para casos em que existe alta complexidade no estimador ou na distribuição de interesse, a obtenção do erro de estimação tornar-se uma tarefa bastante complicada. O uso de metaheurísticas para o ajuste de qualquer distribuição pode ser caracterizado como um caso de estimadores complexos que dificultam a obtenção do erro de estimação. Entretanto, este problema pode ser resolvido por meio do uso de técnicas de reamostragem, como visto em Akman et al. (2011).

Os métodos de reamostragem são técnicas que permitem a estimação da variabilidade (variância, desvio padrão ou viés). Estas técnicas possuem forte apelo teórico e apresentam a característica de reamostrar conjuntos de dados para esboçar o comportamento das estimativas pontuais. Vale apontar que embora estejam consolidados na literatura, os métodos de reamostragem as vezes são vistos como procedimentos artificiais. Dentre as técnicas mais conhecidas, cita-se o método Bootstrap proposto em Efron (1979) e o método Jackknife proposto inicialmente em Quenouille (1949).

3.6 Método Jackknife

O método Jackknife foi introduzido inicialmente em Quenouille (1949) e consiste na proposta de reduzir o viés de um estimador de interesse. Em seguida, em Quenouille (1956), a ideia foi generalizada para várias partições e sua aplicabilidade foi expandida. Posteriormente, em Tukey (1958), a metodologia foi organizada para estimar o erro padrão e algum tempo depois o nome Jackknife foi dado à técnica em um trabalho não publicado de Tukey. Desde então estudos foram realizados e alguns deles podem ser vistos em Efron e Tibshirani (1993, pags. 149-150), por exemplo.

3.6.1 Apresentação do método

Seja $\mathbf{w} = (w_1, w_2, \dots, w_n)^\top$ uma amostra aleatória observada, $s(\cdot)$ uma função qualquer de interesse e $\hat{\theta} = s(\mathbf{w})$ um estimador para o parâmetro de interesse θ . O método Jackknife consiste na construção de n amostras, denotadas por amostras Jackknife, da forma

$$\mathbf{w}_{(i)} = (w_1, w_2, \dots, w_{i-1}, w_{i+1}, \dots, w_n)^\top, \quad (29)$$

no qual $i = 1, 2, \dots, n$ e $\mathbf{w}_{(i)}$ denota a i -ésima amostra Jackknife que consiste na amostra original com a i -ésima observação removida. Em seguida, assumindo $\hat{\theta}_{(i)} = s(\mathbf{w}_{(i)})$ como a i -ésima estimativa Jackknife para θ , temos a estimativa Jackknife do viés definida por

$$\hat{B}_{jk} = (n-1)(\hat{\theta}_{(\cdot)} - \hat{\theta}), \quad (30)$$

em que $\hat{\theta}_{(\cdot)}$ denota a média aritmética das n estimativas Jackknife. Agora, de posse de \hat{B}_{jk} , pode-se melhorar o estimador $\hat{\theta}$ construindo-se o estimador Jackknife da forma

$$\bar{\theta}_{jk} = n\hat{\theta} - (n-1)\hat{\theta}_{(\cdot)}. \quad (31)$$

Em Tukey (1958), indica-se que as estimativas Jackknife também podem proporcionar uma estimativa não paramétrica da variância de $\bar{\theta}_{jk}$. Para isso, Tukey apresenta uma estrutura chamada de pseudo-valores que são calculados da forma

$$\tilde{\theta}_{(i)} = n\hat{\theta} - (n-1)\hat{\theta}_{(i)}, \quad (32)$$

no qual a média aritmética dos pseudo-valores $\tilde{\theta}_{(\cdot)}$ fornece o estimador Jackknife ($\tilde{\theta}_{(\cdot)} = \bar{\theta}_{jk}$). Assim, com base no estimador para a variância da média amostral, o estimador da variância de $\bar{\theta}_{jk}$ é dado por

$$\widehat{\text{VAR}}_{jk} = \frac{\sum_{i=1}^n (\tilde{\theta}_{(i)} - \tilde{\theta}_{(\cdot)})^2}{n(n-1)}, \quad (33)$$

com respectivo desvio padrão da forma $\widehat{\text{DP}}_{jk} = \sqrt{\widehat{\text{VAR}}_{jk}}$.

3.6.2 Intervalo de Confiança

Como comentado em Tukey (1958), os pseudo-valores 32 podem ser tratados como variáveis aleatórias independentes e identicamente distribuídas, possibilitando a construção do intervalo de confiança para θ da forma

$$\bar{\theta}_{jk} \pm t_{n-1}^{(1-\alpha)} \widehat{\text{DP}}_{jk}, \quad (34)$$

em que $t_{n-1}^{(1-\alpha)}$ denota o quantil $1 - \alpha$ da distribuição t de Student com $n - 1$ graus de liberdade e $\alpha \in (0, 1)$. Além disso, como visto em Miller (1974) e Efron (1982), vale apontar que este intervalo tem caráter assintótico. Ainda em Efron e Tibshirani (1993, p. 145) comenta-se que este intervalo não é significativamente melhor que um intervalo baseado na distribuição normal. Além disso, embora o Jackknife seja fácil e prático, a sua metodologia pode vir a falhar com estimadores robustos como a mediana. Assim, para contornar esta problemática geralmente remove-se d observações durante a construção das amostras Jackknife. Este procedimento foi proposto em Shao e Wu (1989) e recebe o nome de método Jackknife com d observações removidas (*Jackknife Delet- d Method*), fazendo com que o método usual possa ser denotado por método Jackknife com uma observação removida (*Jackknife Delet-1 Method*). Vale apontar que o método proposto em Shao e Wu (1989) pode vir a necessitar de um maior poder computacional.

3.7 Método Bootstrap

O método Bootstrap foi introduzido por Efron (1979) e consiste em um método de reamostragem com forte apelo computacional que fornece um mecanismo simples para estimar a distribuição amostral ou a precisão de um estimador de interesse $\hat{\theta}$, não importando o quão complexo seja a sua estrutura. De maneira semelhante ao Jackknife, o Bootstrap procede construindo subamostras com base na amostra em estudo (mecanismo de reamostragem). As subamostras são denotadas por amostras Bootstraps e possuem a mesma dimensão da amostra estudada (original), sendo construídas de forma independente por meio de amostragem aleatória com reposição, usando a amostra original como referência. Além disso, como visto em Efron e Tibshirani (1993), a metodologia é fundamentada com base na função de distribuição empírica.

3.7.1 Apresentação do método

Seja $\mathbf{w} = (w_1, w_2, \dots, w_n)^\top$ uma amostra aleatória observada com distribuição de probabilidade $F_{\mathbf{w}}$ (não necessariamente conhecida) e θ um parâmetro de interesse, no qual $\theta = t(F)$. Nosso problema de interesse consiste em obter a distribuição de probabilidade de um estimador qualquer $\hat{\theta} = s(\mathbf{x})$, para posteriormente obter suas medidas de variabilidade. Um método comum para resolver este problema, consiste na construção da distribuição amostral de $\hat{\theta}$ em que repete-se o experimento um número suficiente de vezes para aproximar a distribuição de $\hat{\theta}$ por meio de sua distribuição empírica, $\hat{F}_{\hat{\theta}}$. Contudo, repetir o experimento diversas vezes pode ser impraticável. Assim, como forma alternativa de solução, temos o método Bootstrap que sugere utilizar \mathbf{w} como alvo de estudo e refazer o processo de amostragem

com base na função de distribuição empírica dos dados, assim emulando processo de obtenção da distribuição amostral de $\hat{\theta}$. Em resumo, o método Bootstrap utiliza a amostra original como representante da população, retira um número suficiente de amostras Bootstrap e aproxima a distribuição amostral de qualquer estimador através de sua distribuição amostral empírica gerada pelas amostras Bootstrap, o que por consequência fornece a informação desejada, tal como variância, desvio padrão e viés.

A seguir, a descrição do algoritmo para construir a distribuição empírica do estimador $\hat{\theta}$ através do processo Bootstrap:

0. Realizar o experimento para obter uma amostra aleatória

$$\mathbf{w} = (w_1, w_2, \dots, w_n)^\top$$

e calcular a estimativa para o estimador de interesse, $\hat{\theta} = s(\mathbf{w})$;

1. Construir a distribuição empírica dos dados que associa probabilidade $1/n$ para cada observação, $\hat{F}_{\mathbf{w}}$;
2. A partir de $\hat{F}_{\mathbf{w}}$, realizar uma amostragem aleatória com reposição para obter uma amostra Bootstrap

$$\mathbf{w}^* = (w_1^*, w_2^*, \dots, w_n^*)^\top;$$

3. Calcular a estimativa Bootstrap para o estimador de interesse, $\hat{\theta}^* = s(\mathbf{w}^*)$;
4. Repetir de maneira independente as etapas 2 e 3 para um certo número B de vezes;
5. Com base nas B estimativas Bootstrap, construir a distribuição empírica Bootstrap para o estimador de interesse, $\hat{F}_{\hat{\theta}}^*$.

Como visto em Efron e Tibshirani (1993), no contexto não paramétrico, pode ser demonstrado que a distribuição empírica $\hat{F}_{\mathbf{w}}$ é uma estatística suficiente para a verdadeira distribuição de probabilidade $F_{\mathbf{w}}$, o que significa que toda informação necessária sobre F contida em \mathbf{w} , também está contida em $\hat{F}_{\mathbf{w}}$. Assim, de posse da distribuição empírica $\hat{F}_{\hat{\theta}}^*$, podemos obter qualquer informação sobre $\hat{\theta}$ por meio de estatísticas amostrais (medidas de resumo) como média, variância e quantis. Ainda em Efron e Tibshirani (1993), apresenta-se o princípio da extensão (*plug-in principle*) como justificativa teórica para o uso de medidas de resumo. O princípio da extensão é um método de inferência que estima o parâmetro de interesse $\theta = t(F_{\mathbf{w}})$, através do estimador consistente e assintoticamente não-viesado construído da forma $\hat{\theta} = t(\hat{F}_{\mathbf{w}})$. Com base nisso, as medidas de resumo aplicadas às B estimativas Bootstrap podem ser vistas como estimadores de extensão (*plug-in*) que usam a distribuição empírica no lugar da distribuição de probabilidade, assim proporcionando estimativas para $\hat{\theta}$. Esta mecânica proporcionada pelo método acarreta a necessidade de um número suficiente de amostras Bootstrap para validar as estimativas. Como sugestão para B , temos Efron e Tibshirani (1993) com a indicação de 25 a 200 reamostras; e Krzanowski (2000) com a indicação de 100 a 1000 reamostras, dependendo da complexidade do estimador usado. Como forma alternativa para encontrar um valor razoável para B , o pesquisador pode realizar testes com diferentes quantidades de reamostras e avaliar quais valores são razoáveis para o problema. Vale ressaltar que embora não exista valor definido para B , esta problemática está restrita somente ao poder computacional que o pesquisador dispõe. Além disso, por construir sua metodologia com base na função de distribuição empírica, este método recebe o nome de Bootstrap não-paramétrico. Para detalhes sobre o método paramétrico, ver Efron e Tibshirani (1993, pags. 53), por exemplo. Neste trabalho adotaremos somente o método Bootstrap não-paramétrico.

Após a aplicação do algoritmo Bootstrap, o estimador Bootstrap para variância de $\hat{\theta}$ é obtido

da forma

$$\widehat{\text{VAR}}_{bs} = \frac{\sum_{i=1}^B (\hat{\theta}_i^* - \hat{\theta}_{(\cdot)}^*)^2}{B-1}, \quad (35)$$

no qual $\hat{\theta}_{(\cdot)}^*$ denota a média aritmética das estimativas Bootstrap. A estimativa do viés de $\hat{\theta}$ baseado nas B estimativas Bootstrap é obtido da forma

$$\hat{B}_{bs} = \hat{\theta}_{(\cdot)}^* - \hat{\theta}, \quad (36)$$

no qual podemos melhorar (no sentido da redução do viés) o estimador $\hat{\theta}$ construindo o estimador Bootstrap da forma

$$\hat{\theta}_{bs} = 2\hat{\theta} - \hat{\theta}_{(\cdot)}^*. \quad (37)$$

3.7.2 Intervalos de Confiança

Ao longo do tempo, o Bootstrap recebeu diversos estudos sobre a temática dos intervalos de confiança não-paramétricos, proporcionando vários métodos de construção. Dentre os vários tipos de intervalos, encontra-se o intervalo não-paramétrico percentil que foi apresentado em Efron (1979) e recebeu maiores detalhes em Efron (1981). Sua metodologia sugere que o intervalo com confiança $1 - 2\alpha$ para θ , com $\alpha \in [0, 1]$, seja construído com base no uso dos percentis obtidos através da distribuição empírica Bootstrap, $\hat{F}_{\hat{\theta}}$, sendo estruturado da forma

$$[\hat{\theta}_{\alpha}^*, \hat{\theta}_{1-\alpha}^*], \quad (38)$$

no qual $\hat{\theta}_{\alpha}^*$ e $\hat{\theta}_{1-\alpha}^*$ denotam respectivamente os percentis α e $1 - \alpha$ da distribuição empírica Bootstrap. Como visto em Tibshirani (1984), o intervalo percentil também pode ser adaptado para gerar o intervalo pivotal da forma

$$[2\hat{\theta} - \hat{\theta}_{1-\alpha}^*, 2\hat{\theta} - \hat{\theta}_{\alpha}^*], \quad (39)$$

no qual esta estrutura é construída aproximando assintoticamente a distribuição da quantidade pivotal $\hat{\theta} - \theta$, por meio da distribuição Bootstrap da quantidade $\hat{\theta}^* - \hat{\theta}$. Vale ressaltar que

$$\hat{\theta} - \theta \sim C, \quad (40)$$

no qual C denota uma distribuição (não necessariamente conhecida) que não depende θ .

O intervalo percentil, embora simples e prático, não apresenta boa cobertura para amostras pequenas e por isso recebeu um primeiro refinamento em Efron (1981) e posteriormente um segundo refinamento em Efron (1987). O primeiro refinamento foi chamado intervalo percentil corrigido pelo viés (*bias-corrected percentile confidence interval - BC*) e foi estruturado para fornecer o intervalo de confiança com o nível $1 - 2\alpha$ da forma

$$[Q^*(\Phi(2z_0 + z_{\alpha})), Q^*(\Phi(2z_0 + z_{1-\alpha}))], \quad (41)$$

em que $z_0 = \Phi^{-1}(Q^*(\hat{\theta}))$ define a correção de viés, $\Phi(\cdot)^{-1}$ define a função quantílica da distribuição Normal padrão e $Q^*(\cdot)$ define a função quantílica da distribuição empírica $\hat{F}_{\hat{\theta}}^*$. Por sua vez, as quantidades z_{α} e $z_{1-\alpha}$ denotam respectivamente os quantis α e $1 - \alpha$ da distribuição Normal padrão. O segundo refinamento foi chamado de intervalo percentil acelerado corrigido pelo viés (*bias-corrected and accelerated percentile confidence interval - BC_a*) e fornece o intervalo de confiança com o nível $1 - 2\alpha$ por meio da estrutura

$$\left[Q^* \left(\Phi \left(h_0 + \frac{h_0 + z_{\alpha}}{1 - \hat{a}(h_0 + z_{\alpha})} \right) \right), Q^* \left(\Phi \left(h_0 + \frac{h_0 + z_{1-\alpha}}{1 - \hat{a}(h_0 + z_{1-\alpha})} \right) \right) \right], \quad (42)$$

em que $\Phi(\cdot)$, h_0 e \hat{a} definem, respectivamente, a função de distribuição acumulada do modelo normal padrão, a correção de viés e a constante de aceleração. O componente h_0 é definido da forma

$$h_0 = \frac{\sum_{i=1}^B \mathbb{1}_{(\hat{\theta}_i^* < \hat{\theta})}(\hat{\theta}_i^*)}{B}. \quad (43)$$

Por sua vez, o componente \hat{a} não possui estrutura específica, mas pode ser facilmente construído por meio do método Jackknife da forma

$$\hat{a} = \frac{\sum_{i=1}^n (\hat{\theta}_{(\cdot)} - \hat{\theta}_{(i)})^3}{6 \left\{ \sum_{i=1}^n (\hat{\theta}_{(\cdot)} - \hat{\theta}_{(i)})^2 \right\}^{3/2}}, \quad (44)$$

como visto em Efron e Tibshirani (1993), por exemplo.

3.8 Testes de Hipóteses

Dado os algoritmos de otimização e reamostragem, testes de hipóteses clássicos como o teste da razão de verossimilhanças generalizada ou testes do tipo Wald também podem ser aplicados ao modelo de regressão L-logístico. Estes testes compartilham da seguinte estruturação de hipóteses:

$$\begin{aligned} H_0 : \boldsymbol{\beta} &= [\beta_0, \beta_1, \beta_2, \dots, \beta_q, 0, \dots, 0]^\top \\ H_1 : \boldsymbol{\beta} &\neq [\beta_0, \beta_1, \beta_2, \dots, \beta_q, 0, \dots, 0]^\top \end{aligned} \quad (45)$$

3.8.1 Teste da Razão de Verossimilhanças Generalizada

Com base na estruturação das hipóteses, como visto em Bolfarini e Sandoval (2010), o teste da razão de verossimilhanças generalizada apresenta a seguinte estatística do teste

$$\Lambda = 2 [l_{H_0}(\boldsymbol{\beta}; \mathbf{y}, X) - l_{H_1}(\boldsymbol{\beta}; \mathbf{y}, X)] \quad (46)$$

em que $l_{H_k}(\cdot)$ denota o logaritmo natural da função de verossimilhança sobre a suposição da veracidade da k -ésima hipótese ($k = 0, 1$) e sob a veracidade de H_0 e condições de regularidade Λ possui distribuição assintótica Qui-quadrado.

3.8.2 Teste do tipo Wald com Bootstrap

Dado as propriedades do estimador de máxima verossimilhança, como visto em McCullagh e Nelder (1989), o clássico teste de Wald apresenta a estatística do teste da forma

$$W = (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_{H_0})^\top \Sigma_{\hat{\boldsymbol{\beta}}}^{-1} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_{H_0}), \quad (47)$$

em que $\boldsymbol{\beta}_{H_0}$ denota o vetor de parâmetros estipulado em H_0 , $\hat{\boldsymbol{\beta}}$ denota o vetor estimado de parâmetros do modelo, $\Sigma_{\hat{\boldsymbol{\beta}}}$ denota a matriz de covariâncias do vetor estimado $\hat{\boldsymbol{\beta}}$ e como no teste da razão de verossimilhança generalizada, sob a veracidade de H_0 e condições de regularidade, a estatística W possui distribuição assintótica Qui-quadrado. O teste também apresenta sua versão univariada com estatística do teste da forma

$$W^j = \frac{\hat{\beta}_j - \beta_0}{\sqrt{\hat{\sigma}_{\hat{\beta}_j}^2}}, \quad (48)$$

no qual $\hat{\beta}_j$ denota a j -ésima componente do vetor $\hat{\boldsymbol{\beta}}$, $\hat{\sigma}_{\hat{\beta}_j}^2$ denota o j -ésimo elemento da diagonal principal da matriz $\Sigma_{\hat{\boldsymbol{\beta}}}$ que corresponde à variância associada a $\hat{\beta}_j$ e W^j possui distribuição t -student com $n - p - 1$ graus de liberdade, em que n denota o número observações da amostra e p denota o número de parâmetros

no vetor β . Contudo, pode-se utilizar o método Bootstrap para fornecer a estimativa para o erro padrão (teste univariado) ou para a matriz de covariâncias (teste multivariado) de modo a construir a estatística do teste. Além disso o método Bootstrap fornece um estimador do erro padrão que possui consistência assegurada pelo teorema de Glivenko–Cantelli, como visto em van der Vaart e Wellner (1996). Como exemplo de uso da combinação do método bootstrap com o teste de Wald, pode-se citar como por exemplo Johnson e Romer (2016) e Wilcox (1995). Na literatura, os testes que utilizam quantidades pivotais e possuem estruturas semelhantes ao tradicional teste de Wald geralmente recebem a notação de testes do tipo Wald, como por exemplo Basu et al. (2017), Fay e Graubard (2001), Dufour (1997) e Silvapulle e Silvapulle (1995).

3.9 Medidas de Diagnóstico

Em virtude da recente proposta do modelo de regressão L-logística, ferramentas (medidas) de diagnóstico são escassas na literatura, quando visamos a não utilização de derivadas da função de verossimilhança. Contudo, podemos detectar observações atípicas (pontos influentes) por meio de uma medida D_i semelhante a usual distância de Cook. Esta medida pode ser entendida como uma distância do tipo Cook e é definida por

$$D_i = \frac{(\beta - \beta^{(i)})^\top (\beta - \beta^{(i)})}{|\mathbf{D}|}, \quad (49)$$

no qual β denota o vetor de parâmetros, $\beta^{(i)}$ denota o vetor de parâmetros com a ausência da i -ésima observação e $|\mathbf{D}|$ denota a norma do vetor das distâncias. Além disso, D_i corresponde a medida de distância da i -ésima observação e pertence ao intervalo $[0, 1]$.

4 SIMULAÇÃO

Esta seção apresenta a implementação do método de simulação de Monte Carlo e utiliza as metaheurísticas junto aos métodos de reamostragem para estimar os parâmetros da distribuição L-logística, visando o método de máxima verossimilhança. A implementação das técnicas foram realizadas através da versão 3.3.3 do software livre R. Por razões de performance, optou-se pelo uso de pacotes computacionais. O GA foi implementado por do pacote **GA**, com autoria de Scrucca (2013). O DE foi implementado por meio do pacote **DEoptim**, com autoria de Mullen et al. (2011). Os demais processos foram implementados manualmente. Os resultados foram obtidos através da construção de 1000 réplicas de Monte Carlo para os tamanhos amostrais 10, 25, 50 e 100. A estimação foi aplicada a nove cenários oriundos da combinação de $m = (0, 2; 0, 5; 0, 9)^\top$ e $b = (0, 5; 1; 2)^\top$ que denotam valores para os parâmetros de posição e dispersão (respectivamente) da distribuição L-logística. Tais valores de m e b foram escolhidos de modo a ilustrar características úteis da distribuição.

As estimativas pontuais obtidas com as 1000 réplicas, $R = 1000$, foram utilizadas para calcular as seguintes estatísticas:

- **Média:** média das réplicas

$$\hat{\theta}_R = \sum_{i=1}^R \frac{\hat{\theta}_i}{R},$$

em que $\hat{\theta}_i$ denotando a estimativa para a i -ésima réplica;

- **EPS:** erro de simulação

$$SD(\hat{\theta}_R) = \sqrt{\sum_{i=1}^R \frac{(\hat{\theta}_i - \hat{\theta}_R)^2}{R - 1}};$$

- **Viés:** diferença da média com o valor verdadeiro do parâmetro

$$B(\hat{\theta}_R) = \hat{\theta}_R - \theta;$$

- **REQM:** raiz do erro quadrático médio

$$REQM(\hat{\theta}_R) = \sqrt{[B(\hat{\theta}_R)]^2 + [SD(\hat{\theta}_R)]^2};$$

- **AVRB:** valor absoluto do viés dividido pelo valor absoluto do verdadeiro valor do parâmetro,

$$AVRB(\hat{\theta}_R) = \frac{|B(\hat{\theta}_R)|}{|\theta|}.$$

Com exceção da média e do viés, para todas as outras estatísticas, quanto menor o seu valor, melhor o desempenho do método de estimação. Vale lembrar, que o viés também tem esse efeito quando analisado em termo de seu valor absoluto. A seguir, na Figura 10 e na Tabela 1, apresentam-se os gráficos dos 9 cenários e a parametrização das metaheurísticas. Além disso, utilizou-se o número máximo de gerações como critério de parada dos algoritmos, no qual o DE foi implementado em seu formato DE/rand/1/bin, enquanto que o GA foi implementado usando o operador de seleção por torneio, o operador de recombinação BLX- α e o operador de mutação uniforme.

Figura 10: Ilustração gráfica do comportamento de todos os cenários usados na simulação. Histogramas produzidos com amostras de tamanho 5×10^5 da distribuição L-logística.

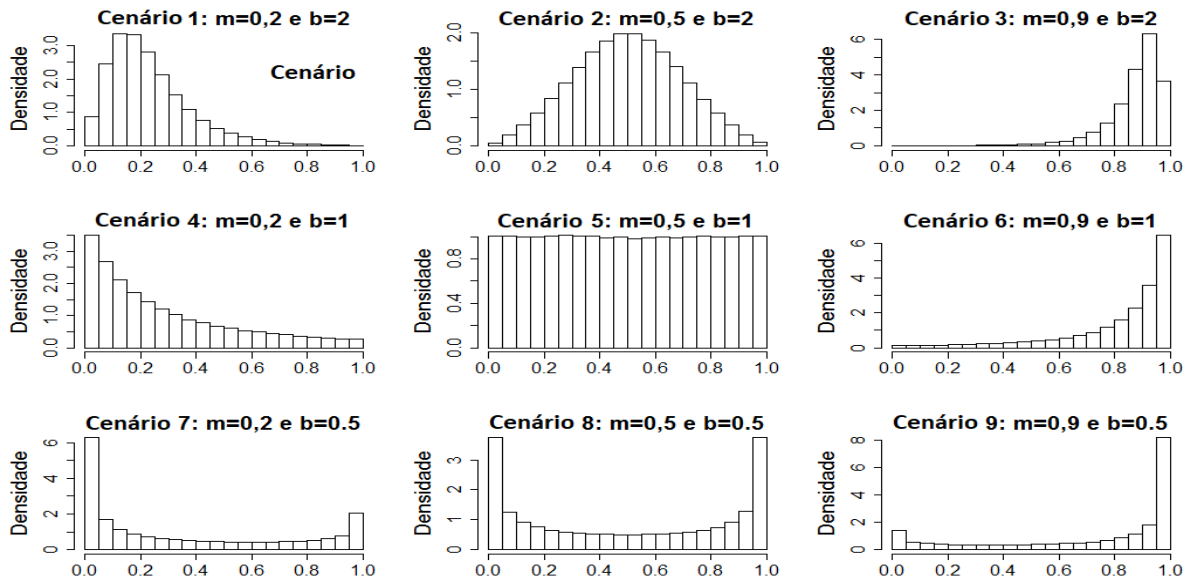


Tabela 1: Configuração dos Algoritmos.

Metaheurística	N_p	N_G	P_r	p_m	F
GA	300	90	0,85	0,001	—
DE	300	90	0,85	—	0,5

Com relação aos métodos de reamostragem, em cada cenário construiu-se os intervalos para 0,95 de confiança e avaliou-se a probabilidade de cobertura (PC), o comprimento médio apresentado pelos intervalos (MComp) e o desvio padrão do comprimento apresentado pelos intervalos (DPCComp). Para o método Bootstrap utilizou-se 500 reamostras e realizou-se a construção dos intervalos pivotal, percentil, BC e BC_a para cada réplica de Monte Carlo. Para o método Jackknife, em cada réplica construiu-se o intervalo baseado na distribuição t de Student. Vale ressaltar que a constante de aceleração do intervalo BC_a foi construída com base no método Jackknife.

4.1 Resultados: Estatísticas da Simulação

Nesta seção apresentam-se os resultados das estatísticas obtidos no processo de simulação para os 9 cenários propostos. Constata-se que em todos os casos, os algoritmos GA e DE apresentaram os mesmos resultados para todos os tamanhos de amostra, $n = (10, 25, 50, 100, 500,)^\top$. Além disso, constata-se que em todos os casos:

- i) Os algoritmos forneceram soluções próximas dos valores reais estipulados e melhoraram a precisão conforme o aumento de n , como esperado;
- ii) Dado a proximidade das estimativas com o verdadeiro valor do parâmetro, todas as estatísticas referentes à variabilidade demonstraram valores baixos.
- iii) Dado a característica do parâmetro de dispersão b , suas estimativas necessitam de maiores tamanhos de n ;
- iv) Existe influência nas estimativas do parâmetro m , conforme os valores adotados para o parâmetro b . O que não ocorre no processo inverso.

A seguir, nas Tabelas 2 a 10, todos os resultados obtidos no processo de simulação.

Tabela 2: Resultados da simulação para o cenário 1. $\theta = (m; b)^\top = (0, 2; 2)^\top$.

Algoritmo	Amostra	θ	Média	EPS	Viés	REQM	AVRB
GA	10	m	0,2036	0,0466	0,0036	0,0468	0,0179
		b	2,2817	0,7039	0,2817	0,7582	0,1409
	25	m	0,2010	0,0286	0,0010	0,0286	0,0052
		b	2,1196	0,3737	0,1196	0,3924	0,0598
	50	m	0,1998	0,0203	-0,0002	0,0203	0,0009
		b	2,0563	0,2521	0,0563	0,2583	0,0281
	100	m	0,2002	0,0141	0,0002	0,0141	0,0008
		b	2,0297	0,1705	0,0297	0,1730	0,0148
	500	m	0,2001	0,0062	0,0001	0,0062	0,0005
		b	2,0055	0,0766	0,0055	0,0768	0,0027
DE	10	m	0,2036	0,0466	0,0036	0,0468	0,0179
		b	2,2817	0,7039	0,2817	0,7582	0,1409
	25	m	0,2010	0,0286	0,0010	0,0286	0,0052
		b	2,1196	0,3737	0,1196	0,3924	0,0598
	50	m	0,1998	0,0203	-0,0002	0,0203	0,0009
		b	2,0563	0,2521	0,0563	0,2583	0,0281
	100	m	0,2002	0,0141	0,0002	0,0141	0,0008
		b	2,0297	0,1705	0,0297	0,1730	0,0148
	500	m	0,2001	0,0062	0,0001	0,0062	0,0005
		b	2,0055	0,0766	0,0055	0,0768	0,0027

Tabela 3: Resultados da simulação para o cenário 2. $\theta = (m; b)^\top = (0, 5; 2)^\top$.

Algoritmo	Amostra	θ	Média	EPS	Viés	REQM	AVRB
GA	10	m	0,4995	0,0706	-0,0005	0,0706	0,0011
		b	2,2817	0,7039	0,2817	0,7582	0,1409
	25	m	0,4993	0,0444	-0,0007	0,0444	0,0015
		b	2,1196	0,3737	0,1196	0,3924	0,0598
	50	m	0,4985	0,0317	-0,0015	0,0318	0,003
		b	2,0563	0,2521	0,0563	0,2583	0,0281
	100	m	0,4997	0,0220	-0,0003	0,0220	0,0007
		b	2,0297	0,1705	0,0297	0,1730	0,0148
	500	m	0,5000	0,0097	$< 10^{-4}$	0,0097	0,0001
		b	2,0055	0,0766	0,0055	0,0768	0,0027
DE	10	m	0,4995	0,0706	-0,0005	0,0706	0,0011
		b	2,2817	0,7039	0,2817	0,7582	0,1409
	25	m	0,4993	0,0444	-0,0007	0,0444	0,0015
		b	2,1196	0,3737	0,1196	0,3924	0,0598
	50	m	0,4985	0,0317	-0,0015	0,0318	0,003
		b	2,0563	0,2521	0,0563	0,2583	0,0281
	100	m	0,4997	0,0220	-0,0003	0,0220	0,0007
		b	2,0297	0,1705	0,0297	0,1730	0,0148
	500	m	0,5000	0,0097	$< 10^{-4}$	0,0097	0,0001
		b	2,0055	0,0766	0,0055	0,0768	0,0027

Tabela 4: Resultados da simulação para o cenário 3. $\theta = (m; b)^\top = (0, 9; 2)^\top$.

Algoritmo	Amostra	θ	Média	EPS	Viés	REQM	AVRB	
GA	10	m	0,8968	0,0267	-0,0032	0,0269	0,0035	
		b	2,2817	0,7039	0,2817	0,7582	0,1409	
	25	m	0,8986	0,0164	-0,0014	0,0165	0,0016	
		b	2,1196	0,3737	0,1196	0,3924	0,0598	
	50	m	0,8989	0,0116	-0,0011	0,0117	0,0013	
		b	2,0563	0,2521	0,0563	0,2583	0,0281	
	100	m	0,8996	0,0079	-0,0004	0,0080	0,0004	
		b	2,0297	0,1705	0,0297	0,1730	0,0148	
	500	m	0,9000	0,0035	$< 10^{-4}$	0,0035	$< 10^{-4}$	
		b	2,0055	0,0766	0,0055	0,0768	0,0027	
	DE	10	m	0,8968	0,0267	-0,0032	0,0269	0,0035
			b	2,2817	0,7039	0,2817	0,7582	0,1409
25		m	0,8986	0,0164	-0,0014	0,0165	0,0016	
		b	2,1196	0,3737	0,1196	0,3924	0,0598	
50		m	0,8989	0,0116	-0,0011	0,0117	0,0013	
		b	2,0563	0,2521	0,0563	0,2583	0,0281	
100		m	0,8996	0,0079	-0,0004	0,0080	0,0004	
		b	2,0297	0,1705	0,0297	0,1730	0,0148	
500		m	0,9000	0,0035	$< 10^{-4}$	0,0035	$< 10^{-4}$	
		b	2,0055	0,0766	0,0055	0,0768	0,0027	

Tabela 5: Resultados da simulação para o cenário 4. $\theta = (m; b)^\top = (0, 2; 1)^\top$.

Algoritmo	Amostra	θ	Média	EPS	Viés	REQM	AVRB	
GA	10	m	0,2141	0,0947	0,0141	0,0957	0,0705	
		b	1,1409	0,3519	0,1409	0,3791	0,1409	
	25	m	0,2050	0,0574	0,0050	0,0576	0,0250	
		b	1,0598	0,1869	0,0598	0,1962	0,0598	
	50	m	0,2011	0,0407	0,0011	0,0407	0,0057	
		b	1,0281	0,1261	0,0281	0,1292	0,0281	
	100	m	0,2011	0,0284	0,0011	0,0284	0,0053	
		b	1,0148	0,0852	0,0148	0,0865	0,0148	
	500	m	0,2003	0,0124	0,0003	0,0124	0,0017	
		b	1,0027	0,0383	0,0027	0,0384	0,0027	
	DE	10	m	0,2141	0,0947	0,0141	0,0957	0,0705
			b	1,1409	0,3519	0,1409	0,3791	0,1409
25		m	0,2050	0,0574	0,0050	0,0576	0,0250	
		b	1,0598	0,1869	0,0598	0,1962	0,0598	
50		m	0,2011	0,0407	0,0011	0,0407	0,0057	
		b	1,0281	0,1261	0,0281	0,1292	0,0281	
100		m	0,2011	0,0284	0,0011	0,0284	0,0053	
		b	1,0148	0,0852	0,0148	0,0865	0,0148	
500		m	0,2003	0,0124	0,0003	0,0124	0,0017	
		b	1,0027	0,0383	0,0027	0,0384	0,0027	

Tabela 6: Resultados da simulação para o cenário 5. $\theta = (m; b)^\top = (0, 5; 1)^\top$.

Algoritmo	Amostra	θ	Média	EPS	Viés	REQM	AVRB	
GA	10	m	0,4989	0,1338	-0,0011	0,1338	0,0022	
		b	1,1409	0,3519	0,1409	0,3791	0,1409	
	25	m	0,4986	0,0868	-0,0014	0,0868	0,0028	
		b	1,0598	0,1869	0,0598	0,1962	0,0598	
	50	m	0,4971	0,0626	-0,0029	0,0627	0,0059	
		b	1,0281	0,1261	0,0281	0,1292	0,0281	
	100	m	0,4993	0,0437	-0,0007	0,0437	0,0013	
		b	1,0148	0,0852	0,0148	0,0865	0,0148	
	500	m	0,5001	0,0193	0,0001	0,0193	0,0002	
		b	1,0027	0,0383	0,0027	0,0384	0,0027	
	DE	10	m	0,4989	0,1338	-0,0011	0,1338	0,0022
			b	1,1409	0,3519	0,1409	0,3791	0,1409
25		m	0,4986	0,0868	-0,0014	0,0868	0,0028	
		b	1,0598	0,1869	0,0598	0,1962	0,0598	
50		m	0,4971	0,0626	-0,0029	0,0627	0,0059	
		b	1,0281	0,1261	0,0281	0,1292	0,0281	
100		m	0,4993	0,0437	-0,0007	0,0437	0,0013	
		b	1,0148	0,0852	0,0148	0,0865	0,0148	
500		m	0,5001	0,0193	0,0001	0,0193	0,0002	
		b	1,0027	0,0383	0,0027	0,0384	0,0027	

Tabela 7: Resultados da simulação para o cenário 6. $\theta = (m; b)^\top = (0, 9; 1)^\top$.

Algoritmo	Amostra	θ	Média	EPS	Viés	REQM	AVRB	
GA	10	m	0,8878	0,0575	-0,0122	0,0588	0,0135	
		b	1,1409	0,3519	0,1409	0,3791	0,1409	
	25	m	0,8948	0,0341	-0,0052	0,0345	0,0057	
		b	1,0598	0,1869	0,0598	0,1962	0,0598	
	50	m	0,8966	0,0239	-0,0034	0,0242	0,0038	
		b	1,0281	0,1261	0,0281	0,1292	0,0281	
	100	m	0,8986	0,0160	-0,0014	0,0161	0,0015	
		b	1,0148	0,0852	0,0148	0,0865	0,0148	
	500	m	0,8998	0,0070	-0,0002	0,0070	0,0002	
		b	1,0027	0,0383	0,0027	0,0384	0,0027	
	DE	10	m	0,8878	0,0575	-0,0122	0,0588	0,0135
			b	1,1409	0,3519	0,1409	0,3791	0,1409
25		m	0,8948	0,0341	-0,0052	0,0345	0,0057	
		b	1,0598	0,1869	0,0598	0,1962	0,0598	
50		m	0,8966	0,0239	-0,0034	0,0242	0,0038	
		b	1,0281	0,1261	0,0281	0,1292	0,0281	
100		m	0,8986	0,0160	-0,0014	0,0161	0,0015	
		b	1,0148	0,0852	0,0148	0,0865	0,0148	
500		m	0,8998	0,0070	-0,0002	0,0070	0,0002	
		b	1,0027	0,0383	0,0027	0,0384	0,0027	

Tabela 8: Resultados da simulação para o cenário 7. $\theta = (m; b)^\top = (0, 2; 0, 5)^\top$.

Algoritmo	Amostra	θ	Média	EPS	Viés	REQM	AVRB
GA	10	m	0,2467	0,1857	0,0467	0,1915	0,2334
		b	0,5704	0,1760	0,0704	0,1895	0,1409
	25	m	0,2201	0,1160	0,0201	0,1177	0,1003
		b	0,5299	0,0934	0,0299	0,0981	0,0598
	50	m	0,2079	0,0820	0,0079	0,0824	0,0394
		b	0,5141	0,0630	0,0141	0,0646	0,0281
	100	m	0,2049	0,0574	0,0049	0,0577	0,0247
		b	0,5074	0,0426	0,0074	0,0433	0,0148
	500	m	0,2013	0,0248	0,0013	0,0248	0,0063
		b	0,5014	0,0192	0,0014	0,0192	0,0027
DE	10	m	0,2467	0,1857	0,0467	0,1915	0,2334
		b	0,5704	0,1760	0,0704	0,1895	0,1409
	25	m	0,2201	0,1160	0,0201	0,1177	0,1003
		b	0,5299	0,0934	0,0299	0,0981	0,0598
	50	m	0,2079	0,0820	0,0079	0,0824	0,0394
		b	0,5141	0,0630	0,0141	0,0646	0,0281
	100	m	0,2049	0,0574	0,0049	0,0577	0,0247
		b	0,5074	0,0426	0,0074	0,0433	0,0148
	500	m	0,2013	0,0248	0,0013	0,0248	0,0063
		b	0,5014	0,0192	0,0014	0,0192	0,0027

Tabela 9: Resultados da simulação para o cenário 8. $\theta = (m; b)^\top = (0, 5; 0, 5)^\top$.

Algoritmo	Amostra	θ	Média	EPS	Viés	REQM	AVRB
GA	10	m	0,4975	0,2288	-0,0025	0,2288	0,0050
		b	0,5704	0,1760	0,0704	0,1895	0,1409
	25	m	0,4975	0,1608	-0,0025	0,1608	0,0050
		b	0,5299	0,0934	0,0299	0,0981	0,0598
	50	m	0,4944	0,1196	-0,0056	0,1198	0,0112
		b	0,5141	0,0630	0,0141	0,0646	0,0281
	100	m	0,4986	0,0854	-0,0014	0,0854	0,0027
		b	0,5074	0,0426	0,0074	0,0433	0,0148
	500	m	0,5002	0,0385	0,0002	0,0385	0,0004
		b	0,5014	0,0192	0,0014	0,0192	0,0027
DE	10	m	0,4975	0,2288	-0,0025	0,2288	0,0050
		b	0,5704	0,1760	0,0704	0,1895	0,1409
	25	m	0,4975	0,1608	-0,0025	0,1608	0,0050
		b	0,5299	0,0934	0,0299	0,0981	0,0598
	50	m	0,4944	0,1196	-0,0056	0,1198	0,0112
		b	0,5141	0,0630	0,0141	0,0646	0,0281
	100	m	0,4986	0,0854	-0,0014	0,0854	0,0027
		b	0,5074	0,0426	0,0074	0,0433	0,0148
	500	m	0,5002	0,0385	0,0002	0,0385	0,0004
		b	0,5014	0,0192	0,0014	0,0192	0,0027

Tabela 10: Resultados da simulação para o cenário 9. $\theta = (m; b)^\top = (0, 9; 0, 5)^\top$.

Algoritmo	Amostra	θ	Média	EPS	Viés	REQM	AVRB	
GA	10	m	0,8558	0,1335	-0,0442	0,1406	0,0492	
		b	0,5704	0,1760	0,0704	0,1895	0,1409	
	25	m	0,8808	0,0762	-0,0192	0,0786	0,0213	
		b	0,5299	0,0934	0,0299	0,0981	0,0598	
	50	m	0,8884	0,0518	-0,0116	0,0531	0,0128	
		b	0,5141	0,0630	0,0141	0,0646	0,0281	
	100	m	0,8951	0,0330	-0,0049	0,0334	0,0055	
		b	0,5074	0,0426	0,0074	0,0433	0,0148	
	500	m	0,8992	0,0141	-0,0008	0,0142	0,0009	
		b	0,5014	0,0192	0,0014	0,0192	0,0027	
	DE	10	m	0,8558	0,1335	-0,0442	0,1406	0,0492
			b	0,5704	0,1760	0,0704	0,1895	0,1409
25		m	0,8808	0,0762	-0,0192	0,0786	0,0213	
		b	0,5299	0,0934	0,0299	0,0981	0,0598	
50		m	0,8884	0,0518	-0,0116	0,0531	0,0128	
		b	0,5141	0,0630	0,0141	0,0646	0,0281	
100		m	0,8951	0,0330	-0,0049	0,0334	0,0055	
		b	0,5074	0,0426	0,0074	0,0433	0,0148	
500		m	0,8992	0,0141	-0,0008	0,0142	0,0009	
		b	0,5014	0,0192	0,0014	0,0192	0,0027	

4.2 Resultados: Métodos de Reamostragem

Nesta seção apresenta-se os resultados dos métodos de reamostragem Jackknife e Bootstrap, obtidos no processo de simulação para os 9 cenários propostos. Constata-se que em todos os casos, novamente, os algoritmos GA e DE apresentaram os mesmos resultados para todos os tamanhos de amostra, n . Além disso, observa-se que:

- i) Conforme aumenta-se o tamanho da amostra, como esperado, a probabilidade de cobertura aproxima-se do valor estipulado de 0,95, para todos os intervalos.
- ii) Para os 9 cenários, o aumento da amostra acarreta uma redução nas estatísticas de média e desvio padrão do comprimento de cada intervalo;
- iii) Em todos os cenários referentes ao parâmetro de dispersão b , para amostras menores ou iguais a 100, a probabilidade de cobertura apresentada pelo intervalo Jackknife é superior ao valor estipulado de 0,95. Para $n = 500$, o intervalo Jackknife demonstra semelhança aos demais intervalos.
- iv) Para o parâmetro de posição m , a utilização dos intervalos surge de maneira particular, tendo para cada cenário e tamanho de amostra um melhor tipo de intervalo, não sendo possível recomendar um intervalo com desempenho superior em todas as situações. Com relação ao parâmetro de dispersão b , com exceção do intervalo Jackknife com $n \leq 100$, a particularidade dos intervalos também ocorre, não sendo possível a recomendação de um intervalo geral.

A seguir, nas Tabelas 11 a 46, todos os resultados obtidos no processo de reamostragem efetuado na simulação.

Tabela 11: Probabilidade de cobertura com Algoritmo Genético para $m = 0, 2$ no cenário 1.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,9330	0,1969	0,0684
		Pivotal	0,8660	0,1687	0,0556
		Percentil	0,9020	0,1687	0,0556
		BC	0,9000	0,1688	0,0558
		BC _a	0,8980	0,1699	0,0563
	25	Jackknife- t	0,9310	0,1138	0,0248
		Pivotal	0,9080	0,1077	0,0232
		Percentil	0,9200	0,1077	0,0232
		BC	0,9180	0,1078	0,0233
		BC _a	0,9180	0,1080	0,0234
	50	Jackknife- t	0,9270	0,0784	0,0118
		Pivotal	0,9150	0,0765	0,0117
		Percentil	0,9290	0,0765	0,0117
		BC	0,9310	0,0765	0,0117
		BC _a	0,9300	0,0766	0,0117
	100	Jackknife- t	0,9430	0,0549	0,0057
		Pivotal	0,9300	0,0540	0,0060
		Percentil	0,9430	0,0540	0,0060
		BC	0,9420	0,0540	0,0060
		BC _a	0,9410	0,0540	0,0060
500	Jackknife- t	0,9480	0,0244	0,0011	
	Pivotal	0,9500	0,0241	0,0015	
	Percentil	0,9460	0,0241	0,0015	
	BC	0,9460	0,0241	0,0015	
	BC _a	0,9460	0,0241	0,0015	

Tabela 12: Probabilidade de cobertura com Evolução Diferencial para $m = 0,2$ no cenário 1.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,9330	0,1969	0,0684
		Pivotal	0,8660	0,1687	0,0556
		Percentil	0,9020	0,1687	0,0556
		BC	0,9000	0,1688	0,0558
		BC _a	0,8980	0,1699	0,0563
	25	Jackknife- t	0,9310	0,1138	0,0248
		Pivotal	0,9080	0,1077	0,0232
		Percentil	0,9200	0,1077	0,0232
		BC	0,9180	0,1078	0,0233
		BC _a	0,9180	0,1080	0,0234
	50	Jackknife- t	0,9270	0,0784	0,0118
		Pivotal	0,9150	0,0765	0,0117
		Percentil	0,9290	0,0765	0,0117
		BC	0,9310	0,0765	0,0117
		BC _a	0,9300	0,0766	0,0117
	100	Jackknife- t	0,9430	0,0549	0,0057
		Pivotal	0,9300	0,0540	0,0060
		Percentil	0,9430	0,0540	0,0060
		BC	0,9420	0,0540	0,0060
		BC _a	0,9410	0,0540	0,0060
500	Jackknife- t	0,9480	0,0244	0,0011	
	Pivotal	0,9500	0,0241	0,0015	
	Percentil	0,9460	0,0241	0,0015	
	BC	0,9460	0,0241	0,0015	
	BC _a	0,9460	0,0241	0,0015	

Tabela 13: Probabilidade de cobertura com Algoritmo Genético para $b = 2$ no cenário 1.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,9580	3,4069	1,9246
		Pivotal	0,8720	3,2766	1,5962
		Percentil	0,9020	3,2766	1,5962
		BC	0,8210	2,4181	1,2889
		BC _a	0,8980	2,1877	1,0693
	25	Jackknife- t	0,9540	1,5494	0,4497
		Pivotal	0,9130	1,4993	0,4126
		Percentil	0,9200	1,4993	0,4126
		BC	0,8840	1,3537	0,3886
		BC _a	0,9180	1,3119	0,3635
	50	Jackknife- t	0,9550	1,0032	0,2025
		Pivotal	0,9330	0,9828	0,1949
		Percentil	0,9290	0,9828	0,1949
		BC	0,9180	0,9358	0,1885
		BC _a	0,9300	0,9238	0,1836
	100	Jackknife- t	0,9610	0,6805	0,0915
		Pivotal	0,9460	0,6682	0,0917
		Percentil	0,9430	0,6682	0,0917
		BC	0,9440	0,6528	0,0905
		BC _a	0,9410	0,6492	0,0906
500	Jackknife- t	0,9480	0,2953	0,0179	
	Pivotal	0,9440	0,2917	0,0215	
	Percentil	0,9460	0,2917	0,0215	
	BC	0,9310	0,2906	0,0213	
	BC _a	0,9460	0,2903	0,0211	

Tabela 14: Probabilidade de cobertura com Evolução Diferencial para $b = 2$ no cenário 1.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,9580	3,4069	1,9246
		Pivotal	0,8720	3,2766	1,5962
		Percentil	0,9020	3,2766	1,5962
		BC	0,8210	2,4181	1,2889
		BC _a	0,8980	2,1877	1,0693
	25	Jackknife- t	0,9540	1,5494	0,4497
		Pivotal	0,9130	1,4993	0,4126
		Percentil	0,9200	1,4993	0,4126
		BC	0,8840	1,3537	0,3886
		BC _a	0,9180	1,3119	0,3635
	50	Jackknife- t	0,9550	1,0032	0,2025
		Pivotal	0,9330	0,9828	0,1949
		Percentil	0,9290	0,9828	0,1949
		BC	0,9180	0,9358	0,1885
		BC _a	0,9300	0,9238	0,1836
	100	Jackknife- t	0,9610	0,6805	0,0915
		Pivotal	0,9460	0,6682	0,0917
		Percentil	0,9430	0,6682	0,0917
		BC	0,9440	0,6528	0,0905
		BC _a	0,9410	0,6492	0,0906
500	Jackknife- t	0,9480	0,2953	0,0179	
	Pivotal	0,9440	0,2917	0,0215	
	Percentil	0,9460	0,2917	0,0215	
	BC	0,9310	0,2906	0,0213	
	BC _a	0,9460	0,2903	0,0211	

Tabela 15: Probabilidade de cobertura com Algoritmo Genético para $m = 0,5$ no cenário 2.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,9280	0,3010	0,0892
		Pivotal	0,8620	0,2512	0,0663
		Percentil	0,9020	0,2512	0,0663
		BC	0,9000	0,2511	0,0665
		BC _a	0,8980	0,2529	0,0668
	25	Jackknife- t	0,9270	0,1765	0,0331
		Pivotal	0,9100	0,1653	0,0299
		Percentil	0,9200	0,1653	0,0299
		BC	0,9180	0,1655	0,0300
		BC _a	0,9180	0,1658	0,0301
	50	Jackknife- t	0,9380	0,1224	0,0160
		Pivotal	0,9220	0,1188	0,0158
		Percentil	0,9290	0,1188	0,0158
		BC	0,9310	0,1188	0,0158
		BC _a	0,9300	0,1189	0,0158
	100	Jackknife- t	0,9410	0,0856	0,0076
		Pivotal	0,9320	0,0840	0,0080
		Percentil	0,9430	0,0840	0,0080
		BC	0,9420	0,0840	0,0081
		BC _a	0,9410	0,0840	0,0081
500	Jackknife- t	0,9490	0,0380	0,0016	
	Pivotal	0,9480	0,0376	0,0022	
	Percentil	0,9460	0,0376	0,0022	
	BC	0,9460	0,0376	0,0022	
	BC _a	0,9460	0,0376	0,0022	

Tabela 16: Probabilidade de cobertura com Evolução Diferencial para $m = 0,5$ no cenário 2.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,9280	0,3010	0,0892
		Pivotal	0,8620	0,2512	0,0663
		Percentil	0,9020	0,2512	0,0663
		BC	0,9000	0,2511	0,0665
		BC _a	0,8980	0,2529	0,0668
	25	Jackknife- t	0,9270	0,1765	0,0331
		Pivotal	0,9100	0,1653	0,0299
		Percentil	0,9200	0,1653	0,0299
		BC	0,9180	0,1655	0,0300
		BC _a	0,9180	0,1658	0,0301
	50	Jackknife- t	0,9380	0,1224	0,0160
		Pivotal	0,9220	0,1188	0,0158
		Percentil	0,9290	0,1188	0,0158
		BC	0,9310	0,1188	0,0158
		BC _a	0,9300	0,1189	0,0158
	100	Jackknife- t	0,9410	0,0856	0,0076
		Pivotal	0,9320	0,0840	0,0080
		Percentil	0,9430	0,0840	0,0080
		BC	0,9420	0,0840	0,0081
		BC _a	0,9410	0,0840	0,0081
500	Jackknife- t	0,9490	0,0380	0,0016	
	Pivotal	0,9480	0,0376	0,0022	
	Percentil	0,9460	0,0376	0,0022	
	BC	0,9460	0,0376	0,0022	
	BC _a	0,9460	0,0376	0,0022	

Tabela 17: Probabilidade de cobertura com Algoritmo Genético para $b = 2$ no cenário 2.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,9580	3,4069	1,9246
		Pivotal	0,8720	3,2766	1,5962
		Percentil	0,9020	3,2766	1,5962
		BC	0,8210	2,4181	1,2889
		BC _a	0,8980	2,1877	1,0693
	25	Jackknife- t	0,9540	1,5494	0,4497
		Pivotal	0,9130	1,4993	0,4126
		Percentil	0,9200	1,4993	0,4126
		BC	0,8840	1,3537	0,3886
		BC _a	0,9180	1,3119	0,3635
	50	Jackknife- t	0,9550	1,0032	0,2025
		Pivotal	0,9330	0,9828	0,1949
		Percentil	0,9290	0,9828	0,1949
		BC	0,9180	0,9358	0,1885
		BC _a	0,9300	0,9238	0,1836
	100	Jackknife- t	0,9610	0,6805	0,0915
		Pivotal	0,9460	0,6682	0,0917
		Percentil	0,9430	0,6682	0,0917
		BC	0,9440	0,6528	0,0905
		BC _a	0,9410	0,6492	0,0906
500	Jackknife- t	0,9480	0,2953	0,0179	
	Pivotal	0,9440	0,2917	0,0215	
	Percentil	0,9460	0,2917	0,0215	
	BC	0,9310	0,2906	0,0213	
	BC _a	0,9460	0,2903	0,0211	

Tabela 18: Probabilidade de cobertura com Evolução Diferencial para $b = 2$ no cenário 2.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,9580	3,4069	1,9246
		Pivotal	0,8720	3,2766	1,5962
		Percentil	0,9020	3,2766	1,5962
		BC	0,8210	2,4181	1,2889
		BC _a	0,8980	2,1877	1,0693
	25	Jackknife- t	0,9540	1,5494	0,4497
		Pivotal	0,9130	1,4993	0,4126
		Percentil	0,9200	1,4993	0,4126
		BC	0,8840	1,3537	0,3886
		BC _a	0,9180	1,3119	0,3635
	50	Jackknife- t	0,9550	1,0032	0,2025
		Pivotal	0,9330	0,9828	0,1949
		Percentil	0,9290	0,9828	0,1949
		BC	0,9180	0,9358	0,1885
		BC _a	0,9300	0,9238	0,1836
	100	Jackknife- t	0,9610	0,6805	0,0915
		Pivotal	0,9460	0,6682	0,0917
		Percentil	0,9430	0,6682	0,0917
		BC	0,9440	0,6528	0,0905
		BC _a	0,9410	0,6492	0,0906
500	Jackknife- t	0,9480	0,2953	0,0179	
	Pivotal	0,9440	0,2917	0,0215	
	Percentil	0,9460	0,2917	0,0215	
	BC	0,9310	0,2906	0,0213	
	BC _a	0,9460	0,2903	0,0211	

Tabela 19: Probabilidade de cobertura com Algoritmo Genético para $m = 0,9$ no cenário 3.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,9170	0,1134	0,0446
		Pivotal	0,8750	0,0993	0,0382
		Percentil	0,9020	0,0993	0,0382
		BC	0,9000	0,0992	0,0382
		BC _a	0,8980	0,0999	0,0388
	25	Jackknife- t	0,9300	0,0646	0,0150
		Pivotal	0,9080	0,0617	0,0143
		Percentil	0,9200	0,0617	0,0143
		BC	0,9180	0,0619	0,0144
		BC _a	0,9190	0,0620	0,0145
	50	Jackknife- t	0,9370	0,0446	0,0075
		Pivotal	0,9230	0,0437	0,0076
		Percentil	0,9290	0,0437	0,0076
		BC	0,9310	0,0437	0,0076
		BC _a	0,9310	0,0438	0,0076
	100	Jackknife- t	0,9440	0,0310	0,0035
		Pivotal	0,9360	0,0305	0,0036
		Percentil	0,9430	0,0305	0,0036
		BC	0,9420	0,0305	0,0036
		BC _a	0,9410	0,0305	0,0036
500	Jackknife- t	0,9430	0,0137	0,0007	
	Pivotal	0,9440	0,0136	0,0009	
	Percentil	0,9460	0,0136	0,0009	
	BC	0,9460	0,0136	0,0009	
	BC _a	0,9460	0,0136	0,0009	

Tabela 20: Probabilidade de cobertura com Evolução Diferencial para $m = 0,9$ no cenário 3.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,9170	0,1134	0,0446
		Pivotal	0,8750	0,0993	0,0382
		Percentil	0,9020	0,0993	0,0382
		BC	0,9000	0,0992	0,0382
		BC _a	0,8980	0,0999	0,0388
	25	Jackknife- t	0,9300	0,0646	0,0150
		Pivotal	0,9080	0,0617	0,0143
		Percentil	0,9200	0,0617	0,0143
		BC	0,9180	0,0619	0,0144
		BC _a	0,9190	0,0620	0,0145
	50	Jackknife- t	0,9370	0,0446	0,0075
		Pivotal	0,9230	0,0437	0,0076
		Percentil	0,9290	0,0437	0,0076
		BC	0,9310	0,0437	0,0076
		BC _a	0,9310	0,0438	0,0076
	100	Jackknife- t	0,9440	0,0310	0,0035
		Pivotal	0,9360	0,0305	0,0036
		Percentil	0,9430	0,0305	0,0036
		BC	0,9420	0,0305	0,0036
		BC _a	0,9410	0,0305	0,0036
500	Jackknife- t	0,9430	0,0137	0,0007	
	Pivotal	0,9440	0,0136	0,0009	
	Percentil	0,9460	0,0136	0,0009	
	BC	0,9460	0,0136	0,0009	
	BC _a	0,9460	0,0136	0,0009	

Tabela 21: Probabilidade de cobertura com Algoritmo Genético para $b = 2$ no cenário 3.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,9580	3,4069	1,9246
		Pivotal	0,8720	3,2766	1,5962
		Percentil	0,9020	3,2766	1,5962
		BC	0,8210	2,4181	1,2889
		BC _a	0,8980	2,1877	1,0693
	25	Jackknife- t	0,9540	1,5494	0,4497
		Pivotal	0,9130	1,4993	0,4126
		Percentil	0,9200	1,4993	0,4126
		BC	0,8840	1,3537	0,3886
		BC _a	0,9190	1,3119	0,3635
	50	Jackknife- t	0,9550	1,0032	0,2025
		Pivotal	0,9330	0,9828	0,1949
		Percentil	0,9290	0,9828	0,1949
		BC	0,9180	0,9358	0,1885
		BC _a	0,9310	0,9238	0,1836
	100	Jackknife- t	0,9610	0,6805	0,0915
		Pivotal	0,9460	0,6682	0,0917
		Percentil	0,9430	0,6682	0,0917
		BC	0,9440	0,6528	0,0905
		BC _a	0,9410	0,6492	0,0906
500	Jackknife- t	0,9480	0,2953	0,0179	
	Pivotal	0,9440	0,2917	0,0215	
	Percentil	0,9460	0,2917	0,0215	
	BC	0,9310	0,2906	0,0213	
	BC _a	0,9460	0,2903	0,0211	

Tabela 22: Probabilidade de cobertura com Evolução Diferencial para $b = 2$ no cenário 3.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,9580	3,4069	1,9246
		Pivotal	0,8720	3,2766	1,5962
		Percentil	0,9020	3,2766	1,5962
		BC	0,8210	2,4181	1,2889
		BC _a	0,8980	2,1877	1,0693
	25	Jackknife- t	0,9540	1,5494	0,4497
		Pivotal	0,9130	1,4993	0,4126
		Percentil	0,9200	1,4993	0,4126
		BC	0,8840	1,3537	0,3886
		BC _a	0,9190	1,3119	0,3635
	50	Jackknife- t	0,9550	1,0032	0,2025
		Pivotal	0,9330	0,9828	0,1949
		Percentil	0,9290	0,9828	0,1949
		BC	0,9180	0,9358	0,1885
		BC _a	0,9310	0,9238	0,1836
	100	Jackknife- t	0,9610	0,6805	0,0915
		Pivotal	0,9460	0,6682	0,0917
		Percentil	0,9430	0,6682	0,0917
		BC	0,9440	0,6528	0,0905
		BC _a	0,9410	0,6492	0,0906
500	Jackknife- t	0,9480	0,2953	0,0179	
	Pivotal	0,9440	0,2917	0,0215	
	Percentil	0,9460	0,2917	0,0215	
	BC	0,9310	0,2906	0,0213	
	BC _a	0,9460	0,2903	0,0211	

Tabela 23: Probabilidade de cobertura com Algoritmo Genético para $m = 0, 2$ no cenário 4.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,8990	0,3916	0,1693
		Pivotal	0,8080	0,3301	0,1246
		Percentil	0,9020	0,3301	0,1246
		BC	0,9000	0,3304	0,1250
		BC _a	0,8980	0,3311	0,1261
	25	Jackknife- t	0,9180	0,2275	0,0639
		Pivotal	0,8800	0,2148	0,0579
		Percentil	0,9200	0,2148	0,0579
		BC	0,9180	0,2150	0,0581
		BC _a	0,9180	0,2151	0,0583
	50	Jackknife- t	0,9180	0,1564	0,0310
		Pivotal	0,9010	0,1525	0,0299
		Percentil	0,9290	0,1525	0,0299
		BC	0,9310	0,1526	0,0300
		BC _a	0,9300	0,1527	0,0300
	100	Jackknife- t	0,9360	0,1097	0,0153
		Pivotal	0,9210	0,1080	0,0155
		Percentil	0,9430	0,1080	0,0155
		BC	0,9420	0,1080	0,0156
		BC _a	0,9410	0,1080	0,0156
500	Jackknife- t	0,9480	0,0487	0,0030	
	Pivotal	0,9430	0,0482	0,0036	
	Percentil	0,9460	0,0482	0,0036	
	BC	0,9460	0,0482	0,0036	
	BC _a	0,9460	0,0482	0,0036	

Tabela 24: Probabilidade de cobertura com Evolução Diferencial para $m = 0,2$ no cenário 4.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,8990	0,3916	0,1693
		Pivotal	0,8080	0,3301	0,1246
		Percentil	0,9020	0,3301	0,1246
		BC	0,9000	0,3304	0,1250
		BC _a	0,8980	0,3311	0,1261
	25	Jackknife- t	0,9180	0,2275	0,0639
		Pivotal	0,8800	0,2148	0,0579
		Percentil	0,9200	0,2148	0,0579
		BC	0,9180	0,2150	0,0581
		BC _a	0,9180	0,2151	0,0583
	50	Jackknife- t	0,9180	0,1564	0,0310
		Pivotal	0,9010	0,1525	0,0299
		Percentil	0,9290	0,1525	0,0299
		BC	0,9310	0,1526	0,0300
		BC _a	0,9300	0,1527	0,0300
	100	Jackknife- t	0,9360	0,1097	0,0153
		Pivotal	0,9210	0,1080	0,0155
		Percentil	0,9430	0,1080	0,0155
		BC	0,9420	0,1080	0,0156
		BC _a	0,9410	0,1080	0,0156
500	Jackknife- t	0,9480	0,0487	0,0030	
	Pivotal	0,9430	0,0482	0,0036	
	Percentil	0,9460	0,0482	0,0036	
	BC	0,9460	0,0482	0,0036	
	BC _a	0,9460	0,0482	0,0036	

Tabela 25: Probabilidade de cobertura com Algoritmo Genético para $b = 1$ no cenário 4.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,9580	1,7034	0,9623
		Pivotal	0,8720	1,6383	0,7981
		Percentil	0,9020	1,6383	0,7981
		BC	0,8210	1,2090	0,6445
		BC _a	0,8980	1,0939	0,5347
	25	Jackknife- t	0,9540	0,7747	0,2249
		Pivotal	0,9130	0,7496	0,2063
		Percentil	0,9200	0,7496	0,2063
		BC	0,8840	0,6769	0,1943
		BC _a	0,9180	0,6559	0,1818
	50	Jackknife- t	0,9550	0,5016	0,1012
		Pivotal	0,9330	0,4914	0,0974
		Percentil	0,9290	0,4914	0,0974
		BC	0,9180	0,4679	0,0942
		BC _a	0,9300	0,4619	0,0918
	100	Jackknife- t	0,9610	0,3402	0,0457
		Pivotal	0,9460	0,3341	0,0458
		Percentil	0,9430	0,3341	0,0458
		BC	0,9440	0,3264	0,0452
		BC _a	0,9410	0,3246	0,0453
500	Jackknife- t	0,9480	0,1477	0,0089	
	Pivotal	0,9440	0,1458	0,0108	
	Percentil	0,9460	0,1458	0,0108	
	BC	0,9310	0,1453	0,0107	
	BC _a	0,9460	0,1452	0,0106	

Tabela 26: Probabilidade de cobertura com Evolução Diferencial para $b = 1$ no cenário 4.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,9580	1,7034	0,9623
		Pivotal	0,8720	1,6383	0,7981
		Percentil	0,9020	1,6383	0,7981
		BC	0,8210	1,2090	0,6445
		BC _a	0,8980	1,0939	0,5347
	25	Jackknife- t	0,9540	0,7747	0,2249
		Pivotal	0,9130	0,7496	0,2063
		Percentil	0,9200	0,7496	0,2063
		BC	0,8840	0,6769	0,1943
		BC _a	0,9180	0,6559	0,1818
	50	Jackknife- t	0,9550	0,5016	0,1012
		Pivotal	0,9330	0,4914	0,0974
		Percentil	0,9290	0,4914	0,0974
		BC	0,9180	0,4679	0,0942
		BC _a	0,9300	0,4619	0,0918
	100	Jackknife- t	0,9610	0,3402	0,0457
		Pivotal	0,9460	0,3341	0,0458
		Percentil	0,9430	0,3341	0,0458
		BC	0,9440	0,3264	0,0452
		BC _a	0,9410	0,3246	0,0453
500	Jackknife- t	0,9480	0,1477	0,0089	
	Pivotal	0,9440	0,1458	0,0108	
	Percentil	0,9460	0,1458	0,0108	
	BC	0,9310	0,1453	0,0107	
	BC _a	0,9460	0,1452	0,0106	

Tabela 27: Probabilidade de cobertura com Algoritmo Genético para $m = 0,5$ no cenário 5.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,8920	0,5675	0,1723
		Pivotal	0,7830	0,4468	0,1088
		Percentil	0,9020	0,4468	0,1088
		BC	0,9000	0,4465	0,1092
		BC _a	0,8950	0,4483	0,1091
	25	Jackknife- t	0,9110	0,3449	0,0663
		Pivotal	0,8710	0,3144	0,0552
		Percentil	0,9200	0,3144	0,0552
		BC	0,9180	0,3148	0,0553
		BC _a	0,9180	0,3153	0,0555
	50	Jackknife- t	0,9250	0,2419	0,0320
		Pivotal	0,8940	0,2314	0,0301
		Percentil	0,9290	0,2314	0,0301
		BC	0,9310	0,2315	0,0301
		BC _a	0,9300	0,2317	0,0301
	100	Jackknife- t	0,9380	0,1702	0,0152
		Pivotal	0,9260	0,1659	0,0157
		Percentil	0,9430	0,1659	0,0157
		BC	0,9420	0,1658	0,0158
		BC _a	0,9410	0,1659	0,0158
500	Jackknife- t	0,9460	0,0760	0,0031	
	Pivotal	0,9470	0,0750	0,0043	
	Percentil	0,9460	0,0750	0,0043	
	BC	0,9460	0,0750	0,0044	
	BC _a	0,9460	0,0750	0,0044	

Tabela 28: Probabilidade de cobertura com Evolução Diferencial para $m = 0,5$ no cenário 5.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,8920	0,5675	0,1723
		Pivotal	0,7830	0,4468	0,1088
		Percentil	0,9020	0,4468	0,1088
		BC	0,9000	0,4465	0,1092
		BC _a	0,8950	0,4483	0,1091
	25	Jackknife- t	0,9110	0,3449	0,0663
		Pivotal	0,8710	0,3144	0,0552
		Percentil	0,9200	0,3144	0,0552
		BC	0,9180	0,3148	0,0553
		BC _a	0,9180	0,3153	0,0555
	50	Jackknife- t	0,9250	0,2419	0,0320
		Pivotal	0,8940	0,2314	0,0301
		Percentil	0,9290	0,2314	0,0301
		BC	0,9310	0,2315	0,0301
		BC _a	0,9300	0,2317	0,0301
	100	Jackknife- t	0,9380	0,1702	0,0152
		Pivotal	0,9260	0,1659	0,0157
		Percentil	0,9430	0,1659	0,0157
		BC	0,9420	0,1658	0,0158
		BC _a	0,9410	0,1659	0,0158
500	Jackknife- t	0,9460	0,0760	0,0031	
	Pivotal	0,9470	0,0750	0,0043	
	Percentil	0,9460	0,0750	0,0043	
	BC	0,9460	0,0750	0,0044	
	BC _a	0,9460	0,0750	0,0044	

Tabela 29: Probabilidade de cobertura com Algoritmo Genético para $b = 1$ no cenário 5.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,9580	1,7034	0,9623
		Pivotal	0,8720	1,6383	0,7981
		Percentil	0,9020	1,6383	0,7981
		BC	0,8210	1,2090	0,6445
		BC _a	0,8950	1,0939	0,5347
	25	Jackknife- t	0,9540	0,7747	0,2249
		Pivotal	0,9130	0,7496	0,2063
		Percentil	0,9200	0,7496	0,2063
		BC	0,8840	0,6769	0,1943
		BC _a	0,9180	0,6559	0,1818
	50	Jackknife- t	0,9550	0,5016	0,1012
		Pivotal	0,9330	0,4914	0,0974
		Percentil	0,9290	0,4914	0,0974
		BC	0,9180	0,4679	0,0942
		BC _a	0,9300	0,4619	0,0918
	100	Jackknife- t	0,9610	0,3402	0,0457
		Pivotal	0,9460	0,3341	0,0458
		Percentil	0,9430	0,3341	0,0458
		BC	0,9440	0,3264	0,0452
		BC _a	0,9410	0,3246	0,0453
500	Jackknife- t	0,9480	0,1477	0,0089	
	Pivotal	0,9440	0,1458	0,0108	
	Percentil	0,9460	0,1458	0,0108	
	BC	0,9310	0,1453	0,0107	
	BC _a	0,9460	0,1452	0,0106	

Tabela 30: Probabilidade de cobertura com Evolução Diferencial para $b = 1$ no cenário 5.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,9580	1,7034	0,9623
		Pivotal	0,8720	1,6383	0,7981
		Percentil	0,9020	1,6383	0,7981
		BC	0,8210	1,2090	0,6445
		BC _a	0,8950	1,0939	0,5347
	25	Jackknife- t	0,9540	0,7747	0,2249
		Pivotal	0,9130	0,7496	0,2063
		Percentil	0,9200	0,7496	0,2063
		BC	0,8840	0,6769	0,1943
		BC _a	0,9180	0,6559	0,1818
	50	Jackknife- t	0,9550	0,5016	0,1012
		Pivotal	0,9330	0,4914	0,0974
		Percentil	0,9290	0,4914	0,0974
		BC	0,9180	0,4679	0,0942
		BC _a	0,9300	0,4619	0,0918
	100	Jackknife- t	0,9610	0,3402	0,0457
		Pivotal	0,9460	0,3341	0,0458
		Percentil	0,9430	0,3341	0,0458
		BC	0,9440	0,3264	0,0452
		BC _a	0,9410	0,3246	0,0453
500	Jackknife- t	0,9480	0,1477	0,0089	
	Pivotal	0,9440	0,1458	0,0108	
	Percentil	0,9460	0,1458	0,0108	
	BC	0,9310	0,1453	0,0107	
	BC _a	0,9460	0,1452	0,0106	

Tabela 31: Probabilidade de cobertura com Algoritmo Genético para $m = 0,9$ no cenário 6.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,8940	0,2393	0,1331
		Pivotal	0,8040	0,2160	0,1096
		Percentil	0,9020	0,2160	0,1096
		BC	0,9000	0,2156	0,1097
		BC _a	0,8980	0,2160	0,1108
	25	Jackknife- t	0,9100	0,1321	0,0439
		Pivotal	0,8680	0,1290	0,0424
		Percentil	0,9200	0,1290	0,0424
		BC	0,9180	0,1294	0,0427
		BC _a	0,9190	0,1295	0,0429
	50	Jackknife- t	0,9290	0,0907	0,0223
		Pivotal	0,9100	0,0900	0,0224
		Percentil	0,9290	0,0900	0,0224
		BC	0,9310	0,0899	0,0223
		BC _a	0,9310	0,0900	0,0224
	100	Jackknife- t	0,9350	0,0623	0,0102
		Pivotal	0,9250	0,0618	0,0103
		Percentil	0,9430	0,0618	0,0103
		BC	0,9420	0,0618	0,0104
		BC _a	0,9410	0,0618	0,0104
500	Jackknife- t	0,9430	0,0274	0,0021	
	Pivotal	0,9450	0,0272	0,0023	
	Percentil	0,9460	0,0272	0,0023	
	BC	0,9460	0,0272	0,0023	
	BC _a	0,9460	0,0272	0,0023	

Tabela 32: Probabilidade de cobertura com Evolução Diferencial para $m = 0,9$ no cenário 6.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,8940	0,2393	0,1331
		Pivotal	0,8040	0,2160	0,1096
		Percentil	0,9020	0,2160	0,1096
		BC	0,9000	0,2156	0,1097
		BC _a	0,8980	0,2160	0,1108
	25	Jackknife- t	0,9100	0,1321	0,0439
		Pivotal	0,8680	0,1290	0,0424
		Percentil	0,9200	0,1290	0,0424
		BC	0,9180	0,1294	0,0427
		BC _a	0,9190	0,1295	0,0429
	50	Jackknife- t	0,9290	0,0907	0,0223
		Pivotal	0,9100	0,0900	0,0224
		Percentil	0,9290	0,0900	0,0224
		BC	0,9310	0,0899	0,0223
		BC _a	0,9310	0,0900	0,0224
	100	Jackknife- t	0,9350	0,0623	0,0102
		Pivotal	0,9250	0,0618	0,0103
		Percentil	0,9430	0,0618	0,0103
		BC	0,9420	0,0618	0,0104
		BC _a	0,9410	0,0618	0,0104
500	Jackknife- t	0,9430	0,0274	0,0021	
	Pivotal	0,9450	0,0272	0,0023	
	Percentil	0,9460	0,0272	0,0023	
	BC	0,9460	0,0272	0,0023	
	BC _a	0,9460	0,0272	0,0023	

Tabela 33: Probabilidade de cobertura com Algoritmo Genético para $b = 1$ no cenário 6.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,9580	1,7034	0,9623
		Pivotal	0,8720	1,6383	0,7981
		Percentil	0,9020	1,6383	0,7981
		BC	0,8210	1,2090	0,6445
		BC _a	0,8980	1,0939	0,5347
	25	Jackknife- t	0,9540	0,7747	0,2249
		Pivotal	0,9130	0,7496	0,2063
		Percentil	0,9200	0,7496	0,2063
		BC	0,8840	0,6769	0,1943
		BC _a	0,9190	0,6559	0,1818
	50	Jackknife- t	0,9550	0,5016	0,1012
		Pivotal	0,9330	0,4914	0,0974
		Percentil	0,9290	0,4914	0,0974
		BC	0,9180	0,4679	0,0942
		BC _a	0,9310	0,4619	0,0918
	100	Jackknife- t	0,9610	0,3402	0,0457
		Pivotal	0,9460	0,3341	0,0458
		Percentil	0,9430	0,3341	0,0458
		BC	0,9440	0,3264	0,0452
		BC _a	0,9410	0,3246	0,0453
500	Jackknife- t	0,9480	0,1477	0,0089	
	Pivotal	0,9440	0,1458	0,0108	
	Percentil	0,9460	0,1458	0,0108	
	BC	0,9310	0,1453	0,0107	
	BC _a	0,9460	0,1452	0,0106	

Tabela 34: Probabilidade de cobertura com Evolução Diferencial para $b = 1$ no cenário 6.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,9580	1,7034	0,9623
		Pivotal	0,8720	1,6383	0,7981
		Percentil	0,9020	1,6383	0,7981
		BC	0,8210	1,2090	0,6445
		BC _a	0,8980	1,0939	0,5347
	25	Jackknife- t	0,9540	0,7747	0,2249
		Pivotal	0,9130	0,7496	0,2063
		Percentil	0,9200	0,7496	0,2063
		BC	0,8840	0,6769	0,1943
		BC _a	0,9190	0,6559	0,1818
	50	Jackknife- t	0,9550	0,5016	0,1012
		Pivotal	0,9330	0,4914	0,0974
		Percentil	0,9290	0,4914	0,0974
		BC	0,9180	0,4679	0,0942
		BC _a	0,9310	0,4619	0,0918
	100	Jackknife- t	0,9610	0,3402	0,0457
		Pivotal	0,9460	0,3341	0,0458
		Percentil	0,9430	0,3341	0,0458
		BC	0,9440	0,3264	0,0452
		BC _a	0,9410	0,3246	0,0453
500	Jackknife- t	0,9480	0,1477	0,0089	
	Pivotal	0,9440	0,1458	0,0108	
	Percentil	0,9460	0,1458	0,0108	
	BC	0,9310	0,1453	0,0107	
	BC _a	0,9460	0,1452	0,0106	

Tabela 35: Probabilidade de cobertura com Algoritmo Genético para $m = 0, 2$ no cenário 7.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,8030	0,7397	0,4070
		Pivotal	0,6400	0,5711	0,2117
		Percentil	0,9020	0,5711	0,2117
		BC	0,9000	0,5713	0,2123
		BC _a	0,8940	0,5669	0,2146
	25	Jackknife- t	0,8690	0,4511	0,1821
		Pivotal	0,7760	0,4132	0,1361
		Percentil	0,9200	0,4132	0,1361
		BC	0,9180	0,4133	0,1367
		BC _a	0,9170	0,4125	0,1374
	50	Jackknife- t	0,8920	0,3106	0,0949
		Pivotal	0,8360	0,3006	0,0823
		Percentil	0,9290	0,3006	0,0823
		BC	0,9310	0,3010	0,0823
		BC _a	0,9300	0,3009	0,0825
	100	Jackknife- t	0,9260	0,2193	0,0487
		Pivotal	0,8710	0,2157	0,0459
		Percentil	0,9430	0,2157	0,0459
		BC	0,9420	0,2157	0,0463
		BC _a	0,9410	0,2156	0,0463
500	Jackknife- t	0,9460	0,0975	0,0098	
	Pivotal	0,9260	0,0964	0,0105	
	Percentil	0,9460	0,0964	0,0105	
	BC	0,9460	0,0965	0,0105	
	BC _a	0,9460	0,0965	0,0105	

Tabela 36: Probabilidade de cobertura com Evolução Diferencial para $m = 0,2$ no cenário 7.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,8030	0,7397	0,4070
		Pivotal	0,6400	0,5711	0,2117
		Percentil	0,9020	0,5711	0,2117
		BC	0,9000	0,5713	0,2123
		BC _a	0,8940	0,5669	0,2146
	25	Jackknife- t	0,8690	0,4511	0,1821
		Pivotal	0,7760	0,4132	0,1361
		Percentil	0,9200	0,4132	0,1361
		BC	0,9180	0,4133	0,1367
		BC _a	0,9170	0,4125	0,1374
	50	Jackknife- t	0,8920	0,3106	0,0949
		Pivotal	0,8360	0,3006	0,0823
		Percentil	0,9290	0,3006	0,0823
		BC	0,9310	0,3010	0,0823
		BC _a	0,9300	0,3009	0,0825
	100	Jackknife- t	0,9260	0,2193	0,0487
		Pivotal	0,8710	0,2157	0,0459
		Percentil	0,9430	0,2157	0,0459
		BC	0,9420	0,2157	0,0463
		BC _a	0,9410	0,2156	0,0463
500	Jackknife- t	0,9460	0,0975	0,0098	
	Pivotal	0,9260	0,0964	0,0105	
	Percentil	0,9460	0,0964	0,0105	
	BC	0,9460	0,0965	0,0105	
	BC _a	0,9460	0,0965	0,0105	

Tabela 37: Probabilidade de cobertura com Algoritmo Genético para $b = 0,5$ no cenário 7.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,9580	0,8517	0,4811
		Pivotal	0,8720	0,8191	0,3991
		Percentil	0,9020	0,8191	0,3991
		BC	0,8210	0,6046	0,3222
		BC _a	0,8940	0,5470	0,2673
	25	Jackknife- t	0,9540	0,3873	0,1124
		Pivotal	0,9130	0,3748	0,1031
		Percentil	0,9200	0,3748	0,1031
		BC	0,8840	0,3384	0,0971
		BC _a	0,9170	0,3280	0,0909
	50	Jackknife- t	0,9550	0,2508	0,0506
		Pivotal	0,9330	0,2457	0,0487
		Percentil	0,9290	0,2457	0,0487
		BC	0,9180	0,2340	0,0471
		BC _a	0,9300	0,2310	0,0459
	100	Jackknife- t	0,9610	0,1701	0,0229
		Pivotal	0,9460	0,1671	0,0229
		Percentil	0,9430	0,1671	0,0229
		BC	0,9440	0,1632	0,0226
		BC _a	0,9410	0,1623	0,0227
500	Jackknife- t	0,9480	0,0738	0,0045	
	Pivotal	0,9440	0,0729	0,0054	
	Percentil	0,9460	0,0729	0,0054	
	BC	0,9310	0,0727	0,0053	
	BC _a	0,9460	0,0726	0,0053	

Tabela 38: Probabilidade de cobertura com Evolução Diferencial para $b = 0,5$ no cenário 7.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,9580	0,8517	0,4811
		Pivotal	0,8720	0,8191	0,3991
		Percentil	0,9020	0,8191	0,3991
		BC	0,8210	0,6046	0,3222
		BC _a	0,8940	0,5470	0,2673
	25	Jackknife- t	0,9540	0,3873	0,1124
		Pivotal	0,9130	0,3748	0,1031
		Percentil	0,9200	0,3748	0,1031
		BC	0,8840	0,3384	0,0971
		BC _a	0,9170	0,3280	0,0909
	50	Jackknife- t	0,9550	0,2508	0,0506
		Pivotal	0,9330	0,2457	0,0487
		Percentil	0,9290	0,2457	0,0487
		BC	0,9180	0,2340	0,0471
		BC _a	0,9300	0,2310	0,0459
	100	Jackknife- t	0,9610	0,1701	0,0229
		Pivotal	0,9460	0,1671	0,0229
		Percentil	0,9430	0,1671	0,0229
		BC	0,9440	0,1632	0,0226
		BC _a	0,9410	0,1623	0,0227
500	Jackknife- t	0,9480	0,0738	0,0045	
	Pivotal	0,9440	0,0729	0,0054	
	Percentil	0,9460	0,0729	0,0054	
	BC	0,9310	0,0727	0,0053	
	BC _a	0,9460	0,0726	0,0053	

Tabela 39: Probabilidade de cobertura com Algoritmo Genético para $m = 0,5$ no cenário 8.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,7840	0,9566	0,3712
		Pivotal	0,6000	0,6701	0,1594
		Percentil	0,9020	0,6701	0,1594
		BC	0,9000	0,6693	0,1598
		BC _a	0,8920	0,6665	0,1614
	25	Jackknife- t	0,8630	0,6370	0,1479
		Pivotal	0,7600	0,5379	0,0961
		Percentil	0,9200	0,5379	0,0961
		BC	0,9180	0,5383	0,0961
		BC _a	0,9170	0,5383	0,0964
	50	Jackknife- t	0,8860	0,4631	0,0705
		Pivotal	0,8350	0,4230	0,0565
		Percentil	0,9290	0,4230	0,0565
		BC	0,9310	0,4232	0,0564
		BC _a	0,9300	0,4233	0,0565
	100	Jackknife- t	0,9230	0,3330	0,0324
		Pivotal	0,8940	0,3164	0,0308
		Percentil	0,9430	0,3164	0,0308
		BC	0,9420	0,3162	0,0309
		BC _a	0,9410	0,3163	0,0309
500	Jackknife- t	0,9420	0,1513	0,0063	
	Pivotal	0,9380	0,1485	0,0085	
	Percentil	0,9460	0,1485	0,0085	
	BC	0,9460	0,1486	0,0085	
	BC _a	0,9460	0,1486	0,0086	

Tabela 40: Probabilidade de cobertura com Evolução Diferencial para $m = 0,5$ no cenário 8.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,7840	0,9566	0,3712
		Pivotal	0,6000	0,6701	0,1594
		Percentil	0,9020	0,6701	0,1594
		BC	0,9000	0,6693	0,1598
		BC _a	0,8920	0,6665	0,1614
	25	Jackknife- t	0,8630	0,6370	0,1479
		Pivotal	0,7600	0,5379	0,0961
		Percentil	0,9200	0,5379	0,0961
		BC	0,9180	0,5383	0,0961
		BC _a	0,9170	0,5383	0,0964
	50	Jackknife- t	0,8860	0,4631	0,0705
		Pivotal	0,8350	0,4230	0,0565
		Percentil	0,9290	0,4230	0,0565
		BC	0,9310	0,4232	0,0564
		BC _a	0,9300	0,4233	0,0565
	100	Jackknife- t	0,9230	0,3330	0,0324
		Pivotal	0,8940	0,3164	0,0308
		Percentil	0,9430	0,3164	0,0308
		BC	0,9420	0,3162	0,0309
		BC _a	0,9410	0,3163	0,0309
500	Jackknife- t	0,9420	0,1513	0,0063	
	Pivotal	0,9380	0,1485	0,0085	
	Percentil	0,9460	0,1485	0,0085	
	BC	0,9460	0,1486	0,0085	
	BC _a	0,9460	0,1486	0,0086	

Tabela 41: Probabilidade de cobertura com Algoritmo Genético para $b = 0,5$ no cenário 8.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,9580	0,8517	0,4811
		Pivotal	0,8720	0,8192	0,3991
		Percentil	0,9020	0,8192	0,3991
		BC	0,8210	0,6045	0,3222
		BC _a	0,8920	0,5469	0,2673
	25	Jackknife- t	0,9540	0,3873	0,1124
		Pivotal	0,9130	0,3748	0,1031
		Percentil	0,9200	0,3748	0,1031
		BC	0,8840	0,3384	0,0971
		BC _a	0,9170	0,3280	0,0909
	50	Jackknife- t	0,9550	0,2508	0,0506
		Pivotal	0,9330	0,2457	0,0487
		Percentil	0,9290	0,2457	0,0487
		BC	0,9180	0,2340	0,0471
		BC _a	0,9300	0,2310	0,0459
	100	Jackknife- t	0,9610	0,1701	0,0229
		Pivotal	0,9460	0,1671	0,0229
		Percentil	0,9430	0,1671	0,0229
		BC	0,9440	0,1632	0,0226
		BC _a	0,9410	0,1623	0,0227
500	Jackknife- t	0,9480	0,0738	0,0045	
	Pivotal	0,9440	0,0729	0,0054	
	Percentil	0,9460	0,0729	0,0054	
	BC	0,9310	0,0727	0,0053	
	BC _a	0,9460	0,0726	0,0053	

Tabela 42: Probabilidade de cobertura com Evolução Diferencial para $b = 0,5$ no cenário 8.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,9580	0,8517	0,4811
		Pivotal	0,8720	0,8192	0,3991
		Percentil	0,9020	0,8192	0,3991
		BC	0,8210	0,6045	0,3222
		BC _a	0,8920	0,5469	0,2673
	25	Jackknife- t	0,9540	0,3873	0,1124
		Pivotal	0,9130	0,3748	0,1031
		Percentil	0,9200	0,3748	0,1031
		BC	0,8840	0,3384	0,0971
		BC _a	0,9170	0,3280	0,0909
	50	Jackknife- t	0,9550	0,2508	0,0506
		Pivotal	0,9330	0,2457	0,0487
		Percentil	0,9290	0,2457	0,0487
		BC	0,9180	0,2340	0,0471
		BC _a	0,9300	0,2310	0,0459
	100	Jackknife- t	0,9610	0,1701	0,0229
		Pivotal	0,9460	0,1671	0,0229
		Percentil	0,9430	0,1671	0,0229
		BC	0,9440	0,1632	0,0226
		BC _a	0,9410	0,1623	0,0227
500	Jackknife- t	0,9480	0,0738	0,0045	
	Pivotal	0,9440	0,0729	0,0054	
	Percentil	0,9460	0,0729	0,0054	
	BC	0,9310	0,0727	0,0053	
	BC _a	0,9460	0,0726	0,0053	

Tabela 43: Probabilidade de cobertura com Algoritmo Genético para $m = 0,9$ no cenário 9.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,8190	0,5273	0,3958
		Pivotal	0,7040	0,4540	0,2335
		Percentil	0,9020	0,4540	0,2335
		BC	0,9000	0,45301	0,234
		BC _a	0,8950	0,4478	0,2360
	25	Jackknife- t	0,8630	0,2822	0,1508
		Pivotal	0,7860	0,2849	0,1303
		Percentil	0,9200	0,2849	0,1303
		BC	0,9180	0,2859	0,1311
		BC _a	0,9190	0,2851	0,1317
	50	Jackknife- t	0,9020	0,1900	0,0780
		Pivotal	0,8510	0,1943	0,0742
		Percentil	0,9290	0,1943	0,0742
		BC	0,9310	0,1942	0,0742
		BC _a	0,9310	0,1940	0,0743
	100	Jackknife- t	0,9220	0,1273	0,0364
		Pivotal	0,8880	0,1288	0,0357
		Percentil	0,9430	0,1288	0,0357
		BC	0,9420	0,1289	0,0361
		BC _a	0,9410	0,1288	0,0361
500	Jackknife- t	0,9470	0,0551	0,0072	
	Pivotal	0,9310	0,0548	0,0075	
	Percentil	0,9460	0,0548	0,0075	
	BC	0,9460	0,0548	0,0076	
	BC _a	0,9460	0,0548	0,0076	

Tabela 44: Probabilidade de cobertura com Evolução Diferencial para $m = 0,9$ no cenário 9.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,8190	0,5273	0,3958
		Pivotal	0,7040	0,4540	0,2335
		Percentil	0,9020	0,4540	0,2335
		BC	0,9000	0,45301	0,234
		BC _a	0,8950	0,4478	0,2360
	25	Jackknife- t	0,8630	0,2822	0,1508
		Pivotal	0,7860	0,2849	0,1303
		Percentil	0,9200	0,2849	0,1303
		BC	0,9180	0,2859	0,1311
		BC _a	0,9190	0,2851	0,1317
	50	Jackknife- t	0,9020	0,1900	0,0780
		Pivotal	0,8510	0,1943	0,0742
		Percentil	0,9290	0,1943	0,0742
		BC	0,9310	0,1942	0,0742
		BC _a	0,9310	0,1940	0,0743
	100	Jackknife- t	0,9220	0,1273	0,0364
		Pivotal	0,8880	0,1288	0,0357
		Percentil	0,9430	0,1288	0,0357
		BC	0,9420	0,1289	0,0361
		BC _a	0,9410	0,1288	0,0361
500	Jackknife- t	0,9470	0,0551	0,0072	
	Pivotal	0,9310	0,0548	0,0075	
	Percentil	0,9460	0,0548	0,0075	
	BC	0,9460	0,0548	0,0076	
	BC _a	0,9460	0,0548	0,0076	

Tabela 45: Probabilidade de cobertura com Algoritmo Genético para $b = 0,5$ no cenário 9.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
GA	10	Jackknife- t	0,9580	0,8517	0,4811
		Pivotal	0,8720	0,8189	0,3993
		Percentil	0,9020	0,8189	0,3993
		BC	0,8210	0,6046	0,3223
		BC _a	0,8950	0,5470	0,2674
	25	Jackknife- t	0,9540	0,3873	0,1124
		Pivotal	0,9130	0,3748	0,1031
		Percentil	0,9200	0,3748	0,1031
		BC	0,8840	0,3384	0,0971
		BC _a	0,9190	0,3280	0,0909
	50	Jackknife- t	0,9550	0,2508	0,0506
		Pivotal	0,9330	0,2457	0,0487
		Percentil	0,9290	0,2457	0,0487
		BC	0,9180	0,2340	0,0471
		BC _a	0,9310	0,2310	0,0459
	100	Jackknife- t	0,9610	0,1701	0,0229
		Pivotal	0,9460	0,1671	0,0229
		Percentil	0,9430	0,1671	0,0229
		BC	0,9440	0,1632	0,0226
		BC _a	0,9410	0,1623	0,0227
500	Jackknife- t	0,9480	0,0738	0,0045	
	Pivotal	0,9440	0,0729	0,0054	
	Percentil	0,9460	0,0729	0,0054	
	BC	0,9310	0,0727	0,0053	
	BC _a	0,9460	0,0726	0,0053	

Tabela 46: Probabilidade de cobertura com Evolução Diferencial para $b = 0,5$ no cenário 9.

Algoritmo	Amostra	Intervalo	PC	MComp	DPComp
DE	10	Jackknife- t	0,9580	0,8517	0,4811
		Pivotal	0,8720	0,8189	0,3993
		Percentil	0,9020	0,8189	0,3993
		BC	0,8210	0,6046	0,3223
		BC _a	0,8950	0,5470	0,2674
	25	Jackknife- t	0,9540	0,3873	0,1124
		Pivotal	0,9130	0,3748	0,1031
		Percentil	0,9200	0,3748	0,1031
		BC	0,8840	0,3384	0,0971
		BC _a	0,9190	0,3280	0,0909
	50	Jackknife- t	0,9550	0,2508	0,0506
		Pivotal	0,9330	0,2457	0,0487
		Percentil	0,9290	0,2457	0,0487
		BC	0,9180	0,2340	0,0471
		BC _a	0,9310	0,2310	0,0459
	100	Jackknife- t	0,9610	0,1701	0,0229
		Pivotal	0,9460	0,1671	0,0229
		Percentil	0,9430	0,1671	0,0229
		BC	0,9440	0,1632	0,0226
		BC _a	0,9410	0,1623	0,0227
500	Jackknife- t	0,9480	0,0738	0,0045	
	Pivotal	0,9440	0,0729	0,0054	
	Percentil	0,9460	0,0729	0,0054	
	BC	0,9310	0,0727	0,0053	
	BC _a	0,9460	0,0726	0,0053	

5 APLICAÇÃO EM DADOS REAIS

5.1 Introdução

Neste capítulo utiliza-se dois conjuntos de dados reais para ilustrar as metodologias discutidas até aqui, assim como avaliar a qualidade do modelo de regressão L-logística com relação aos tradicionais modelo de regressão Beta e Simplex apresentados em Ferrari e Cribari-Neto (2004) e Miyashiro (2008). As estimativas de máxima verossimilhança dos modelos de regressão (L-logística, Beta e Simplex) são obtidas através dos algoritmos GA e DE. As estimativas de variabilidade são obtidas por meio dos métodos de reamostragem Jackknife e Bootstrap. O primeiro conjunto apresenta dados referentes a taxa de problemas cardíacos nos municípios dos Estados Unidos. O segundo conjunto apresenta dados com foco no índice de percepção da corrupção de 2016.

Assim como na seção 4, para a modelagem dos dois conjuntos de dados, utiliza-se os algoritmos GA e DE por meio da função `ga` do pacote `GA` e da função `DEoptim` do pacote `DEoptim` que estão na versão 3.3.3 do software R. Com relação ao algoritmo Bootstrap, utiliza-se 1000 reamostras para a obtenção da estimativa do erro padrão. A seguir, na Tabela 47, a configuração das metaheurísticas.

Tabela 47: Configuração dos Algoritmos.

Metaheurística	N_p	N_G	P_r	p_m	F
GA	5000	300	0,85	0,001	—
DE	5000	300	0,85	—	0,5

5.2 Aplicação I

Nesta primeira aplicação utiliza-se dados referentes a problemas cardíacos em adultos maiores de 35 anos nos estados dos Estados Unidos. Utiliza-se sexo e etnia como variáveis explicativas (covariáveis), enquanto que a taxa de problemas cardíacos é utilizada como variável resposta nos modelos. O conjunto de dados também apresenta a taxa de problemas cardíacos expressa em valores por 10000 habitantes. Contudo, para a utilização dos modelos, as informações são transformadas para valores percententes ao intervalo $(0, 1)$. Estes dados podem ser obtidos em <https://catalog.data.gov/dataset>. A seguir, nas Figuras 11 e 12, apresenta-se o gráfico de caixa (boxplot) para a taxa em função das covariáveis. Em cada gráfico percebe-se o desnível entre as categorias, o caráter assimétrico (não muito acentuado) e a presença de observações discrepantes.

Figura 11: Gráfico de caixa da taxa de doenças cardíacas segundo a variável sexo.

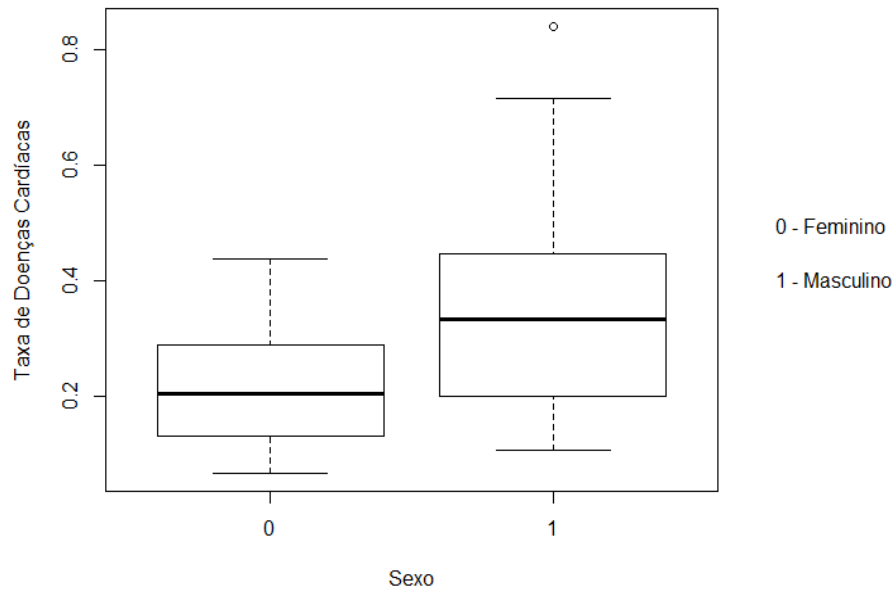
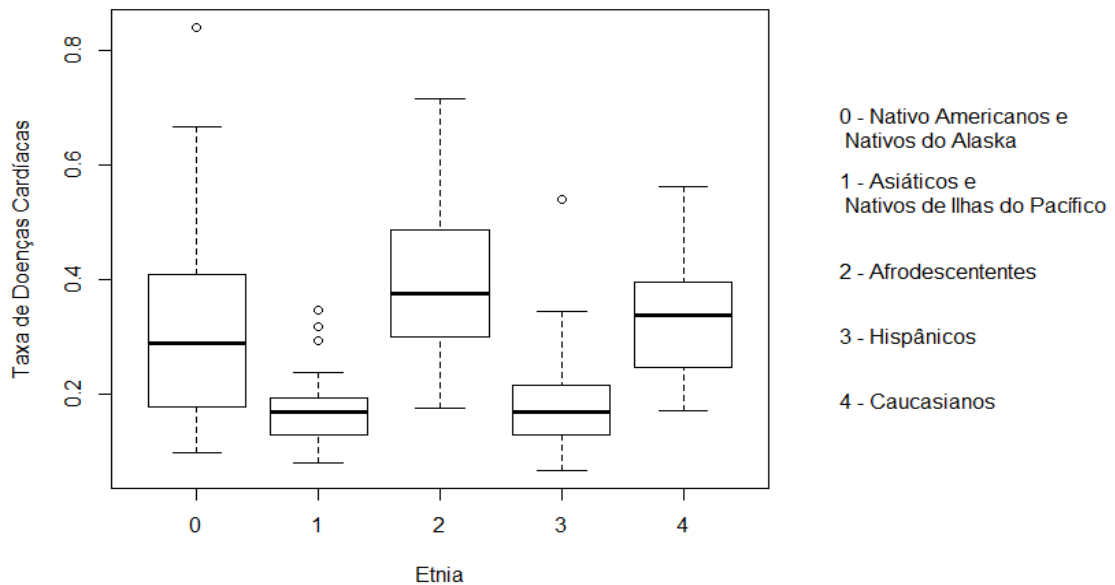


Figura 12: Gráfico de caixa da taxa de doenças cardíacas segundo a variável etnia.



5.2.1 Modelos de Regressão: Resultados e Conclusões

Nesta seção, como cenário inicial, utiliza-se as funções de ligação Logit para ilustrar o desempenho dos modelos com relação aos dados. Tendo em vista a natureza qualitativa das covariáveis, utiliza-se a metodologia das variáveis Dummy para o ajuste dos modelos. Como apresentado na Figura 11, a covariável sexo é codificada por meio de uma variável binária no qual 0 representa o sexo feminino e 1 representa o sexo masculino. Por sua vez, a covariável etnia é codificada por meio de 4 variáveis binárias, no qual a etnia Nativos Americanos e Nativos do Alaska é utilizada como categoria de referência. Além

disso, todos modelos apresentam o seguinte preditor linear

$$\eta_i = \beta_0 + \beta_1 x_{11,i} + \beta_2 x_{21,i} + \beta_3 x_{22,i} + \beta_4 x_{23,i} + \beta_5 x_{24,i}, \quad (50)$$

cujas variáveis são descritas na Tabela 48 a seguir.

Tabela 48: Notação das covariáveis nos modelos.

Variável Dummy	Categorias
x_{11}	Sexo: Masculino
x_{21}	Etnia: Asiáticos e Nativos de ilhas do Pacífico
x_{22}	Etnia: Afrodescendentes
x_{23}	Etnia: Hispânicos
x_{24}	Etnia: Caucasianos

Para avaliar a significância dos parâmetros, utilizou-se inicialmente os testes de Wald com variância Bootstrap e Jackknife, respectivamente. Constatou-se que para uma precisão de 4 casas decimais, as metaheurísticas apresentaram o mesmo valor para as estimativas de máxima verossimilhança. Com relação às estimativas do desvio padrão, constatou-se que as metaheurísticas forneceram os mesmos resultados para cada método de amostragem. Além disso, constatou-se certa proximidade entre as estimativas do método Bootstrap e do método Jackknife. Ainda na ideia do teste do tipo Wald, para os modelos de regressão L-logística e Simplex, constatou-se que a categoria Caucasianos (variável Dummy x_{24}) da covariável Etnia demonstrou ser não significativa para um nível de significância fixado de 0,05 e um nível descritivo de 0,1147 para o modelo de regressão L-logística e um nível descritivo de 0,1204 para o modelo de regressão Simplex. A seguir, nas Tabelas 49 a 50, os resultados do teste tipo Wald para o método Bootstrap e as metehaurísticas. Logo após, nas Tabelas 51 e 52, os resultados para o método Jackknife.

Para constatar os resultados do teste tipo Wald, utilizou-se o teste da razão de verossimilhanças generalizada para verificar a significância do parâmetro β_5 associado a categoria Caucasianos. Contudo, para todos os modelos, constatou-se a significância da categoria para um nível fixado de 0,05 e os níveis descritivos 0,0096, 0,0009 e 0,0415 para os modelos de regressão L-logística, Beta e Simplex. A seguir, na Tabela 53, o resultado do teste da razão de verossimilhanças generalizada.

Para comparar a qualidade do ajuste dos modelos, utilizou-se o coeficiente de determinação, R^2 , do coeficiente de determinação ajustado (R_a^2), do critério de Akaike (AIC), do critério Bayesiano (BIC) e do valor negativo do logaritmo da verossimilhança, $-l(\theta)$. Dado a divergência dos testes de significância, decidiu-se medir a qualidade do ajuste para ambos os casos. Assim, sobre a perspectiva dos resultados do teste de Wald, os dados da categoria Caucasianos foram agrupados com a categoria de referência para estimar os valores de máxima verossimilhança dos modelos regressão L-logística e Simplex. Os resultados podem ser vistos nas Tabelas 54 a 57, no qual as metaheurísticas apresentaram as mesmas estimativas para uma precisão de 4 casas decimais.

De posse de ambos os casos, calculou-se as medidas de qualidade do ajuste, no qual constatou-se que o modelo de regressão L-logística apresentava o melhor ajuste para estes dados. Contudo, houve divergência entre as medidas no caso do teste de Wald, no qual as medidas R^2 e R_a^2 apontavam o modelo de regressão Beta como o melhor ajustado. No entanto, vale ressaltar que neste cenário o modelo de regressão Beta possui uma covariável a mais do que os outros e isto pode ser visto como um possível motivo para qual este modelo demonstrou melhor qualidade de ajuste nestas medidas. A seguir, nas Tabelas 60 e 61, apresenta-se todos os intervalos de 95% de confiança propostos na seção 4, para o modelo de regressão L-logística nos casos do teste de Wald e do teste da razão de verossimilhanças generalizada. Vale ressaltar que para a precisão de quatro casas decimais, não encontrou-se diferença nos valores dos

intervalos apresentados pelas metaheurísticas. No que diz respeito à identificação de observações atípicas, nas Figuras 13 e 16, apresenta-se os resultados da distância do tipo Cook para os modelos de regressão, visando os cenários com e sem a categoria Caucasianos (perspectivas do teste da razão de verossimilhanças generalizada e do teste do tipo Wald). Pode-se concluir que para o cenário com todas as categorias, os modelos de regressão Beta e Simplex apresentaram a observação 151 como atípica ao conjunto de dados. Já no cenário sem a covariável Caucasianos, nota-se uma redução nas distâncias do modelo de regressão L-logística e Simplex, mas o modelo Simplex apresentou as observações 100, 101 e 151 como atípicas. Vale comentar que essa redução na distância das observações indica uma melhora na modelagem e que a observação 151 representa os indivíduos do estado do Kansas que apresentam o sexo masculino e a etnia correspondente a categoria Nativo Americanos e Nativos do Alaska. Ainda visando a observação 151 (estado do Kansas), verifica-se que indivíduos com essa característica de sexo e etnia possuem taxa de doença cardíaca no valor de 0,8395, enquanto que indivíduos que estão em outros estados e possuem as mesmas características, apresentam taxa média de 0,3732 e desvio padrão no valor de 0,1784.

Como alternativa ao cenário com função de ligação Logit, também efetuou-se os mesmos procedimentos para as funções de ligação Probit e Complemento Log-log. Para a função de ligação Probit, observou-se mesmo comportamento da função de ligação Logit, com exceção das medidas R^2 e R_a^2 que apontam o melhor ajuste para o modelo de regressão Simplex no caso do teste da razão de verossimilhanças generalizada. No que diz respeito ao caso do teste de Wald, novamente o modelo de regressão Beta apresenta o melhor ajuste em termos das medidas R^2 e R_a^2 , sendo o número de covariáveis a possível explicação. Assim, embora haja discordância entre as medidas de qualidade do ajuste nos casos dos testes, o modelo de regressão L-logística apresenta o melhor ajuste geral. No que diz respeito a função de ligação Complemento Log-log, observou-se o mesmo comportamento da função de ligação Probit, com diferença no caso do teste de Wald em que observou-se que o modelo de regressão Simplex apresenta o melhor ajuste em termos de R^2 e R_a^2 . Contudo, Morrison (1973) apresenta evidências empíricas sobre o insucesso da qualidade de ajuste fornecida pela medida R^2 , o que por consequência pode ocorrer com a medida R_a^2 , dado que a mesma é função da medida R^2 . A seguir, nas Tabelas 62 a 87, os resultados para as funções de ligação Probit e Complemento Log-log. Vale destacar que as Tabelas 73, 74, 86 e 87 apresentam os intervalos de 95% de confiança propostos na seção 4, para o modelo de regressão L-logística com função de ligação Probit e Complemento Log-log. Com relação ao processo de identificação de observações atípicas nos cenários da função de ligação Probit e Complemento Log-log, constatou-se os mesmos resultados observados no cenário da função de ligação Logit. A seguir, nas Figuras 14, 15, 17 e 18, apresenta-se os resultados para a funções de ligação Probit e Complemento Log-log, para as perspectivas do teste da razão de verossimilhanças generalizada e do teste do tipo Wald.

No que remete ao processo de escolha do modelo mais adequado para este conjunto de dados, constatou-se que o modelo de regressão L-logística apresentou melhor desempenho em relação aos demais modelos de regressão, independente da função de ligação proposta. Contudo, vale comentar que visando as medidas de qualidade do ajuste AIC e BIC, assim como a medida de distância do tipo Cook, constatou-se que a função de ligação Complemento Log-log proporcionou o melhor ajuste, assim tendo como melhor escolha o modelo de regressão L-logística com função de ligação Complemento Log-log. Além disso, vale apontar que embora o teste da razão de verossimilhanças generalizada seja bastante popular por possuir menos requisitos para sua utilização, o cenário do teste tipo Wald proporcionou valores mais homogêneos das distâncias do tipo Cook, o que fornece um indicativo de proximidade da realidade dos dados. O leitor interessado pode consultar as Tabelas 88 e 89 que apresenta algumas estatísticas da distância do tipo Cook, tais como mínimo, máximo, média, desvio padrão, mediana e intervalo interquartil, para as perspectivas dos dois testes.

Tabela 49: Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Logit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,2220	0,1136	-10,7530	$< 10^{-4}$
	β_1	0,6147	0,0387	15,8676	$< 10^{-4}$
	β_2	-0,7379	0,1197	-6,1632	$< 10^{-4}$
	β_3	0,4548	0,1266	3,5935	0,0003
	β_4	-0,6671	0,1249	-5,3433	$< 10^{-4}$
	β_5	0,1892	0,1199	1,5775	0,1147
	b	4,3501	0,1974	22,0413	$< 10^{-4}$
Beta	β_0	-1,1951	0,0848	-14,0853	$< 10^{-4}$
	β_1	0,6044	0,0397	15,2233	$< 10^{-4}$
	β_2	-0,6746	0,0919	-7,3378	$< 10^{-4}$
	β_3	0,4358	0,1020	4,2726	$< 10^{-4}$
	β_4	-0,6049	0,0972	-6,2223	$< 10^{-4}$
	β_5	0,2079	0,0926	2,2448	0,0248
	ϕ	28,2602	3,3288	8,4897	$< 10^{-4}$
Simplex	β_0	-1,1515	0,0799	-14,4073	$< 10^{-4}$
	β_1	0,6092	0,0413	14,7335	$< 10^{-4}$
	β_2	-0,7584	0,0878	-8,6364	$< 10^{-4}$
	β_3	0,3812	0,0958	3,9784	0,0001
	β_4	-0,6563	0,0944	-6,9557	$< 10^{-4}$
	β_5	0,1349	0,0869	1,5533	0,1204
	σ^2	1,0111	0,1070	9,4480	$< 10^{-4}$

Tabela 50: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Logit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,2220	0,1136	-10,7530	$< 10^{-4}$
	β_1	0,6147	0,0387	15,8676	$< 10^{-4}$
	β_2	-0,7379	0,1197	-6,1632	$< 10^{-4}$
	β_3	0,4548	0,1266	3,5935	0,0003
	β_4	-0,6671	0,1249	-5,3433	$< 10^{-4}$
	β_5	0,1892	0,1199	1,5775	0,1147
	b	4,3501	0,1974	22,0413	$< 10^{-4}$
Beta	β_0	-1,1951	0,0848	-14,0853	$< 10^{-4}$
	β_1	0,6044	0,0397	15,2233	$< 10^{-4}$
	β_2	-0,6746	0,0919	-7,3378	$< 10^{-4}$
	β_3	0,4358	0,1020	4,2726	$< 10^{-4}$
	β_4	-0,6049	0,0972	-6,2223	$< 10^{-4}$
	β_5	0,2079	0,0926	2,2448	0,0248
	ϕ	28,2602	3,3288	8,4897	$< 10^{-4}$
Simplex	β_0	-1,1515	0,0799	-14,4073	$< 10^{-4}$
	β_1	0,6092	0,0413	14,7335	$< 10^{-4}$
	β_2	-0,7584	0,0878	-8,6364	$< 10^{-4}$
	β_3	0,3812	0,0958	3,9784	0,0001
	β_4	-0,6563	0,0944	-6,9557	$< 10^{-4}$
	β_5	0,1349	0,0869	1,5533	0,1204
	σ^2	1,0111	0,1070	9,4480	$< 10^{-4}$

Tabela 51: Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Logit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,2220	0,1161	-10,5298	$< 10^{-4}$
	β_1	0,6147	0,0386	15,9120	$< 10^{-4}$
	β_2	-0,7379	0,1207	-6,1115	$< 10^{-4}$
	β_3	0,4548	0,1261	3,6070	0,0003
	β_4	-0,6671	0,1259	-5,3002	$< 10^{-4}$
	β_5	0,1892	0,1204	1,5717	0,1160
	b	4,3501	0,1920	22,6616	$< 10^{-4}$
Beta	β_0	-1,1951	0,0844	-14,1527	$< 10^{-4}$
	β_1	0,6044	0,0403	15,0062	$< 10^{-4}$
	β_2	-0,6746	0,0920	-7,3328	$< 10^{-4}$
	β_3	0,4358	0,0997	4,3709	$< 10^{-4}$
	β_4	-0,6049	0,0980	-6,1705	$< 10^{-4}$
	β_5	0,2079	0,0916	2,2694	0,0232
	ϕ	28,2602	3,2800	8,6160	$< 10^{-4}$
Simplex	β_0	-1,1515	0,0812	-14,1799	$< 10^{-4}$
	β_1	0,6092	0,0431	14,1244	$< 10^{-4}$
	β_2	-0,7584	0,0885	-8,5650	$< 10^{-4}$
	β_3	0,3812	0,0954	3,9981	0,0001
	β_4	-0,6563	0,0969	-6,7741	$< 10^{-4}$
	β_5	0,1349	0,0882	1,5298	0,1261
	σ^2	1,0111	0,1083	9,3345	$< 10^{-4}$

Tabela 52: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Logit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,2220	0,1161	-10,5298	$< 10^{-4}$
	β_1	0,6147	0,0386	15,9120	$< 10^{-4}$
	β_2	-0,7379	0,1207	-6,1115	$< 10^{-4}$
	β_3	0,4548	0,1261	3,6070	0,0003
	β_4	-0,6671	0,1259	-5,3002	$< 10^{-4}$
	β_5	0,1892	0,1204	1,5717	0,1160
	b	4,3501	0,1920	22,6616	$< 10^{-4}$
Beta	β_0	-1,1951	0,0844	-14,1527	$< 10^{-4}$
	β_1	0,6044	0,0403	15,0062	$< 10^{-4}$
	β_2	-0,6746	0,0920	-7,3328	$< 10^{-4}$
	β_3	0,4358	0,0997	4,3709	$< 10^{-4}$
	β_4	-0,6049	0,0980	-6,1705	$< 10^{-4}$
	β_5	0,2079	0,0916	2,2694	0,0232
	ϕ	28,2602	3,2800	8,6160	$< 10^{-4}$
Simplex	β_0	-1,1515	0,0812	-14,1799	$< 10^{-4}$
	β_1	0,6092	0,0431	14,1244	$< 10^{-4}$
	β_2	-0,7584	0,0885	-8,5650	$< 10^{-4}$
	β_3	0,3812	0,0954	3,9981	0,0001
	β_4	-0,6563	0,0969	-6,7741	$< 10^{-4}$
	β_5	0,1349	0,0882	1,5298	0,1261
	σ^2	1,0111	0,1083	9,3345	$< 10^{-4}$

Tabela 53: Teste da Razão de Verossimilhanças Generalizada para a função de ligação Logit.

Modelo de Regressão	Hipótese	$l(\theta)$	Gráus de Liberdade	Estatística	Valor-P
L-logística	H_0	514,2903	1	6,7062	0,0096
	H_1	517,6434			
Beta	H_0	477,5625	1	11,04134	0,0009
	H_1	483,0831			
Simplex	H_0	508,7697	1	4,1536	0,0415
	H_1	510,8465			

Tabela 54: Teste tipo Wald com combinação do Algoritmo Genético com método Bootstrap para a função de ligação Logit e remoção da categoria Caucasianos.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,0800	0,0355	-30,4401	$< 10^{-4}$
	β_1	0,6165	0,0379	16,2570	$< 10^{-4}$
	β_2	-0,8807	0,0429	-20,5524	$< 10^{-4}$
	β_3	0,3119	0,0581	5,3645	$< 10^{-4}$
	β_4	-0,8100	0,0566	-14,3013	$< 10^{-4}$
	b	4,3352	0,1983	21,8658	$< 10^{-4}$
Simplex	β_0	-1,0763	0,0388	-27,7251	$< 10^{-4}$
	β_1	0,6042	0,0412	14,6685	$< 10^{-4}$
	β_2	-0,8318	0,0488	-17,0555	$< 10^{-4}$
	β_3	0,3086	0,0602	5,1282	$< 10^{-4}$
	β_4	-0,7295	0,0618	-11,8119	$< 10^{-4}$
	σ^2	1,0210	0,1047	9,7472	$< 10^{-4}$

Tabela 55: Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Bootstrap para a função de ligação Logit e remoção da categoria Caucasianos.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,0800	0,0355	-30,4401	$< 10^{-4}$
	β_1	0,6165	0,0379	16,2570	$< 10^{-4}$
	β_2	-0,8807	0,0429	-20,5524	$< 10^{-4}$
	β_3	0,3119	0,0581	5,3645	$< 10^{-4}$
	β_4	-0,8100	0,0566	-14,3013	$< 10^{-4}$
	b	4,3352	0,1983	21,8658	$< 10^{-4}$
Simplex	β_0	-1,0763	0,0388	-27,7251	$< 10^{-4}$
	β_1	0,6042	0,0412	14,6685	$< 10^{-4}$
	β_2	-0,8318	0,0488	-17,0555	$< 10^{-4}$
	β_3	0,3086	0,0602	5,1282	$< 10^{-4}$
	β_4	-0,7295	0,0618	-11,8119	$< 10^{-4}$
	σ^2	1,0210	0,1047	9,7472	$< 10^{-4}$

Tabela 56: Teste tipo Wald com combinação do Algoritmo Genético com método Jackknife para a função de ligação Logit e remoção da categoria Caucasianos.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,0800	0,0360	-29,9696	$< 10^{-4}$
	β_1	0,6165	0,0380	16,2324	$< 10^{-4}$
	β_2	-0,8807	0,0430	-20,4788	$< 10^{-4}$
	β_3	0,3119	0,0572	5,4561	$< 10^{-4}$
	β_4	-0,8100	0,0559	-14,4785	$< 10^{-4}$
	b	4,3352	0,1970	22,0035	$< 10^{-4}$
Simplex	β_0	-1,0763	0,0392	-27,4893	$< 10^{-4}$
	β_1	0,6042	0,0426	14,1972	$< 10^{-4}$
	β_2	-0,8318	0,0490	-16,9842	$< 10^{-4}$
	β_3	0,3086	0,0589	5,2377	$< 10^{-4}$
	β_4	-0,7295	0,0625	-11,6742	$< 10^{-4}$
	σ^2	1,0210	0,1050	9,7197	$< 10^{-4}$

Tabela 57: Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Jackknife para a função de ligação Logit e remoção da categoria Caucasianos.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,0800	0,0360	-29,9696	$< 10^{-4}$
	β_1	0,6165	0,0380	16,2324	$< 10^{-4}$
	β_2	-0,8807	0,0430	-20,4788	$< 10^{-4}$
	β_3	0,3119	0,0572	5,4561	$< 10^{-4}$
	β_4	-0,8100	0,0559	-14,4785	$< 10^{-4}$
	b	4,3352	0,1970	22,0035	$< 10^{-4}$
Simplex	β_0	-1,0763	0,0392	-27,4893	$< 10^{-4}$
	β_1	0,6042	0,0426	14,1972	$< 10^{-4}$
	β_2	-0,8318	0,0490	-16,9842	$< 10^{-4}$
	β_3	0,3086	0,0589	5,2377	$< 10^{-4}$
	β_4	-0,7295	0,0625	-11,6742	$< 10^{-4}$
	σ^2	1,0210	0,1050	9,7197	$< 10^{-4}$

Tabela 58: Medidas de qualidade do ajuste com todas a covariáveis e função de ligação Logit.

Modelo de Regressão	$-l(\theta)$	AIC	BIC	R^2	R_a^2
L-logística	-517,6434	-1021,2867	-950,4590	0,6207	0,6189
Beta	-483,0831	-952,1663	-881,3385	0,6206	0,6188
Simplex	-510,8465	-1007,6931	-936,8654	0,6224	0,6207

Tabela 59: Medidas de qualidade do ajuste para a função de ligação Logit e a remoção da categoria Caucasianos dos modelos de regressão L-logística e Simplex.

Modelo de Regressão	$-l(\theta)$	AIC	BIC	R^2	R_a^2
L-logística	-514,2903	-1016,5806	-955,8711	0,6173	0,6155
Beta	-483,0831	-952,1663	-881,3385	0,6206	0,6188
Simplex	-508,7697	-1005,5393	-944,8299	0,6168	0,6150

Tabela 60: Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Logit.

Intervalo	Parâmetro	Limite Inferior	Limite Superior
Jackknife- <i>t</i>	β_0	-1,4134	-1,0307
	β_1	0,5510	0,6784
	β_2	-0,9369	-0,5389
	β_3	0,2469	0,6626
	β_4	-0,8746	-0,4596
	β_5	-0,0092	0,3876
	b	4,0337	4,6665
Pivotal	β_0	-1,4263	-0,9954
	β_1	0,5391	0,6888
	β_2	-0,9794	-0,5292
	β_3	0,2054	0,6875
	β_4	-0,9139	-0,4380
	β_5	-0,0577	0,4085
	b	3,8819	4,6602
Percentil	β_0	-1,4487	-1,0178
	β_1	0,5406	0,6904
	β_2	-0,9465	-0,4963
	β_3	0,2221	0,7042
	β_4	-0,8963	-0,4203
	β_5	-0,0301	0,4361
	b	4,0400	4,8183
BC	β_0	-1,4561	-1,0213
	β_1	0,5414	0,6933
	β_2	-0,9404	-0,4862
	β_3	0,2310	0,7286
	β_4	-0,8743	-0,4082
	β_5	-0,0195	0,4446
	b	3,9631	4,7183
BC _{<i>a</i>}	β_0	-1,4569	-1,0217
	β_1	0,5415	0,6933
	β_2	-0,9395	-0,4848
	β_3	0,2313	0,7297
	β_4	-0,8734	-0,4078
	β_5	-0,0190	0,4454
	b	3,9386	4,7011

Tabela 61: Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Logit e remoção da categoria Caucasianos.

Intervalo	Parâmetro	Limite Inferior	Limite Superior
Jackknife- <i>t</i>	β_0	-1,1394	-1,0206
	β_1	0,5539	0,6791
	β_2	-0,9516	-0,8098
	β_3	0,2177	0,4061
	β_4	-0,9022	-0,7178
	b	4,0105	4,6600
Pivotal	β_0	-1,1457	-1,0082
	β_1	0,5418	0,6877
	β_2	-0,9691	-0,8044
	β_3	0,1974	0,4297
	β_4	-0,9284	-0,7001
	b	3,8961	4,6640
Percentil	β_0	-1,1518	-1,0143
	β_1	0,5452	0,6912
	β_2	-0,9571	-0,7924
	β_3	0,1941	0,4263
	β_4	-0,9199	-0,6916
	b	4,0065	4,7744
BC	β_0	-1,1528	-1,0153
	β_1	0,5459	0,6920
	β_2	-0,9552	-0,7915
	β_3	0,1879	0,4194
	β_4	-0,9154	-0,6880
	b	3,9431	4,7036
BC _{<i>a</i>}	β_0	-1,1537	-1,0158
	β_1	0,5459	0,6920
	β_2	-0,9551	-0,7912
	β_3	0,1877	0,4193
	β_4	-0,9154	-0,6878
	b	3,9201	4,6873

Tabela 62: Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Probit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-0,7425	0,0673	-11,0324	$< 10^{-4}$
	β_1	0,3563	0,0228	15,6247	$< 10^{-4}$
	β_2	-0,4214	0,0706	-5,9725	$< 10^{-4}$
	β_3	0,2770	0,0755	3,6705	0,0002
	β_4	-0,3827	0,0734	-5,2155	$< 10^{-4}$
	β_5	0,1148	0,0711	1,6138	0,1066
	b	4,3238	0,1952	22,1530	$< 10^{-4}$
Beta	β_0	-0,7276	0,0504	-14,4462	$< 10^{-4}$
	β_1	0,3566	0,0235	15,1533	$< 10^{-4}$
	β_2	-0,3909	0,0544	-7,1904	$< 10^{-4}$
	β_3	0,2658	0,0611	4,3514	$< 10^{-4}$
	β_4	-0,3522	0,0572	-6,1612	$< 10^{-4}$
	β_5	0,1253	0,0551	2,2748	0,0229
	ϕ	28,0664	3,2868	8,5391	$< 10^{-4}$
Simplex	β_0	-0,6991	0,0482	-14,4924	$< 10^{-4}$
	β_1	0,3540	0,0246	14,4083	$< 10^{-4}$
	β_2	-0,4358	0,0523	-8,3273	$< 10^{-4}$
	β_3	0,2327	0,0583	3,9934	0,0001
	β_4	-0,3805	0,0557	-6,8255	$< 10^{-4}$
	β_5	0,0803	0,0525	1,5304	0,1259
	σ^2	1,0194	0,1078	9,4567	$< 10^{-4}$

Tabela 63: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Probit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-0,7425	0,0673	-11,0324	$< 10^{-4}$
	β_1	0,3563	0,0228	15,6247	$< 10^{-4}$
	β_2	-0,4214	0,0706	-5,9725	$< 10^{-4}$
	β_3	0,2770	0,0755	3,6705	0,0002
	β_4	-0,3827	0,0734	-5,2155	$< 10^{-4}$
	β_5	0,1148	0,0711	1,6138	0,1066
	b	4,3238	0,1952	22,1530	$< 10^{-4}$
Beta	β_0	-0,7276	0,0504	-14,4462	$< 10^{-4}$
	β_1	0,3566	0,0235	15,1533	$< 10^{-4}$
	β_2	-0,3909	0,0544	-7,1904	$< 10^{-4}$
	β_3	0,2658	0,0611	4,3514	$< 10^{-4}$
	β_4	-0,3522	0,0572	-6,1612	$< 10^{-4}$
	β_5	0,1253	0,0551	2,2748	0,0229
	ϕ	28,0664	3,2868	8,5391	$< 10^{-4}$
Simplex	β_0	-0,6991	0,0482	-14,4924	$< 10^{-4}$
	β_1	0,3540	0,0246	14,4083	$< 10^{-4}$
	β_2	-0,4358	0,0523	-8,3273	$< 10^{-4}$
	β_3	0,2327	0,0583	3,9934	0,0001
	β_4	-0,3805	0,0557	-6,8255	$< 10^{-4}$
	β_5	0,0803	0,0525	1,5304	0,1259
	σ^2	1,0194	0,1078	9,4567	$< 10^{-4}$

Tabela 64: Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Probit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-0,7425	0,0689	-10,7713	$< 10^{-4}$
	β_1	0,3563	0,0227	15,7221	$< 10^{-4}$
	β_2	-0,4214	0,0714	-5,9002	$< 10^{-4}$
	β_3	0,2770	0,0753	3,6798	0,0002
	β_4	-0,3827	0,0741	-5,1653	$< 10^{-4}$
	β_5	0,1148	0,0716	1,6037	0,1088
	b	4,3238	0,1892	22,8478	$< 10^{-4}$
Beta	β_0	-0,7276	0,0502	-14,5075	$< 10^{-4}$
	β_1	0,3566	0,0238	14,9635	$< 10^{-4}$
	β_2	-0,3909	0,0544	-7,1846	$< 10^{-4}$
	β_3	0,2658	0,0596	4,4562	$< 10^{-4}$
	β_4	-0,3522	0,0576	-6,1120	$< 10^{-4}$
	β_5	0,1253	0,0545	2,2987	0,0215
	ϕ	28,0664	3,2373	8,6697	$< 10^{-4}$
Simplex	β_0	-0,6991	0,0489	-14,2818	$< 10^{-4}$
	β_1	0,3540	0,0256	13,8313	$< 10^{-4}$
	β_2	-0,4358	0,0527	-8,2706	$< 10^{-4}$
	β_3	0,2327	0,0578	4,0258	0,0001
	β_4	-0,3805	0,0571	-6,6611	$< 10^{-4}$
	β_5	0,0803	0,0532	1,5093	0,1312
	σ^2	1,0194	0,1090	9,3513	$< 10^{-4}$

Tabela 65: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Probit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-0,7425	0,0689	-10,7713	$< 10^{-4}$
	β_1	0,3563	0,0227	15,7221	$< 10^{-4}$
	β_2	-0,4214	0,0714	-5,9002	$< 10^{-4}$
	β_3	0,2770	0,0753	3,6798	0,0002
	β_4	-0,3827	0,0741	-5,1653	$< 10^{-4}$
	β_5	0,1148	0,0716	1,6037	0,1088
	b	4,3238	0,1892	22,8478	$< 10^{-4}$
Beta	β_0	-0,7276	0,0502	-14,5075	$< 10^{-4}$
	β_1	0,3566	0,0238	14,9635	$< 10^{-4}$
	β_2	-0,3909	0,0544	-7,1846	$< 10^{-4}$
	β_3	0,2658	0,0596	4,4562	$< 10^{-4}$
	β_4	-0,3522	0,0576	-6,1120	$< 10^{-4}$
	β_5	0,1253	0,0545	2,2987	0,0215
	ϕ	28,0664	3,2373	8,6697	$< 10^{-4}$
Simplex	β_0	-0,6991	0,0489	-14,2818	$< 10^{-4}$
	β_1	0,3540	0,0256	13,8313	$< 10^{-4}$
	β_2	-0,4358	0,0527	-8,2706	$< 10^{-4}$
	β_3	0,2327	0,0578	4,0258	0,0001
	β_4	-0,3805	0,0571	-6,6611	$< 10^{-4}$
	β_5	0,0803	0,0532	1,5093	0,1312
	σ^2	1,0194	0,1090	9,3513	$< 10^{-4}$

Tabela 66: Teste da Razão de Verossimilhanças Generalizada para a função de ligação Probit.

Modelo de Regressão	Hipótese	$l(\theta)$	Gráus de Liberdade	Estatística	Valor-P
L-logística	H_0	512,0248	1	6,747	0,0094
	H_1	515,3983			
Simplex	H_0	507,1172	1	3,9712	0,0463
	H_1	509,1028			

Tabela 67: Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Probit e remoção da categoria Caucasianos.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-0,6567	0,0213	-30,8809	$< 10^{-4}$
	β_1	0,3574	0,0223	16,0440	$< 10^{-4}$
	β_2	-0,5077	0,0252	-20,1102	$< 10^{-4}$
	β_3	0,1906	0,0357	5,3417	$< 10^{-4}$
	β_4	-0,4690	0,0326	-14,3759	$< 10^{-4}$
	b	4,3084	0,1961	21,9759	$< 10^{-4}$
Simplex	β_0	-0,6544	0,0235	-27,8475	$< 10^{-4}$
	β_1	0,3513	0,0244	14,3833	$< 10^{-4}$
	β_2	-0,4796	0,0286	-16,7478	$< 10^{-4}$
	β_3	0,1894	0,0370	5,1139	$< 10^{-4}$
	β_4	-0,4242	0,0356	-11,9159	$< 10^{-4}$
	σ^2	1,0289	0,1054	9,7587	$< 10^{-4}$

Tabela 68: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Probit e remoção da categoria Caucasianos.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-0,6567	0,0213	-30,8809	$< 10^{-4}$
	β_1	0,3574	0,0223	16,0440	$< 10^{-4}$
	β_2	-0,5077	0,0252	-20,1102	$< 10^{-4}$
	β_3	0,1906	0,0357	5,3417	$< 10^{-4}$
	β_4	-0,4690	0,0326	-14,3759	$< 10^{-4}$
	b	4,3084	0,1961	21,9759	$< 10^{-4}$
Simplex	β_0	-0,6544	0,0235	-27,8475	$< 10^{-4}$
	β_1	0,3513	0,0244	14,3833	$< 10^{-4}$
	β_2	-0,4796	0,0286	-16,7478	$< 10^{-4}$
	β_3	0,1894	0,0370	5,1139	$< 10^{-4}$
	β_4	-0,4242	0,0356	-11,9159	$< 10^{-4}$
	σ^2	1,0289	0,1054	9,7587	$< 10^{-4}$

Tabela 69: Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Probit e remoção da categoria Caucasionos.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-0,6567	0,0216	-30,4316	$< 10^{-4}$
	β_1	0,3574	0,0222	16,0638	$< 10^{-4}$
	β_2	-0,5077	0,0254	-20,0238	$< 10^{-4}$
	β_3	0,1906	0,0351	5,4370	$< 10^{-4}$
	β_4	-0,4690	0,0322	-14,5876	$< 10^{-4}$
	b	4,3084	0,1942	22,1882	$< 10^{-4}$
Simplex	β_0	-0,6544	0,0237	-27,6458	$< 10^{-4}$
	β_1	0,3513	0,0252	13,9501	$< 10^{-4}$
	β_2	-0,4796	0,0287	-16,7077	$< 10^{-4}$
	β_3	0,1894	0,0362	5,2352	$< 10^{-4}$
	β_4	-0,4242	0,0360	-11,7894	$< 10^{-4}$
	σ^2	1,0289	0,1056	9,7427	$< 10^{-4}$

Tabela 70: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Probit e remoção da categoria Caucasionos.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-0,6567	0,0216	-30,4316	$< 10^{-4}$
	β_1	0,3574	0,0222	16,0638	$< 10^{-4}$
	β_2	-0,5077	0,0254	-20,0238	$< 10^{-4}$
	β_3	0,1906	0,0351	5,4370	$< 10^{-4}$
	β_4	-0,4690	0,0322	-14,5876	$< 10^{-4}$
	b	4,3084	0,1942	22,1882	$< 10^{-4}$
Simplex	β_0	-0,6544	0,0237	-27,6458	$< 10^{-4}$
	β_1	0,3513	0,0252	13,9501	$< 10^{-4}$
	β_2	-0,4796	0,0287	-16,7077	$< 10^{-4}$
	β_3	0,1894	0,0362	5,2352	$< 10^{-4}$
	β_4	-0,4242	0,0360	-11,7894	$< 10^{-4}$
	σ^2	1,0289	0,1056	9,7427	$< 10^{-4}$

Tabela 71: Medidas de qualidade do ajuste com todas a covariáveis e função de ligação Probit.

Modelo de Regressão	$-l(\theta)$	AIC	BIC	R^2	R_a^2
L-logística	-515,3983	-1016,7967	-945,9689	0,6178	0,6160
Beta	-481,5987	-949,1975	-878,3697	0,6181	0,6163
Simplex	-509,1028	-1004,2057	-933,3780	0,6194	0,6176

Tabela 72: Medidas de qualidade do ajuste para a função de ligação Probit e a remoção da categoria Caucasianos dos modelos de regressão L-logística e Simplex.

Modelo de Regressão	$-l(\theta)$	AIC	BIC	R^2	R_a^2
L-logística	-512,0248	-1012,0497	-951,3402	0,6147	0,6129
Beta	-481,5987	-949,1975	-878,3697	0,6181	0,6163
Simplex	-507,1172	-1002,2344	-941,5249	0,6138	0,6119

Tabela 73: Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Probit.

Intervalo	Parâmetro	Limite Inferior	Limite Superior
Jackknife- <i>t</i>	β_0	-0,8562	-0,6289
	β_1	0,3189	0,3936
	β_2	-0,5391	-0,3036
	β_3	0,1529	0,4011
	β_4	-0,5048	-0,2605
	β_5	-0,0032	0,2328
	b	4,0118	4,6357
Pivotal	β_0	-0,8656	-0,6083
	β_1	0,3110	0,4003
	β_2	-0,5631	-0,2957
	β_3	0,1303	0,4186
	β_4	-0,5247	-0,2481
	β_5	-0,0299	0,2452
	b	3,8543	4,6292
Percentil	β_0	-0,8768	-0,6194
	β_1	0,3123	0,4016
	β_2	-0,5470	-0,2796
	β_3	0,1355	0,4237
	β_4	-0,5172	-0,2407
	β_5	-0,0155	0,2595
	b	4,0183	4,7932
BC	β_0	-0,8797	-0,6208
	β_1	0,3139	0,4022
	β_2	-0,5419	-0,2768
	β_3	0,1429	0,4420
	β_4	-0,5082	-0,2333
	β_5	-0,0097	0,2629
	b	3,9428	4,6870
BC _{<i>a</i>}	β_0	-0,8798	-0,621
	β_1	0,3139	0,4022
	β_2	-0,5419	-0,2751
	β_3	0,1431	0,4420
	β_4	-0,5082	-0,2322
	β_5	-0,0094	0,2631
	b	3,9182	4,6688

Tabela 74: Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Probit e remoção da categoria Caucasianos.

Intervalo	Parâmetro	Limite Inferior	Limite Superior
Jackknife- <i>t</i>	β_0	-0,6923	-0,6211
	β_1	0,3207	0,3941
	β_2	-0,5495	-0,4659
	β_3	0,1328	0,2485
	β_4	-0,5220	-0,4160
	b	3,9883	4,6285
Pivotal	β_0	-0,6962	-0,6139
	β_1	0,3130	0,3996
	β_2	-0,5595	-0,4624
	β_3	0,1208	0,2637
	β_4	-0,5352	-0,4047
	b	3,8732	4,6326
Percentil	β_0	-0,6994	-0,6172
	β_1	0,3152	0,4018
	β_2	-0,5530	-0,4559
	β_3	0,1176	0,2605
	β_4	-0,5333	-0,4028
	b	3,9843	4,7436
BC	β_0	-0,6998	-0,6175
	β_1	0,3155	0,4020
	β_2	-0,5512	-0,4556
	β_3	0,1134	0,2565
	β_4	-0,5292	-0,3985
	b	3,9197	4,6811
BC _{<i>a</i>}	β_0	-0,7001	-0,6178
	β_1	0,3155	0,4020
	β_2	-0,5511	-0,4555
	β_3	0,1133	0,2563
	β_4	-0,5292	-0,3982
	b	3,9049	4,6675

Tabela 75: Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Complemento Log-log.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,3557	0,0977	-13,8708	$< 10^{-4}$
	β_1	0,5288	0,0326	16,2141	$< 10^{-4}$
	β_2	-0,6540	0,1032	-6,3386	$< 10^{-4}$
	β_3	0,3707	0,1078	3,4389	0,0006
	β_4	-0,5893	0,1079	-5,4624	$< 10^{-4}$
	β_5	0,1572	0,1030	1,5262	0,1270
	b	4,3711	0,1991	21,9549	$< 10^{-4}$
Beta	β_0	-1,3293	0,0725	-18,3288	$< 10^{-4}$
	β_1	0,5106	0,0329	15,5075	$< 10^{-4}$
	β_2	-0,5901	0,0783	-7,5357	$< 10^{-4}$
	β_3	0,3525	0,0860	4,1002	$< 10^{-4}$
	β_4	-0,5269	0,0834	-6,3161	$< 10^{-4}$
	β_5	0,1724	0,0790	2,1835	0,0290
	ϕ	28,3987	3,3657	8,4378	$< 10^{-4}$
Simplex	β_0	-1,2984	0,0674	-19,2546	$< 10^{-4}$
	β_1	0,5254	0,0346	15,1990	$< 10^{-4}$
	β_2	-0,6704	0,0747	-8,9708	$< 10^{-4}$
	β_3	0,3081	0,0797	3,8638	0,0001
	β_4	-0,5758	0,0811	-7,0984	$< 10^{-4}$
	β_5	0,1134	0,0732	1,5496	0,1212
	σ^2	1,0044	0,1064	9,4396	$< 10^{-4}$

Tabela 76: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Complemento Log-log.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,3557	0,0977	-13,8708	$< 10^{-4}$
	β_1	0,5288	0,0326	16,2141	$< 10^{-4}$
	β_2	-0,6540	0,1032	-6,3386	$< 10^{-4}$
	β_3	0,3707	0,1078	3,4389	0,0006
	β_4	-0,5893	0,1079	-5,4624	$< 10^{-4}$
	β_5	0,1572	0,1030	1,5262	0,1270
	b	4,3711	0,1991	21,9549	$< 10^{-4}$
Beta	β_0	-1,3293	0,0725	-18,3288	$< 10^{-4}$
	β_1	0,5106	0,0329	15,5075	$< 10^{-4}$
	β_2	-0,5901	0,0783	-7,5357	$< 10^{-4}$
	β_3	0,3525	0,0860	4,1002	$< 10^{-4}$
	β_4	-0,5269	0,0834	-6,3161	$< 10^{-4}$
	β_5	0,1724	0,0790	2,1835	0,0290
	ϕ	28,3987	3,3657	8,4378	$< 10^{-4}$
Simplex	β_0	-1,2984	0,0674	-19,2546	$< 10^{-4}$
	β_1	0,5254	0,0346	15,1990	$< 10^{-4}$
	β_2	-0,6704	0,0747	-8,9708	$< 10^{-4}$
	β_3	0,3081	0,0797	3,8638	0,0001
	β_4	-0,5758	0,0811	-7,0984	$< 10^{-4}$
	β_5	0,1134	0,0732	1,5496	0,1212
	σ^2	1,0044	0,1064	9,4396	$< 10^{-4}$

Tabela 77: Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Complemento Log-log.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,3557	0,0995	-13,6253	$< 10^{-4}$
	β_1	0,5288	0,0327	16,1842	$< 10^{-4}$
	β_2	-0,6540	0,1037	-6,3069	$< 10^{-4}$
	β_3	0,3707	0,1073	3,4559	0,0005
	β_4	-0,5893	0,1087	-5,4225	$< 10^{-4}$
	β_5	0,1572	0,1031	1,5249	0,1273
	b	4,3711	0,1943	22,4972	$< 10^{-4}$
Beta	β_0	-1,3293	0,0722	-18,4238	$< 10^{-4}$
	β_1	0,5106	0,0335	15,2615	$< 10^{-4}$
	β_2	-0,5901	0,0783	-7,5360	$< 10^{-4}$
	β_3	0,3525	0,0842	4,1883	$< 10^{-4}$
	β_4	-0,5269	0,0841	-6,2636	$< 10^{-4}$
	β_5	0,1724	0,0780	2,2091	0,0272
	ϕ	28,3987	3,3182	8,5583	$< 10^{-4}$
Simplex	β_0	-1,2984	0,0687	-18,9096	$< 10^{-4}$
	β_1	0,5254	0,0361	14,5479	$< 10^{-4}$
	β_2	-0,6704	0,0755	-8,8765	$< 10^{-4}$
	β_3	0,3081	0,0798	3,8618	0,0001
	β_4	-0,5758	0,0835	-6,8950	$< 10^{-4}$
	β_5	0,1134	0,0745	1,5227	0,1278
	σ^2	1,0044	0,1079	9,3083	$< 10^{-4}$

Tabela 78: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Complemento Log-log.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,3557	0,0995	-13,6253	$< 10^{-4}$
	β_1	0,5288	0,0327	16,1842	$< 10^{-4}$
	β_2	-0,6540	0,1037	-6,3069	$< 10^{-4}$
	β_3	0,3707	0,1073	3,4559	0,0005
	β_4	-0,5893	0,1087	-5,4225	$< 10^{-4}$
	β_5	0,1572	0,1031	1,5249	0,1273
	b	4,3711	0,1943	22,4972	$< 10^{-4}$
Beta	β_0	-1,3293	0,0722	-18,4238	$< 10^{-4}$
	β_1	0,5106	0,0335	15,2615	$< 10^{-4}$
	β_2	-0,5901	0,0783	-7,5360	$< 10^{-4}$
	β_3	0,3525	0,0842	4,1883	$< 10^{-4}$
	β_4	-0,5269	0,0841	-6,2636	$< 10^{-4}$
	β_5	0,1724	0,0780	2,2091	0,0272
	ϕ	28,3987	3,3182	8,5583	$< 10^{-4}$
Simplex	β_0	-1,2984	0,0687	-18,9096	$< 10^{-4}$
	β_1	0,5254	0,0361	14,5479	$< 10^{-4}$
	β_2	-0,6704	0,0755	-8,8765	$< 10^{-4}$
	β_3	0,3081	0,0798	3,8618	0,0001
	β_4	-0,5758	0,0835	-6,8950	$< 10^{-4}$
	β_5	0,1134	0,0745	1,5227	0,1278
	σ^2	1,0044	0,1079	9,3083	$< 10^{-4}$

Tabela 79: Teste da Razão de Verossimilhanças Generalizada para a função de ligação Complemento Log-log.

Modelo de Regressão	Hipótese	$l(\theta)$	Graus de Liberdade	Estatística	Valor-P
L-logística	H_0	516,0725	1	6,6838	0,0097
	H_1	519,4144			
Simplex	H_0	510,1031	1	4,327	0,0375
	H_1	512,2666			

Tabela 80: Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Complemento Log-log e remoção da categoria Caucasianos.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,2371	0,0301	-41,0526	$< 10^{-4}$
	β_1	0,5303	0,0321	16,5438	$< 10^{-4}$
	β_2	-0,7733	0,0365	-21,1687	$< 10^{-4}$
	β_3	0,2512	0,0471	5,3343	$< 10^{-4}$
	β_4	-0,7087	0,0497	-14,2687	$< 10^{-4}$
	b	4,3568	0,2000	21,7859	$< 10^{-4}$
Simplex	β_0	-1,2351	0,0327	-37,7561	$< 10^{-4}$
	β_1	0,5210	0,0346	15,0627	$< 10^{-4}$
	β_2	-0,7319	0,0419	-17,4564	$< 10^{-4}$
	β_3	0,2473	0,0489	5,0607	$< 10^{-4}$
	β_4	-0,6371	0,0543	-11,7305	$< 10^{-4}$
	σ^2	1,0146	0,1042	9,7415	$< 10^{-4}$

Tabela 81: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Complemento Log-log e remoção da categoria Caucasianos.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,2371	0,0301	-41,0526	$< 10^{-4}$
	β_1	0,5303	0,0321	16,5438	$< 10^{-4}$
	β_2	-0,7733	0,0365	-21,1687	$< 10^{-4}$
	β_3	0,2512	0,0471	5,3343	$< 10^{-4}$
	β_4	-0,7087	0,0497	-14,2687	$< 10^{-4}$
	b	4,3568	0,2000	21,7859	$< 10^{-4}$
Simplex	β_0	-1,2351	0,0327	-37,7561	$< 10^{-4}$
	β_1	0,5210	0,0346	15,0627	$< 10^{-4}$
	β_2	-0,7319	0,0419	-17,4564	$< 10^{-4}$
	β_3	0,2473	0,0489	5,0607	$< 10^{-4}$
	β_4	-0,6371	0,0543	-11,7305	$< 10^{-4}$
	σ^2	1,0146	0,1042	9,7415	$< 10^{-4}$

Tabela 82: Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Complemento Log-log e remoção da categoria Caucasianos.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,2371	0,0307	-40,3335	$< 10^{-4}$
	β_1	0,5303	0,0322	16,4485	$< 10^{-4}$
	β_2	-0,7733	0,0367	-21,0796	$< 10^{-4}$
	β_3	0,2512	0,0463	5,4215	$< 10^{-4}$
	β_4	-0,7087	0,0492	-14,4144	$< 10^{-4}$
	b	4,3568	0,1995	21,8381	$< 10^{-4}$
Simplex	β_0	-1,2351	0,0331	-37,3467	$< 10^{-4}$
	β_1	0,5210	0,0358	14,5409	$< 10^{-4}$
	β_2	-0,7319	0,0422	-17,3349	$< 10^{-4}$
	β_3	0,2473	0,0480	5,1495	$< 10^{-4}$
	β_4	-0,6371	0,0550	-11,5857	$< 10^{-4}$
	σ^2	1,0146	0,1046	9,6957	$< 10^{-4}$

Tabela 83: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Complemento Log-log e remoção da categoria Caucasianos.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,2371	0,0307	-40,3335	$< 10^{-4}$
	β_1	0,5303	0,0322	16,4485	$< 10^{-4}$
	β_2	-0,7733	0,0367	-21,0796	$< 10^{-4}$
	β_3	0,2512	0,0463	5,4215	$< 10^{-4}$
	β_4	-0,7087	0,0492	-14,4144	$< 10^{-4}$
	b	4,3568	0,1995	21,8381	$< 10^{-4}$
Simplex	β_0	-1,2351	0,0331	-37,3467	$< 10^{-4}$
	β_1	0,5210	0,0358	14,5409	$< 10^{-4}$
	β_2	-0,7319	0,0422	-17,3349	$< 10^{-4}$
	β_3	0,2473	0,0480	5,1495	$< 10^{-4}$
	β_4	-0,6371	0,0550	-11,5857	$< 10^{-4}$
	σ^2	1,0146	0,1046	9,6957	$< 10^{-4}$

Tabela 84: Medidas de qualidade do ajuste com todas a covariáveis e função de ligação Complemento Log-log.

Modelo de Regressão	$-l(\theta)$	AIC	BIC	R^2	R_a^2
L-logística	-519,4144	-1024,8289	-954,0012	0,6224	0,6207
Beta	-484,3211	-954,6422	-883,8145	0,6224	0,6206
Simplex	-512,2666	-1010,5331	-939,7054	0,6247	0,6229

Tabela 85: Medidas de qualidade do ajuste para a função de ligação Complemento Log-log e a remoção da categoria Caucasianos dos modelos de regressão L-logística e Simplex.

Modelo de Regressão	$-l(\theta)$	AIC	BIC	R^2	R_a^2
L-logística	-516,0725	-1020,1450	-959,4355	0,6187	0,6169
Beta	-484,3211	-954,6422	-883,8145	0,6224	0,6206
Simplex	-510,1031	-1008,2063	-947,4968	0,6190	0,6172

Tabela 86: Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Complemento Log-log.

Intervalo	Parâmetro	Limite Inferior	Limite Superior
Jackknife- <i>t</i>	β_0	-1,5197	-1,1916
	β_1	0,4749	0,5827
	β_2	-0,8249	-0,4830
	β_3	0,1939	0,5475
	β_4	-0,7685	-0,4102
	β_5	-0,0127	0,3271
	b	4,0508	4,6914
Pivotal	β_0	-1,5275	-1,1576
	β_1	0,4637	0,5917
	β_2	-0,8635	-0,4799
	β_3	0,1497	0,5628
	β_4	-0,8069	-0,3957
	β_5	-0,0610	0,3415
	b	3,8979	4,6834
Percentil	β_0	-1,5537	-1,1838
	β_1	0,4659	0,5939
	β_2	-0,8280	-0,4444
	β_3	0,1785	0,5916
	β_4	-0,7829	-0,3717
	β_5	-0,0271	0,3754
	b	4,0588	4,8443
BC	β_0	-1,5643	-1,1881
	β_1	0,4688	0,5954
	β_2	-0,8244	-0,4316
	β_3	0,1817	0,5990
	β_4	-0,7700	-0,3643
	β_5	-0,0202	0,3795
	b	3,9814	4,7394
BC _{<i>a</i>}	β_0	-1,565	-1,1881
	β_1	0,4688	0,5954
	β_2	-0,8244	-0,4309
	β_3	0,1817	0,6001
	β_4	-0,7685	-0,3638
	β_5	-0,0194	0,3799
	b	3,9617	4,7191

Tabela 87: Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Complemento Log-log e remoção da categoria Caucasianos.

Intervalo	Parâmetro	Limite Inferior	Limite Superior
Jackknife- <i>t</i>	β_0	-1,2876	-1,1865
	β_1	0,4772	0,5835
	β_2	-0,8338	-0,7128
	β_3	0,1748	0,3276
	β_4	-0,7897	-0,6276
	b	4,0279	4,6857
Pivotal	β_0	-1,2926	-1,1764
	β_1	0,4664	0,5902
	β_2	-0,8495	-0,7075
	β_3	0,1575	0,3457
	β_4	-0,8116	-0,6123
	b	3,9078	4,6892
Percentil	β_0	-1,2978	-1,1815
	β_1	0,4704	0,5943
	β_2	-0,8391	-0,6971
	β_3	0,1568	0,3449
	β_4	-0,805	-0,6057
	b	4,0243	4,8058
BC	β_0	-1,2980	-1,1816
	β_1	0,4713	0,5949
	β_2	-0,8391	-0,6967
	β_3	0,1502	0,3382
	β_4	-0,8021	-0,6017
	b	3,9654	4,7292
BC _{<i>a</i>}	β_0	-1,2985	-1,182
	β_1	0,4713	0,5949
	β_2	-0,8385	-0,6966
	β_3	0,1499	0,3380
	β_4	-0,8020	-0,6016
	b	3,9399	4,7096

Tabela 88: Estatísticas da distância do tipo Cook para todos os modelos e todas as funções de ligação. Perspectiva do teste da razão de verossimilhanças generalizadas.

Função de Ligação	Estatística	Modelo de Regressão		
		L-logística	Beta	Simplex
Logit	Mínimo	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
	Máximo	0,1740	0,7154	0,9634
	Média	0,0191	0,0123	0,0068
	Desvio Padrão	0,0445	0,0468	0,0479
	Mediana	0,0017	0,0011	0,0006
	IQR	0,0050	0,0043	0,0025
Probit	Mínimo	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
	Máximo	0,1820	0,7152	0,9637
	Média	0,0191	0,0124	0,0068
	Desvio Padrão	0,0445	0,0468	0,0479
	Mediana	0,0016	0,0010	0,0006
	IQR	0,0047	0,0044	0,0024
Complemento Log-log	Mínimo	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
	Máximo	0,1838	0,7189	0,9648
	Média	0,0189	0,0120	0,0066
	Desvio Padrão	0,0445	0,0469	0,0479
	Mediana	0,0016	0,0009	0,0005
	IQR	0,0048	0,0039	0,0024

Tabela 89: Estatísticas da distância do tipo Cook para todos os modelos e todas as funções de ligação. Perspectiva do teste do tipo Wald.

Função de Ligação	Estatística	Modelo de Regressão		
		L-logística	Beta	Simplex
Logit	Mínimo	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
	Máximo	0,1321	0,7154	0,8003
	Média	0,0347	0,0123	0,0131
	Desvio Padrão	0,0337	0,0468	0,0466
	Mediana	0,0228	0,0011	0,0033
	IQR	0,0550	0,0043	0,0116
Probit	Mínimo	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
	Máximo	0,1399	0,7152	0,8190
	Média	0,0345	0,0124	0,0131
	Desvio Padrão	0,0339	0,0468	0,0466
	Mediana	0,0228	0,0010	0,0034
	IQR	0,0525	0,0044	0,0117
Complemento Log-log	Mínimo	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
	Máximo	0,1249	0,7189	0,7803
	Média	0,0348	0,0120	0,0130
	Desvio Padrão	0,0335	0,0469	0,0466
	Mediana	0,0225	0,0009	0,0032
	IQR	0,0542	0,0039	0,0112

Figura 13: Gráfico da distância do tipo Cook para os modelo de regressão com função de ligação Logit. Perspectiva do teste da razão de verossimilhanças generalizada.

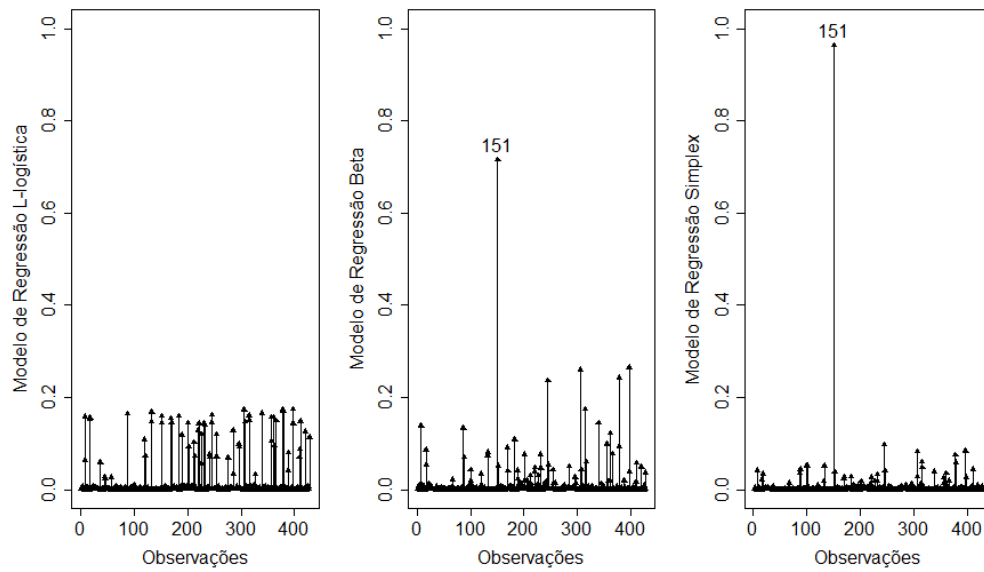


Figura 14: Gráfico da distância do tipo Cook para o modelo de regressão com função de ligação Probit. Perspectiva do teste da razão de verossimilhanças generalizada.

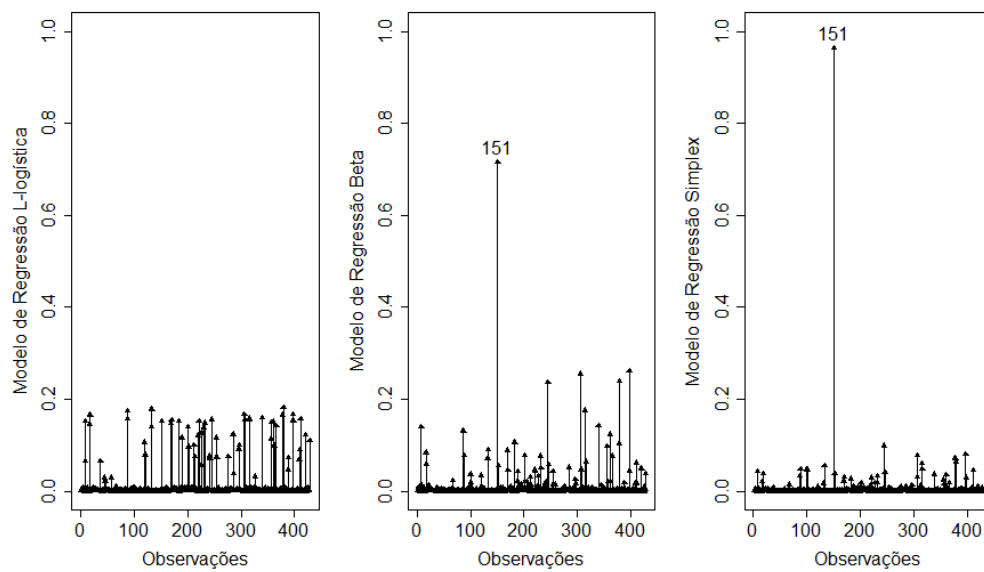


Figura 15: Gráfico da distância do tipo Cook para o modelo de regressão com função de ligação Complemento Log-log. Perspectiva do teste da razão de verossimilhanças generalizada.

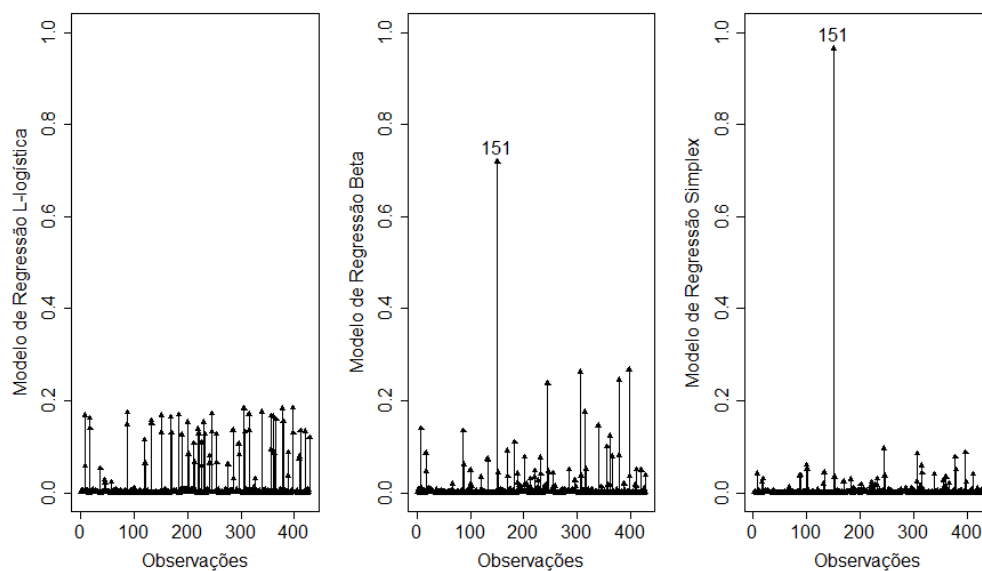


Figura 16: Gráfico da distância do tipo Cook para o modelo de regressão com função de ligação Logit. Perspectiva do teste do tipo Wald.

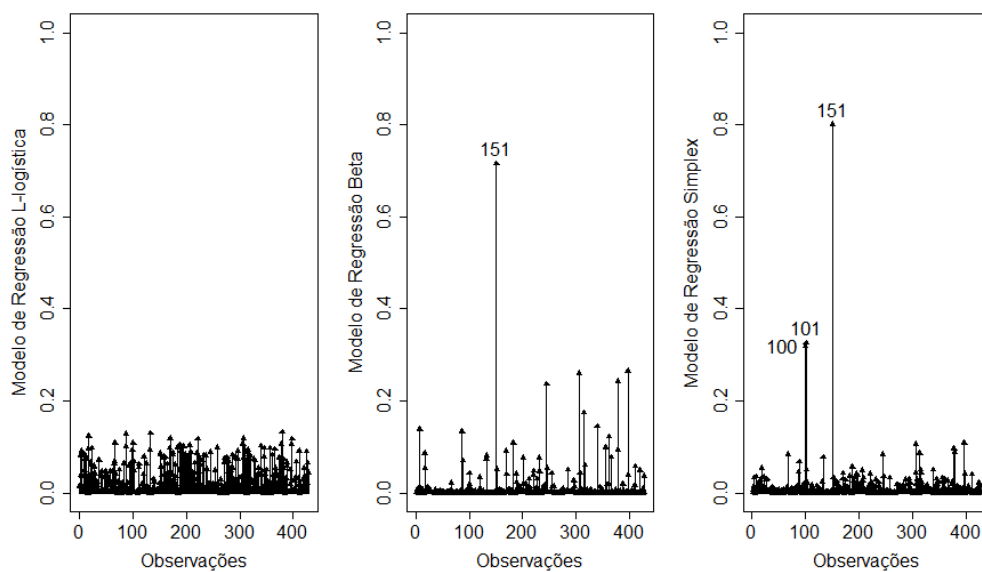


Figura 17: Gráfico da distância do tipo Cook para o modelo de regressão com função de ligação Probit. Perspectiva do teste do tipo Wald.

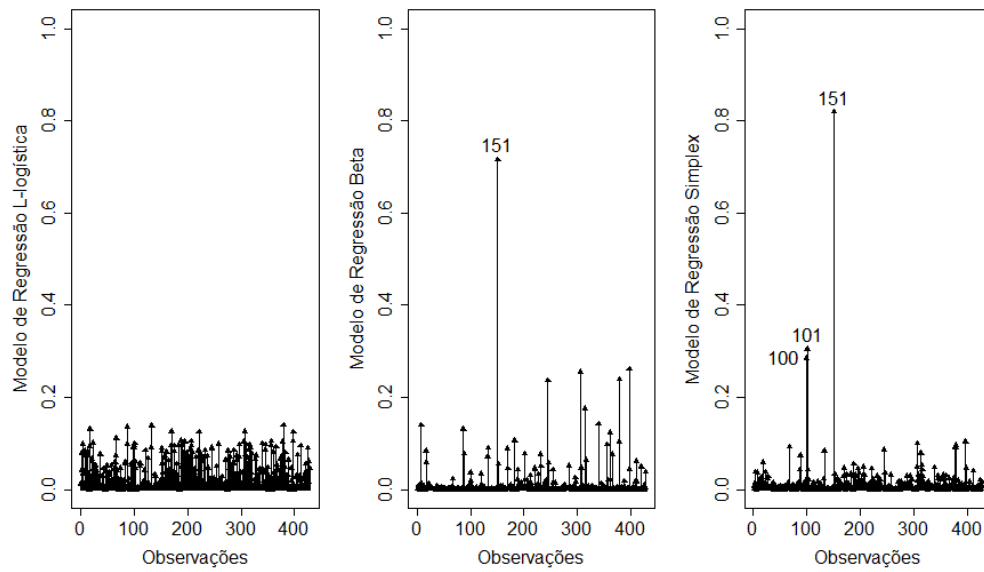
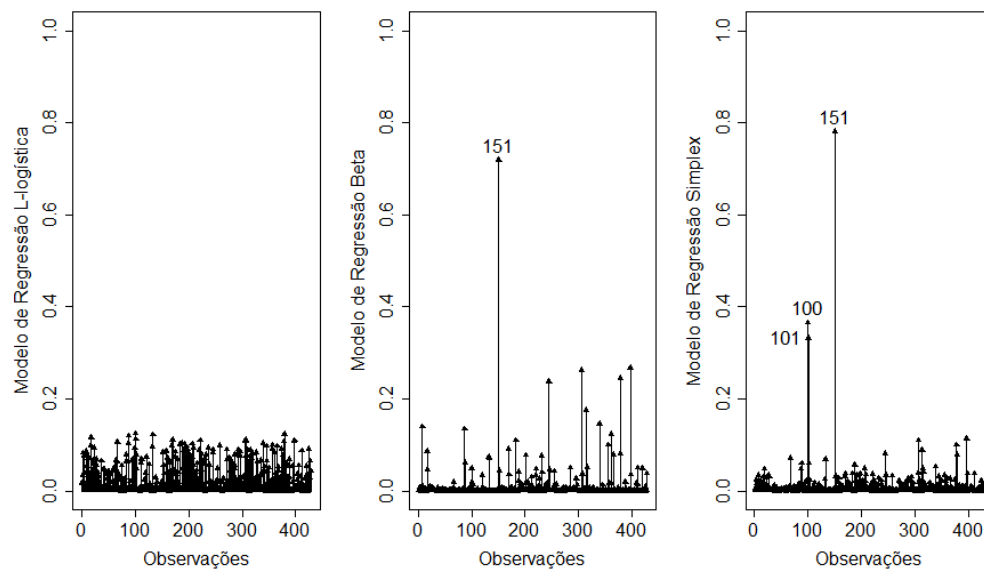


Figura 18: Gráfico da distância do tipo Cook para o modelo de regressão com função de ligação Complemento Log-log. Perspectiva do teste do tipo Wald.



5.3 Aplicação II

Nesta segunda aplicação utiliza-se dados de diferentes fontes para modelar o índice de percepção da corrupção (CPI16) para o ano de 2016. Como covariáveis do modelo, utiliza-se o indicador de liberdade econômica (EF16), o produto interno bruto per capita (PIBp16), o índice de democracia (ID16), o índice de desenvolvimento humano (IDH15), a densidade populacional (DP15) e o índice de inflação (II16). Todas as variáveis (indicadores) são referentes ao ano de 2016, com exceção da densidade populacional e do índice de desenvolvimento humano que são referentes ao ano de 2015. Cada covariável tem como finalidade descrever algum aspecto populacional dos países analisados e possuem graus de correlação com variável resposta CPI. A unificação destas informações acarreta uma base de dados com 147 registros (tamanho amostral) e vale apontar que embora a estruturação destes indicadores forneça subsídios para trabalharmos com modelos que envolvem erros nas variáveis, este trabalho não abordará esta temática. A seguir, na Tabela 90, apresenta-se a notação e as fontes de cada variável na modelagem. Logo após, na Tabela 91 e na Figura 19, as informações de correlação e o gráfico de dispersão entre a variável resposta e as covariáveis. Pode-se perceber que existe uma forte correlação entre a variável resposta e as covariáveis EF16, PIBp16, ID16 e IDH15. No gráfico de dispersão, pode-se observar tendências que diferenciam-se da linearidade. Com a exceção da covariável DP15, todas as outras demonstram uma aparente relação polinomial com a variável resposta. A seguir, na Tabela 92, apresenta-se as transformações utilizada para promover a tendência linear entre a variável resposta e as covariáveis PIBp16, ID16 e IDH15. Para as covariáveis restantes, não identificou-se transformações que melhorassem o comportamento linear com a variável resposta. Em seguida, na Tabela 93 e na Figura 20, as informações de correlação e os gráfico de dispersão entre a variável resposta CPI e as covariáveis. Pode-se perceber um aumento na correlação após a transformação das variáveis e uma melhora no comportamento linear no gráfico de dispersão.

Tabela 90: Informações das variáveis.

Variável	Nome	Fonte
y^*	CPI16	https://www.transparency.org
x_1^*	EF16	https://www.heritage.org/index/download
x_2^*	PIBp16	http://www.imf.org/external/datamapper/NGDPDPC@WEO/OEMDC/ADVEC/WEOWORLD
x_3^*	ID16	https://www.eiu.com/topic/democracy-index
x_4^*	IDH15	http://hdr.undp.org/en/data#
x_5^*	DP15	http://www.imf.org/external/datamapper/LP@WEO/OEMDC/ADVEC/WEOWORLD
x_6^*	II16	http://www.imf.org/external/datamapper/PCPIPCH@WEO/OEMDC/ADVEC/WEOWORLD

Tabela 91: Tabela de correlação das informações.

	y^*	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	x_6^*
y^*	1,0000	0,8158	0,8285	0,7316	0,7350	0,2217	-0,1991
x_1^*	0,8158	1,0000	0,6697	0,6430	0,6820	0,3385	-0,3239
x_2^*	0,8285	0,6697	1,0000	0,5703	0,7129	0,2366	-0,1042
x_3^*	0,7316	0,6430	0,5703	1,0000	0,6140	0,0422	-0,1074
x_4^*	0,7350	0,6820	0,7129	0,6140	1,0000	0,1734	-0,0536
x_5^*	0,2217	0,3386	0,2366	0,0422	0,1734	1,0000	-0,0409
x_6^*	-0,1991	-0,3239	-0,1042	-0,1074	-0,0536	-0,0409	1,0000

Figura 19: Gráfico de dispersão da variáveis.

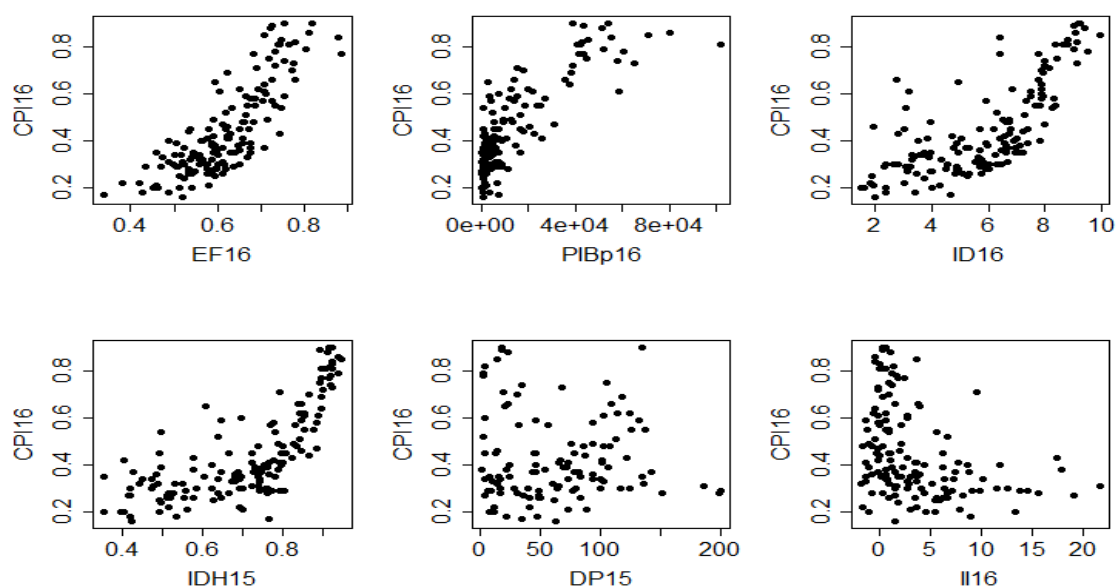


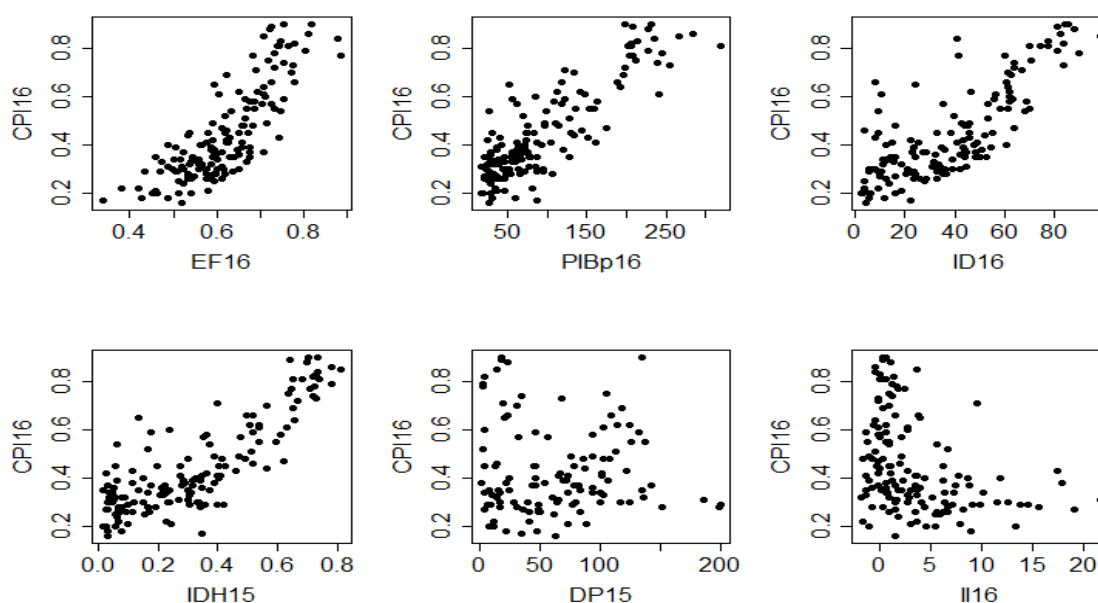
Tabela 92: Transformação das covariáveis.

Covariável	Transformação
x_1	x_1^*
x_2	$\sqrt{x_2^*}$
x_3	$(x_3^*)^2$
x_4	$(x_4^*)^4$
x_5	x_5^*
x_6	x_6^*

Tabela 93: Tabela de correlação das informações transformadas.

	y	x_1	x_2	x_3	x_4	x_5	x_6
y	1,0000	0,8158	0,8597	0,7876	0,8364	0,2217	-0,1991
x_1	0,8158	1,0000	0,7262	0,6596	0,7304	0,3386	-0,3239
x_2	0,8597	0,7262	1,0000	0,6936	0,9363	0,2350	-0,0970
x_3	0,7876	0,6596	0,6936	1,0000	0,7043	0,0243	-0,1240
x_4	0,8364	0,7304	0,9363	0,7043	1,0000	0,2194	-0,0840
x_5	0,2217	0,3386	0,2350	0,0243	0,2194	1,0000	-0,0409
x_6	-0,1991	-0,3239	-0,0970	-0,1240	-0,0840	-0,0409	1,0000

Figura 20: Gráfico de dispersão da variáveis após a transformação.



5.3.1 Modelos de Regressão: Resultados e Conclusões

Nesta seção, como na seção 5.2.1, utiliza-se a função de ligação Logit como cenário inicial para ilustrar o desempenho dos modelos com relação ao conjunto de dados. A seguir, na Equação 51, o predito linear que todos os modelos apresentam.

$$\eta_i = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \beta_3 x_{3,i} + \beta_4 x_{4,i} + \beta_5 x_{5,i} + \beta_6 x_{6,i}. \quad (51)$$

A significância dos parâmetros foi inicialmente verificada por meio do teste do tipo Wald com variância Bootstrap e Jackknife, respectivamente. Constatou-se que para uma precisão de 4 casas decimais, as metaheurísticas apresentaram o mesmo valor para as estimativas pontuais. Com relação às estimativas do desvio padrão, constatou-se que as metaheurísticas forneceram mesmos resultados para o método Bootstrap e para o método Jackknife. Além disso, assim como na seção 5.2.1, os métodos de reamostragem apresentaram estimativas relativamente próximas. No que diz respeito ao teste do tipo Wald, independente do método de reamostragem e dos modelos avaliados, os resultados indicaram significância somente para o intercepto e as variáveis x_1 , x_2 e x_3 para um nível de significância de 0,005. Estes resultados estão apresentados nas Tabelas 94 e 97.

Em seguida, utilizou-se o teste da razão de verossimilhanças generalizada no qual constatou-se a mesma conclusão do teste do tipo Wald para o nível de significância de 0,005. A seguir, na Tabela 98, os resultados do teste da razão de verossimilhanças generalizada. Após a constatação da significância

das variáveis, o processo de estimação foi refeito, assim como o teste do tipo Wald, e resultados podem ser vistos nas Tabelas 99 a 102. Para verificar a qualidade do ajuste dos modelos, novamente utilizou-se as medidas $-l(\theta)$, AIC, BIC, R^2 e R_a^2 . Constatou-se que para as medidas $-l(\theta)$, AIC e BIC o modelo de regressão L-logística demonstrou um melhor ajuste aos dados. Com relação às medidas R^2 e R_a^2 , o modelo de regressão Beta apresentou os valores 0,8525 e 0,8505, enquanto que o modelo de regressão L-logística apresentou os valores 0,8520 e 0,8499. Os resultados destas medidas estão bastante próximos e como citado na seção 5.2.1, estas medidas podem falhar na indicação da qualidade do ajuste. A seguir, na Tabela 103, os resultados das medidas de qualidade do ajuste. Logo após, nas Tabelas 104 e 105, os resultados dos métodos de reamostragem com relação aos intervalos de confiança para 95% para o modelo de regressão L-logística para o caso com todas as variáveis e o caso com a remoção das variáveis não significativas.

Alternativamente ao cenário com função de ligação Logit, utilizou-se as funções de ligação Probit e Complemento Log-log, no qual para o cenário com função de ligação Probit, constatou-se que não houve mudança no comportamento nos teste tipo Wald e no teste da razão de verossimilhanças generalizado. Com relação às medidas de qualidade de ajuste, novamente o modelo de regressão L-logística demonstrou uma melhor qualidade do ajuste, com exceção nas medidas R^2 e R_a^2 , no qual indicou-se o modelo de regressão Beta como melhor ajuste e o modelo de regressão L-logística com valores próximos. A seguir, nas Tabelas 106 a 114, os resultados dos teste do tipo Wald e do teste da razão de verossimilhanças generalizada. Em seguida, na Tabela 115, os resultados das medidas de qualidade de ajuste. Logo após, nas Tabelas 116 e 117, os resultados dos métodos de reamostragem com relação aos intervalos de confiança para 95%, assim como no cenário da função de ligação Logit. Com relação à função de ligação Complemento Log-log, constatou-se o mesmo comportamento das demais funções de ligação. Vale ressaltar que para um nível de significância de 0,05, na Tabela 122, o teste da razão de verossimilhanças generalizada apresentou nível descritivo no valor de 0,053 para o modelo de regressão Simplex. A seguir, Tabelas 118 a 126, os resultados do teste do tipo Wald. Em seguida, na Tabela 127, os resultados das medidas de qualidade de ajuste, no qual o modelo de regressão L-logística apresentou melhor qualidade do ajuste para as medidas $-l(\theta)$, AIC e BIC. Além disso, para as medidas R^2 e R_a^2 , o modelo de regressão Beta apresentou melhor qualidade de ajuste com valores próximos ao do modelo de regressão L-logística. A seguir, nas Tabelas 128 e 129, os intervalos de confiança para 95%, assim como no cenário da função de ligação Logit e Probit.

Com relação a identificação de observações atípicas, nas Figuras 21 a 23, apresentam-se os resultados gráficos da distância do tipo Cook, para todos os modelos de regressão e todas as funções de ligação. Em visão geral, para todas as funções de ligação, identificou-se valores atípicos em todos os modelos. Também constatou-se uma semelhança entre o cenário da função de ligação Logit e Probit, em termo das distâncias das observações. Com relação a função de ligação Complemento Log-log, verificou-se que a mesma ocasionou mais valores atípicos que as outras funções de ligação. No que diz respeito à performance dos modelos, independente da função de ligação, graficamente o modelo de regressão L-logística apresentou um menor número a valores atípicos. Dentre as observações atípicas detectadas, todos os modelos (independente da função de ligação) indicaram a observação 144 que corresponde as informações da Venezuela. Avaliando as informações deste país, no que concerne ao índice de percepção da corrupção de 2016 (CPI16), este país está na posição 146 da lista dos países menos corruptos do mundo (com um valor de 0,1700). No que diz respeito ao indicador de liberdade econômica de 2016 (EF16), dentre a lista dos países considerados livres economicamente, a Venezuela ocupa a posição 147 (com indicador no valor de 0,3370). Quanto ao PIB per capita de 2016 (PIBp16), dentre a lista dos países com os maiores valores, a Venezuela está na posição 57 com o valor de 7609,6200. Por fim, com relação ao índice de democracia de 2016 (ID16), a Venezuela está na posição 103 (com valor de 4,68) da lista dos países mais democráticos. Em uma visão global, considerando todos os países, nota-se que a Venezuela

sempre ocupa as últimas posições das listas citadas anteriormente. Em outras palavras, este país está sempre perto do terceiro quartil da lista de posições. Contudo, visando a covariável PIBp16, nota-se uma disparidade na posição. Este pode ser o indicativo do motivo pelo qual este país está sendo considerado como observação atípica por todos os modelos de regressão.

No que diz respeito à escolha da melhor modelagem para este conjunto de dados, constatou-se através das medidas de qualidade do ajuste AIC e BIC que o modelo de regressão L-logística apresentou melhor desempenho (independente da função de ligação). No que diz respeito à função de ligação mais adequada, a função Complemento Log-log apresentou maiores valores nas medidas de qualidade do ajuste. No entanto, com base na distância do tipo Cook, nota-se que esta função acarreta mais valores atípicos (pontos influentes) para modelo de regressão L-logística e os modelos de regressão Beta e Simplex. Assim, embora o uso da mediana torne o modelo de regressão L-logística mais robusto a valores atípicos, torna-se mais adequado o uso da função de ligação Logit ou Probit. Com relação ao processo de escolha de uma dessas funções de ligação, verifica-se que a função de ligação Probit apresenta valores um pouco inferiores aos valores da função de ligação Logit para as medidas de qualidade do ajuste AIC e BIC. Contudo, visando a distância do tipo Cook, verifica-se que a função de ligação Probit também apresenta valores um pouco menores em estatísticas tais como média, mediana, máximo e intervalo interquartil. Esta pequena diferença nas estatísticas pode ser o motivo pelo qual o modelo de regressão L-logística com função de ligação Probit apresente as medidas R^2 e R_a^2 um pouco maiores do que o mesmo modelo com função de ligação Logit. Assim, visando a distância do tipo Cook e a proposta de robustez do modelo de regressão L-logística, opta-se pela escolha da função de ligação Probit como melhor opção para a modelagem deste conjunto de dados. O leitor interessado, pode consultar as estatísticas das distâncias do tipo Cook na Tabela 130 logo a frente.

Tabela 94: Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Logit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,5337	0,2345	-10,8052	$< 10^{-4}$
	β_1	2,1809	0,4763	4,5793	$< 10^{-4}$
	β_2	0,0046	0,0011	4,0352	0,0001
	β_3	0,0107	0,0018	5,9554	$< 10^{-4}$
	β_4	0,4588	0,3970	1,1559	0,2477
	β_5	$< 10^{-4}$	0,0001	0,1936	0,8465
	β_6	-0,0019	0,0034	-0,5505	0,5820
	b	5,2843	0,3804	13,8919	$< 10^{-4}$
Beta	β_0	-2,6053	0,2378	-10,9551	$< 10^{-4}$
	β_1	2,4004	0,4869	4,9303	$< 10^{-4}$
	β_2	0,0048	0,0012	4,1248	$< 10^{-4}$
	β_3	0,0100	0,0017	5,8735	$< 10^{-4}$
	β_4	0,2792	0,3922	0,7119	0,4765
	β_5	$< 10^{-4}$	0,0001	0,0569	0,9546
	β_6	-0,0013	0,0032	-0,4002	0,6890
	ϕ	40,4936	6,6720	6,0692	$< 10^{-4}$
Simplex	β_0	-2,6707	0,2417	-11,0503	$< 10^{-4}$
	β_1	2,5457	0,5039	5,0522	$< 10^{-4}$
	β_2	0,0043	0,0011	3,9500	0,0001
	β_3	0,0104	0,0018	5,8817	$< 10^{-4}$
	β_4	0,4001	0,3766	1,0622	0,2882
	β_5	$< 10^{-4}$	0,0001	0,1430	0,8863
	β_6	-0,0017	0,0031	-0,5526	0,5805
	σ^2	0,5833	0,0857	6,8061	$< 10^{-4}$

Tabela 95: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Logit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,5337	0,2345	-10,8052	$< 10^{-4}$
	β_1	2,1809	0,4763	4,5793	$< 10^{-4}$
	β_2	0,0046	0,0011	4,0352	0,0001
	β_3	0,0107	0,0018	5,9554	$< 10^{-4}$
	β_4	0,4588	0,3970	1,1559	0,2477
	β_5	$< 10^{-4}$	0,0001	0,1936	0,8465
	β_6	-0,0019	0,0034	-0,5505	0,5820
	b	5,2843	0,3804	13,8919	$< 10^{-4}$
Beta	β_0	-2,6053	0,2378	-10,9551	$< 10^{-4}$
	β_1	2,4004	0,4869	4,9303	$< 10^{-4}$
	β_2	0,0048	0,0012	4,1248	$< 10^{-4}$
	β_3	0,0100	0,0017	5,8735	$< 10^{-4}$
	β_4	0,2792	0,3922	0,7119	0,4765
	β_5	$< 10^{-4}$	0,0001	0,0569	0,9546
	β_6	-0,0013	0,0032	-0,4002	0,6890
	ϕ	40,4936	6,6720	6,0692	$< 10^{-4}$
Simplex	β_0	-2,6707	0,2417	-11,0503	$< 10^{-4}$
	β_1	2,5457	0,5039	5,0522	$< 10^{-4}$
	β_2	0,0043	0,0011	3,9500	0,0001
	β_3	0,0104	0,0018	5,8817	$< 10^{-4}$
	β_4	0,4001	0,3766	1,0622	0,2882
	β_5	$< 10^{-4}$	0,0001	0,1430	0,8863
	β_6	-0,0017	0,0031	-0,5526	0,5805
	σ^2	0,5833	0,0857	6,8061	$< 10^{-4}$

Tabela 96: Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Logit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,5337	0,2374	-10,6748	$< 10^{-4}$
	β_1	2,1809	0,4802	4,5414	$< 10^{-4}$
	β_2	0,0046	0,0011	4,1404	$< 10^{-4}$
	β_3	0,0107	0,0018	5,9911	$< 10^{-4}$
	β_4	0,4588	0,4022	1,1408	0,2539
	β_5	$< 10^{-4}$	$< 10^{-4}$	0,5020	0,6156
	β_6	-0,0019	0,0043	-0,4314	0,6662
	b	5,2843	0,3719	14,2106	$< 10^{-4}$
Beta	β_0	-2,6053	0,2400	-10,8549	$< 10^{-4}$
	β_1	2,4004	0,4881	4,9178	$< 10^{-4}$
	β_2	0,0048	0,0012	4,1035	$< 10^{-4}$
	β_3	0,0100	0,0017	5,7507	$< 10^{-4}$
	β_4	0,2792	0,4013	0,6957	0,4866
	β_5	$< 10^{-4}$	$< 10^{-4}$	0,1470	0,8831
	β_6	-0,0013	0,0037	-0,3402	0,7337
	ϕ	40,4936	6,7347	6,0126	$< 10^{-4}$
Simplex	β_0	-2,6707	0,2411	-11,0755	$< 10^{-4}$
	β_1	2,5457	0,5037	5,0536	$< 10^{-4}$
	β_2	0,0043	0,0011	4,0499	0,0001
	β_3	0,0104	0,0018	5,8912	$< 10^{-4}$
	β_4	0,4001	0,3790	1,0555	0,2912
	β_5	$< 10^{-4}$	0,0001	0,3205	0,7486
	β_6	-0,0017	0,0030	-0,5552	0,5787
	σ^2	0,5833	0,0927	6,2913	$< 10^{-4}$

Tabela 97: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Logit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,5337	0,2374	-10,6748	$< 10^{-4}$
	β_1	2,1809	0,4802	4,5414	$< 10^{-4}$
	β_2	0,0046	0,0011	4,1404	$< 10^{-4}$
	β_3	0,0107	0,0018	5,9911	$< 10^{-4}$
	β_4	0,4588	0,4022	1,1408	0,2539
	β_5	$< 10^{-4}$	$< 10^{-4}$	0,5020	0,6156
	β_6	-0,0019	0,0043	-0,4314	0,6662
	b	5,2843	0,3719	14,2106	$< 10^{-4}$
Beta	β_0	-2,6053	0,2400	-10,8549	$< 10^{-4}$
	β_1	2,4004	0,4881	4,9178	$< 10^{-4}$
	β_2	0,0048	0,0012	4,1035	$< 10^{-4}$
	β_3	0,0100	0,0017	5,7507	$< 10^{-4}$
	β_4	0,2792	0,4013	0,6957	0,4866
	β_5	$< 10^{-4}$	$< 10^{-4}$	0,1470	0,8831
	β_6	-0,0013	0,0037	-0,3402	0,7337
	ϕ	40,4936	6,7347	6,0126	$< 10^{-4}$
Simplex	β_0	-2,6707	0,2411	-11,0755	$< 10^{-4}$
	β_1	2,5457	0,5037	5,0536	$< 10^{-4}$
	β_2	0,0043	0,0011	4,0499	0,0001
	β_3	0,0104	0,0018	5,8912	$< 10^{-4}$
	β_4	0,4001	0,3790	1,0555	0,2912
	β_5	$< 10^{-4}$	0,0001	0,3205	0,7486
	β_6	-0,0017	0,0030	-0,5552	0,5787
	σ^2	0,5833	0,0927	6,2913	$< 10^{-4}$

Tabela 98: Teste da Razão de Verossimilhanças Generalizada para a função de ligação Logit.

Modelo de Regressão	Hipótese	$l(\theta)$	Graus de Liberdade	Estatística	Valor-P
L-logística	H_0	181,2172	4	2,2071	0,6977
	H_1	182,3207			
Beta	H_0	179,2161	4	0,7685	0,9426
	H_1	179,6004			
Simplex	H_0	177,8456	4	1,4849	0,2230
	H_1	178,5880			

Tabela 99: Teste tipo Wald com combinação do Algoritmo Genético com método Bootstrap para a função de ligação Logit e remoção das covariáveis x_4 , x_5 e x_6 .

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,7584	0,2200	-12,5371	$< 10^{-4}$
	β_1	2,5759	0,4492	5,7344	$< 10^{-4}$
	β_2	0,0056	0,0007	8,2336	$< 10^{-4}$
	β_3	0,0107	0,0017	6,3113	$< 10^{-4}$
	b	5,2076	0,3522	14,7848	$< 10^{-4}$
Beta	β_0	-2,7364	0,2000	-13,6846	$< 10^{-4}$
	β_1	2,6157	0,4182	6,2550	$< 10^{-4}$
	β_2	0,0054	0,0006	8,4472	$< 10^{-4}$
	β_3	0,0101	0,0016	6,4102	$< 10^{-4}$
	ϕ	40,2014	6,2718	6,4098	$< 10^{-4}$
Simplex	β_0	-2,8572	0,2084	-13,7077	$< 10^{-4}$
	β_1	2,8694	0,4361	6,5789	$< 10^{-4}$
	β_2	0,0052	0,0006	8,6833	$< 10^{-4}$
	β_3	0,0103	0,0017	6,1482	$< 10^{-4}$
	σ^2	0,593	0,0849	6,9818	$< 10^{-4}$

Tabela 100: Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Bootstrap para a função de ligação Logit e remoção das covariáveis x_4 , x_5 e x_6 .

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,7584	0,2200	-12,5371	$< 10^{-4}$
	β_1	2,5759	0,4492	5,7344	$< 10^{-4}$
	β_2	0,0056	0,0007	8,2336	$< 10^{-4}$
	β_3	0,0107	0,0017	6,3113	$< 10^{-4}$
	b	5,2076	0,3522	14,7848	$< 10^{-4}$
Beta	β_0	-2,7364	0,2000	-13,6846	$< 10^{-4}$
	β_1	2,6157	0,4182	6,2550	$< 10^{-4}$
	β_2	0,0054	0,0006	8,4472	$< 10^{-4}$
	β_3	0,0101	0,0016	6,4102	$< 10^{-4}$
	ϕ	40,2014	6,2718	6,4098	$< 10^{-4}$
Simplex	β_0	-2,8572	0,2084	-13,7077	$< 10^{-4}$
	β_1	2,8694	0,4361	6,5789	$< 10^{-4}$
	β_2	0,0052	0,0006	8,6833	$< 10^{-4}$
	β_3	0,0103	0,0017	6,1482	$< 10^{-4}$
	σ^2	0,593	0,0849	6,9818	$< 10^{-4}$

Tabela 101: Teste tipo Wald com combinação do Algoritmo Genético com método Jackknife para a função de ligação Logit e remoção das covariáveis x_4 , x_5 e x_6 .

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,7584	0,2282	-12,0856	$< 10^{-4}$
	β_1	2,5759	0,4676	5,5093	$< 10^{-4}$
	β_2	0,0056	0,0007	8,2031	$< 10^{-4}$
	β_3	0,0107	0,0017	6,2520	$< 10^{-4}$
	b	5,2076	0,3548	14,6762	$< 10^{-4}$
Beta	β_0	-2,7364	0,2021	-13,5386	$< 10^{-4}$
	β_1	2,6157	0,4214	6,2066	$< 10^{-4}$
	β_2	0,0054	0,0006	8,4759	$< 10^{-4}$
	β_3	0,0101	0,0016	6,2311	$< 10^{-4}$
	ϕ	40,2014	6,4731	6,2105	$< 10^{-4}$
Simplex	β_0	-2,8572	0,2116	-13,5025	$< 10^{-4}$
	β_1	2,8694	0,4450	6,4483	$< 10^{-4}$
	β_2	0,0052	0,0006	8,5657	$< 10^{-4}$
	β_3	0,0103	0,0017	6,0482	$< 10^{-4}$
	σ^2	0,5930	0,0900	6,5919	$< 10^{-4}$

Tabela 102: Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Jackknife para a função de ligação Logit e remoção das covariáveis x_4 , x_5 e x_6 .

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,7584	0,2282	-12,0856	$< 10^{-4}$
	β_1	2,5759	0,4676	5,5093	$< 10^{-4}$
	β_2	0,0056	0,0007	8,2031	$< 10^{-4}$
	β_3	0,0107	0,0017	6,2520	$< 10^{-4}$
	b	5,2076	0,3548	14,6762	$< 10^{-4}$
Beta	β_0	-2,7364	0,2021	-13,5386	$< 10^{-4}$
	β_1	2,6157	0,4214	6,2066	$< 10^{-4}$
	β_2	0,0054	0,0006	8,4759	$< 10^{-4}$
	β_3	0,0101	0,0016	6,2311	$< 10^{-4}$
	ϕ	40,2014	6,4731	6,2105	$< 10^{-4}$
Simplex	β_0	-2,8572	0,2116	-13,5025	$< 10^{-4}$
	β_1	2,8694	0,4450	6,4483	$< 10^{-4}$
	β_2	0,0052	0,0006	8,5657	$< 10^{-4}$
	β_3	0,0103	0,0017	6,0482	$< 10^{-4}$
	σ^2	0,5930	0,0900	6,5919	$< 10^{-4}$

Tabela 103: Medidas de qualidade do ajuste para a função de ligação Logit.

Modelo de Regressão	$-l(\theta)$	AIC	BIC	R^2	R_a^2
L-logística	-181,2172	-352,4344	-312,5301	0,8520	0,8499
Beta	-179,2161	-348,4322	-308,5279	0,8525	0,8505
Simplex	-177,8456	-345,6912	-305,7868	0,8503	0,8482

Tabela 104: Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Logit.

Intervalo	Parâmetro	Limite Inferior	Limite Superior
Jackknife- <i>t</i>	β_0	-2,9753	-2,1895
	β_1	1,4511	3,0411
	β_2	0,0037	0,0073
	β_3	0,0081	0,0140
	β_4	-0,5819	0,7497
	β_5	-0,0001	0,0001
	β_6	-0,0089	0,0054
	b	4,6443	5,8754
Pivotal	β_0	-3,0731	-2,1507
	β_1	1,3776	3,2000
	β_2	0,0036	0,0082
	β_3	0,0080	0,0149
	β_4	-1,0007	0,5845
	β_5	$< 10^{-4}$	0,0004
	β_6	-0,0114	0,0019
	b	4,2528	5,7625
Percentil	β_0	-3,0141	-2,0918
	β_1	1,2923	3,1146
	β_2	0,0028	0,0074
	β_3	0,0071	0,0141
	β_4	-0,4167	1,1685
	β_5	-0,0003	0,0001
	β_6	-0,0055	0,0079
	b	4,7572	6,2668
BC	β_0	-3,1411	-2,1753
	β_1	1,4046	3,2320
	β_2	0,0042	0,0089
	β_3	0,0076	0,0145
	β_4	-0,8818	0,5255
	β_5	-0,0003	0,0001
	β_6	-0,0073	0,0064
	b	4,5293	5,8119
BC _{<i>a</i>}	β_0	-3,1504	-2,1843
	β_1	1,4083	3,2605
	β_2	0,0041	0,0088
	β_3	0,0076	0,0145
	β_4	-0,8697	0,5307
	β_5	-0,0002	0,0002
	β_6	-0,0150	0,0044
	b	4,4865	5,7841

Tabela 105: Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Logit e remoção das covariáveis x_4 , x_5 e x_6 .

Intervalo	Parâmetro	Limite Inferior	Limite Superior
Jackknife- t	β_0	-3,1362	-2,3806
	β_1	1,8020	3,3499
	β_2	0,0045	0,0068
	β_3	0,0078	0,0135
	b	4,6202	5,7950
Pivotal	β_0	-3,2131	-2,3563
	β_1	1,8137	3,4978
	β_2	0,0041	0,0069
	β_3	0,0074	0,0140
	b	4,3381	5,7499
Percentil	β_0	-3,1605	-2,3037
	β_1	1,6541	3,3382
	β_2	0,0044	0,0071
	β_3	0,0073	0,0139
	b	4,6653	6,0771
BC	β_0	-3,1727	-2,3066
	β_1	1,6691	3,3478
	β_2	0,0044	0,0071
	β_3	0,0071	0,0137
	b	4,5233	5,8080
BC $_a$	β_0	-3,1848	-2,3264
	β_1	1,6918	3,3700
	β_2	0,0043	0,0070
	β_3	0,0071	0,0137
	b	4,5026	5,7901

Tabela 106: Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Probit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,5784	0,1402	-11,2621	$< 10^{-4}$
	β_1	1,3680	0,2899	4,7195	$< 10^{-4}$
	β_2	0,0033	0,0007	4,6808	$< 10^{-4}$
	β_3	0,0067	0,0011	6,2598	$< 10^{-4}$
	β_4	0,0882	0,2166	0,4070	0,6840
	β_5	$< 10^{-4}$	0,0001	0,2154	0,8295
	β_6	-0,0011	0,0020	-0,5434	0,5868
	b	5,2548	0,3749	14,0183	$< 10^{-4}$
Beta	β_0	-1,6228	0,1441	-11,2628	$< 10^{-4}$
	β_1	1,5056	0,2999	5,0205	$< 10^{-4}$
	β_2	0,0035	0,0007	5,1603	$< 10^{-4}$
	β_3	0,0063	0,0010	6,2211	$< 10^{-4}$
	β_4	-0,0327	0,2101	-0,1554	0,8765
	β_5	$< 10^{-4}$	0,0001	0,0667	0,9468
	β_6	-0,0007	0,0019	-0,3922	0,6949
	ϕ	40,4541	6,5213	6,2034	$< 10^{-4}$
Simplex	β_0	-1,6553	0,1448	-11,4323	$< 10^{-4}$
	β_1	1,5790	0,3060	5,1603	$< 10^{-4}$
	β_2	0,0030	0,0006	4,5875	$< 10^{-4}$
	β_3	0,0066	0,0011	5,9845	$< 10^{-4}$
	β_4	0,0890	0,2035	0,4375	0,6618
	β_5	$< 10^{-4}$	0,0001	0,1768	0,8597
	β_6	-0,0010	0,0018	-0,5349	0,5927
	σ^2	0,5860	0,0854	6,8613	$< 10^{-4}$

Tabela 107: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Probit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,5784	0,1402	-11,2621	$< 10^{-4}$
	β_1	1,3680	0,2899	4,7195	$< 10^{-4}$
	β_2	0,0033	0,0007	4,6808	$< 10^{-4}$
	β_3	0,0067	0,0011	6,2598	$< 10^{-4}$
	β_4	0,0882	0,2166	0,4070	0,6840
	β_5	$< 10^{-4}$	0,0001	0,2154	0,8295
	β_6	-0,0011	0,0020	-0,5434	0,5868
	b	5,2548	0,3749	14,0183	$< 10^{-4}$
Beta	β_0	-1,6228	0,1441	-11,2628	$< 10^{-4}$
	β_1	1,5056	0,2999	5,0205	$< 10^{-4}$
	β_2	0,0035	0,0007	5,1603	$< 10^{-4}$
	β_3	0,0063	0,0010	6,2211	$< 10^{-4}$
	β_4	-0,0327	0,2101	-0,1554	0,8765
	β_5	$< 10^{-4}$	0,0001	0,0667	0,9468
	β_6	-0,0007	0,0019	-0,3922	0,6949
	ϕ	40,4541	6,5213	6,2034	$< 10^{-4}$
Simplex	β_0	-1,6553	0,1448	-11,4323	$< 10^{-4}$
	β_1	1,5790	0,3060	5,1603	$< 10^{-4}$
	β_2	0,0030	0,0006	4,5875	$< 10^{-4}$
	β_3	0,0066	0,0011	5,9845	$< 10^{-4}$
	β_4	0,0890	0,2035	0,4375	0,6618
	β_5	$< 10^{-4}$	0,0001	0,1768	0,8597
	β_6	-0,0010	0,0018	-0,5349	0,5927
	σ^2	0,5860	0,0854	6,8613	$< 10^{-4}$

Tabela 108: Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Probit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,5784	0,1409	-11,2003	$< 10^{-4}$
	β_1	1,3680	0,2913	4,6959	$< 10^{-4}$
	β_2	0,0033	0,0007	4,5354	$< 10^{-4}$
	β_3	0,0067	0,0011	6,3134	$< 10^{-4}$
	β_4	0,0882	0,2254	0,3911	0,6957
	β_5	$< 10^{-4}$	$< 10^{-4}$	0,5872	0,5571
	β_6	-0,0011	0,0022	-0,4808	0,6307
	b	5,2548	0,3662	14,3502	$< 10^{-4}$
Beta	β_0	-1,6228	0,1448	-11,205	$< 10^{-4}$
	β_1	1,5056	0,3006	5,0093	$< 10^{-4}$
	β_2	0,0035	0,0007	5,0098	$< 10^{-4}$
	β_3	0,0063	0,001	6,0796	$< 10^{-4}$
	β_4	-0,0327	0,2178	-0,1499	0,8808
	β_5	$< 10^{-4}$	$< 10^{-4}$	0,1795	0,8575
	β_6	-0,0007	0,0019	-0,3833	0,7015
	ϕ	40,4541	6,5309	6,1943	$< 10^{-4}$
Simplex	β_0	-1,6553	0,1840	-8,9977	$< 10^{-4}$
	β_1	1,5790	0,5163	3,0582	0,0022
	β_2	0,0030	0,0013	2,3185	0,0204
	β_3	0,0066	0,0052	1,2702	0,2040
	β_4	0,0890	0,4310	0,2065	0,8364
	β_5	$< 10^{-4}$	$< 10^{-4}$	0,3065	0,7593
	β_6	-0,0010	0,0034	-0,2829	0,7772
	σ^2	0,5860	0,1387	4,2262	$< 10^{-4}$

Tabela 109: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Probit.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-1,5784	0,1409	-11,2003	$< 10^{-4}$
	β_1	1,3680	0,2913	4,6959	$< 10^{-4}$
	β_2	0,0033	0,0007	4,5354	$< 10^{-4}$
	β_3	0,0067	0,0011	6,3134	$< 10^{-4}$
	β_4	0,0882	0,2254	0,3911	0,6957
	β_5	$< 10^{-4}$	$< 10^{-4}$	0,5872	0,5571
	β_6	-0,0011	0,0022	-0,4808	0,6307
	b	5,2548	0,3662	14,3502	$< 10^{-4}$
Beta	β_0	-1,6228	0,1448	-11,205	$< 10^{-4}$
	β_1	1,5056	0,3006	5,0093	$< 10^{-4}$
	β_2	0,0035	0,0007	5,0098	$< 10^{-4}$
	β_3	0,0063	0,001	6,0796	$< 10^{-4}$
	β_4	-0,0327	0,2178	-0,1499	0,8808
	β_5	$< 10^{-4}$	$< 10^{-4}$	0,1795	0,8575
	β_6	-0,0007	0,0019	-0,3833	0,7015
	ϕ	40,4541	6,5309	6,1943	$< 10^{-4}$
Simplex	β_0	-1,6553	0,1840	-8,9977	$< 10^{-4}$
	β_1	1,5790	0,5163	3,0582	0,0022
	β_2	0,0030	0,0013	2,3185	0,0204
	β_3	0,0066	0,0052	1,2702	0,2040
	β_4	0,0890	0,4310	0,2065	0,8364
	β_5	$< 10^{-4}$	$< 10^{-4}$	0,3065	0,7593
	β_6	-0,0010	0,0034	-0,2829	0,7772
	σ^2	0,5860	0,1387	4,2262	$< 10^{-4}$

Tabela 110: Teste da Razão de Verossimilhanças Generalizada para a função de ligação Probit.

Modelo de Regressão	Hipótese	$l(\theta)$	Gráus de Liberdade	Estatística	Valor-P
L-logística	H_0	181,0120	4	2,4790	0,6484
	H_1	182,2515			
Beta	H_0	179,2611	4	0,7685	0,9420
	H_1	179,6479			
Simplex	H_0	177,8389	4	1,7465	0,1863
	H_1	178,7121			

Tabela 111: Teste tipo Wald com combinação do Algoritmo Genético com método Bootstrap para a função de ligação Probit e remoção das covariáveis x_4 , x_5 e x_6 .

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,5639	0,1556	-16,4748	$< 10^{-4}$
	β_1	2,2245	0,3104	7,1657	$< 10^{-4}$
	β_2	0,0032	0,0005	6,2483	$< 10^{-4}$
	β_3	0,0078	0,0013	6,2352	$< 10^{-4}$
	b	5,3122	0,3506	15,1531	$< 10^{-4}$
Beta	β_0	-2,4961	0,1394	-17,9056	$< 10^{-4}$
	β_1	2,1433	0,2818	7,6068	$< 10^{-4}$
	β_2	0,0033	0,0005	7,3979	$< 10^{-4}$
	β_3	0,0074	0,0011	6,5873	$< 10^{-4}$
	ϕ	42,9422	6,6706	6,4375	$< 10^{-4}$
Simplex	β_0	-2,6099	0,1461	-17,8693	$< 10^{-4}$
	β_1	2,3989	0,2864	8,3765	$< 10^{-4}$
	β_2	0,0028	0,0004	6,4995	$< 10^{-4}$
	β_3	0,0078	0,0013	6,2189	$< 10^{-4}$
	σ^2	0,5679	0,0807	7,0340	$< 10^{-4}$

Tabela 112: Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Bootstrap para a função de ligação Probit e remoção das covariáveis x_4 , x_5 e x_6 .

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,5639	0,1556	-16,4748	$< 10^{-4}$
	β_1	2,2245	0,3104	7,1657	$< 10^{-4}$
	β_2	0,0032	0,0005	6,2483	$< 10^{-4}$
	β_3	0,0078	0,0013	6,2352	$< 10^{-4}$
	b	5,3122	0,3506	15,1531	$< 10^{-4}$
Beta	β_0	-2,4961	0,1394	-17,9056	$< 10^{-4}$
	β_1	2,1433	0,2818	7,6068	$< 10^{-4}$
	β_2	0,0033	0,0005	7,3979	$< 10^{-4}$
	β_3	0,0074	0,0011	6,5873	$< 10^{-4}$
	ϕ	42,9422	6,6706	6,4375	$< 10^{-4}$
Simplex	β_0	-2,6099	0,1461	-17,8693	$< 10^{-4}$
	β_1	2,3989	0,2864	8,3765	$< 10^{-4}$
	β_2	0,0028	0,0004	6,4995	$< 10^{-4}$
	β_3	0,0078	0,0013	6,2189	$< 10^{-4}$
	σ^2	0,5679	0,0807	7,0340	$< 10^{-4}$

Tabela 113: Teste tipo Wald com combinação do Algoritmo Genético com método Jackknife para a função de ligação Probit e remoção das covariáveis x_4 , x_5 e x_6 .

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,5639	0,156	-16,4361	$< 10^{-4}$
	β_1	2,2245	0,3105	7,165	$< 10^{-4}$
	β_2	0,0032	0,0005	6,0999	$< 10^{-4}$
	β_3	0,0078	0,0012	6,3314	$< 10^{-4}$
	b	5,3122	0,3512	15,1242	$< 10^{-4}$
Beta	β_0	-2,4961	0,1393	-17,9246	$< 10^{-4}$
	β_1	2,1433	0,2815	7,6127	$< 10^{-4}$
	β_2	0,0033	0,0005	7,3395	$< 10^{-4}$
	β_3	0,0074	0,0011	6,4989	$< 10^{-4}$
	ϕ	42,9422	6,9739	6,1575	$< 10^{-4}$
Simplex	β_0	-2,6099	0,1469	-17,7659	$< 10^{-4}$
	β_1	2,3989	0,2895	8,2871	$< 10^{-4}$
	β_2	0,0028	0,0004	6,2945	$< 10^{-4}$
	β_3	0,0078	0,0013	6,0951	$< 10^{-4}$
	σ^2	0,5679	0,0856	6,6361	$< 10^{-4}$

Tabela 114: Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Jackknife para a função de ligação Probit e remoção das covariáveis x_4 , x_5 e x_6 .

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,5639	0,156	-16,4361	$< 10^{-4}$
	β_1	2,2245	0,3105	7,165	$< 10^{-4}$
	β_2	0,0032	0,0005	6,0999	$< 10^{-4}$
	β_3	0,0078	0,0012	6,3314	$< 10^{-4}$
	b	5,3122	0,3512	15,1242	$< 10^{-4}$
Beta	β_0	-2,4961	0,1393	-17,9246	$< 10^{-4}$
	β_1	2,1433	0,2815	7,6127	$< 10^{-4}$
	β_2	0,0033	0,0005	7,3395	$< 10^{-4}$
	β_3	0,0074	0,0011	6,4989	$< 10^{-4}$
	ϕ	42,9422	6,9739	6,1575	$< 10^{-4}$
Simplex	β_0	-2,6099	0,1469	-17,7659	$< 10^{-4}$
	β_1	2,3989	0,2895	8,2871	$< 10^{-4}$
	β_2	0,0028	0,0004	6,2945	$< 10^{-4}$
	β_3	0,0078	0,0013	6,0951	$< 10^{-4}$
	σ^2	0,5679	0,0856	6,6361	$< 10^{-4}$

Tabela 115: Medidas de qualidade do ajuste para a função de ligação Probit.

Modelo de Regressão	$-l(\theta)$	AIC	BIC	R^2	R_a^2
L-logística	-181,0120	-352,0240	-312,1196	0,8524	0,8503
Beta	-179,2611	-348,5222	-308,6179	0,8527	0,8506
Simplex	-177,8389	-345,6777	-305,7734	0,8506	0,8485

Tabela 116: Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Probit.

Intervalo	Parâmetro	Limite Inferior	Limite Superior
Jackknife- <i>t</i>	β_0	-1,8117	-1,3451
	β_1	0,8857	1,8502
	β_2	0,0021	0,0044
	β_3	0,0050	0,0085
	β_4	-0,2849	0,4612
	β_5	$< 10^{-4}$	0,0001
	β_6	-0,0047	0,0026
	<i>b</i>	4,6486	5,8609
Pivotal	β_0	-1,8431	-1,2967
	β_1	0,8080	1,8965
	β_2	0,0014	0,0043
	β_3	0,0047	0,0088
	β_4	-0,2669	0,5777
	β_5	$< 10^{-4}$	0,0002
	β_6	-0,0066	0,0018
	<i>b</i>	4,2978	5,7715
Percentil	β_0	-1,8600	-1,3136
	β_1	0,8395	1,9279
	β_2	0,0023	0,0051
	β_3	0,0046	0,0088
	β_4	-0,4014	0,4432
	β_5	-0,0002	$< 10^{-4}$
	β_6	-0,0039	0,0044
	<i>b</i>	4,7381	6,2118
BC	β_0	-1,8546	-1,3106
	β_1	0,8343	1,9091
	β_2	0,0022	0,0048
	β_3	0,0045	0,0087
	β_4	-0,3746	0,4646
	β_5	-0,0002	$< 10^{-4}$
	β_6	-0,0060	0,0029
	<i>b</i>	4,5733	5,8102
BC _a	β_0	-1,8587	-1,3119
	β_1	0,8404	1,9284
	β_2	0,0021	0,0046
	β_3	0,0044	0,0087
	β_4	-0,3489	0,4924
	β_5	-0,0001	0,0001
	β_6	-0,0097	0,0020
	<i>b</i>	4,5187	5,7916

Tabela 117: Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Probit e remoção das covariáveis x_4 , x_5 e x_6 .

Intervalo	Parâmetro	Limite Inferior	Limite Superior
Jackknife- t	β_0	-1,9237	-1,4664
	β_1	1,1208	2,0608
	β_2	0,0027	0,0041
	β_3	0,0048	0,0082
	b	4,614	5,7791
Pivotal	β_0	-1,9748	-1,4546
	β_1	1,1347	2,1457
	β_2	0,0025	0,0042
	β_3	0,0045	0,0085
	b	4,3272	5,7397
Percentil	β_0	-1,9355	-1,4153
	β_1	1,0359	2,0469
	β_2	0,0027	0,0043
	β_3	0,0045	0,0085
	b	4,6534	6,0660
BC	β_0	-1,9425	-1,4196
	β_1	1,0393	2,0610
	β_2	0,0027	0,0043
	β_3	0,0044	0,0084
	b	4,5145	5,7855
BC $_a$	β_0	-1,9496	-1,4306
	β_1	1,0577	2,0826
	β_2	0,0026	0,0043
	β_3	0,0044	0,0084
	b	4,4932	5,7713

Tabela 118: Teste tipo Wald com combinação do Algoritmo Genético e método Bootstrap para a função de ligação Complemento Log-Log.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,3969	0,3614	-6,6316	$< 10^{-4}$
	β_1	1,8896	0,4093	4,6166	$< 10^{-4}$
	β_2	0,0024	0,0010	2,4329	0,0150
	β_3	0,0076	0,0022	3,4193	0,0006
	β_4	0,4037	0,2770	1,4572	0,1450
	β_5	$< 10^{-4}$	0,0001	0,1096	0,9127
	β_6	-0,0015	0,0027	-0,5526	0,5805
	b	5,4036	1,2629	4,2788	$< 10^{-4}$
Beta	β_0	-2,4240	0,1740	-13,9289	$< 10^{-4}$
	β_1	2,0122	0,3550	5,6687	$< 10^{-4}$
	β_2	0,0029	0,0008	3,7763	0,0002
	β_3	0,0072	0,0013	5,6037	$< 10^{-4}$
	β_4	0,1949	0,2624	0,7427	0,4577
	β_5	$< 10^{-4}$	0,0001	-0,0224	0,9821
	β_6	-0,0010	0,0025	-0,4024	0,6874
	ϕ	43,2747	7,1061	6,0898	$< 10^{-4}$
Simplex	β_0	-2,4466	0,1791	-13,6623	$< 10^{-4}$
	β_1	2,0707	0,3666	5,6489	$< 10^{-4}$
	β_2	0,0021	0,0007	3,0346	0,0024
	β_3	0,0077	0,0014	5,6071	$< 10^{-4}$
	β_4	0,3687	0,2381	1,5485	0,1215
	β_5	$< 10^{-4}$	0,0001	0,1156	0,9080
	β_6	-0,0014	0,0024	-0,5863	0,5576
	σ^2	0,5537	0,0811	6,8301	$< 10^{-4}$

Tabela 119: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Bootstrap para a função de ligação Complemento Log-Log.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,3969	0,3614	-6,6316	$< 10^{-4}$
	β_1	1,8896	0,4093	4,6166	$< 10^{-4}$
	β_2	0,0024	0,0010	2,4329	0,0150
	β_3	0,0076	0,0022	3,4193	0,0006
	β_4	0,4037	0,2770	1,4572	0,1450
	β_5	$< 10^{-4}$	0,0001	0,1096	0,9127
	β_6	-0,0015	0,0027	-0,5526	0,5805
	b	5,4036	1,2629	4,2788	$< 10^{-4}$
Beta	β_0	-2,4240	0,1740	-13,9289	$< 10^{-4}$
	β_1	2,0122	0,3550	5,6687	$< 10^{-4}$
	β_2	0,0029	0,0008	3,7763	0,0002
	β_3	0,0072	0,0013	5,6037	$< 10^{-4}$
	β_4	0,1949	0,2624	0,7427	0,4577
	β_5	$< 10^{-4}$	0,0001	-0,0224	0,9821
	β_6	-0,0010	0,0025	-0,4024	0,6874
	ϕ	43,2747	7,1061	6,0898	$< 10^{-4}$
Simplex	β_0	-2,4466	0,1791	-13,6623	$< 10^{-4}$
	β_1	2,0707	0,3666	5,6489	$< 10^{-4}$
	β_2	0,0021	0,0007	3,0346	0,0024
	β_3	0,0077	0,0014	5,6071	$< 10^{-4}$
	β_4	0,3687	0,2381	1,5485	0,1215
	β_5	$< 10^{-4}$	0,0001	0,1156	0,9080
	β_6	-0,0014	0,0024	-0,5863	0,5576
	σ^2	0,5537	0,0811	6,8301	$< 10^{-4}$

Tabela 120: Teste tipo Wald com combinação do Algoritmo Genético e método Jackknife para a função de ligação Complemento Log-Log.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,3969	0,1797	-13,3376	$< 10^{-4}$
	β_1	1,8896	0,3615	5,2274	$< 10^{-4}$
	β_2	0,0024	0,0008	2,9363	0,0033
	β_3	0,0076	0,0015	5,2293	$< 10^{-4}$
	β_4	0,4037	0,2809	1,4369	0,1507
	β_5	$< 10^{-4}$	$< 10^{-4}$	0,3519	0,7249
	β_6	-0,0015	0,0034	-0,4300	0,6672
	b	5,4036	0,3685	14,6634	$< 10^{-4}$
Beta	β_0	-2,424	0,1746	-13,8807	$< 10^{-4}$
	β_1	2,0122	0,3553	5,6626	$< 10^{-4}$
	β_2	0,0029	0,0008	3,5966	0,0003
	β_3	0,0072	0,0013	5,4377	$< 10^{-4}$
	β_4	0,1949	0,2747	0,7095	0,4780
	β_5	$< 10^{-4}$	$< 10^{-4}$	-0,0662	0,9472
	β_6	-0,001	0,0029	-0,3495	0,7267
	ϕ	43,2747	7,2642	5,9572	$< 10^{-4}$
Simplex	β_0	-2,4466	0,1790	-13,666	$< 10^{-4}$
	β_1	2,0707	0,3640	5,6881	$< 10^{-4}$
	β_2	0,0021	0,0010	2,1519	0,0314
	β_3	0,0077	0,0021	3,6921	0,0002
	β_4	0,3687	0,2626	1,4038	0,1604
	β_5	$< 10^{-4}$	$< 10^{-4}$	0,3071	0,7587
	β_6	-0,0014	0,0026	-0,5336	0,5936
	σ^2	0,5537	0,1176	4,7069	$< 10^{-4}$

Tabela 121: Teste tipo Wald com combinação do algoritmo Evolução Diferencial e método Jackknife para a função de ligação Complemento Log-Log.

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,3969	0,1797	-13,3376	$< 10^{-4}$
	β_1	1,8896	0,3615	5,2274	$< 10^{-4}$
	β_2	0,0024	0,0008	2,9363	0,0033
	β_3	0,0076	0,0015	5,2293	$< 10^{-4}$
	β_4	0,4037	0,2809	1,4369	0,1507
	β_5	$< 10^{-4}$	$< 10^{-4}$	0,3519	0,7249
	β_6	-0,0015	0,0034	-0,4300	0,6672
	b	5,4036	0,3685	14,6634	$< 10^{-4}$
Beta	β_0	-2,424	0,1746	-13,8807	$< 10^{-4}$
	β_1	2,0122	0,3553	5,6626	$< 10^{-4}$
	β_2	0,0029	0,0008	3,5966	0,0003
	β_3	0,0072	0,0013	5,4377	$< 10^{-4}$
	β_4	0,1949	0,2747	0,7095	0,4780
	β_5	$< 10^{-4}$	$< 10^{-4}$	-0,0662	0,9472
	β_6	-0,001	0,0029	-0,3495	0,7267
	ϕ	43,2747	7,2642	5,9572	$< 10^{-4}$
Simplex	β_0	-2,4466	0,1790	-13,666	$< 10^{-4}$
	β_1	2,0707	0,3640	5,6881	$< 10^{-4}$
	β_2	0,0021	0,0010	2,1519	0,0314
	β_3	0,0077	0,0021	3,6921	0,0002
	β_4	0,3687	0,2626	1,4038	0,1604
	β_5	$< 10^{-4}$	$< 10^{-4}$	0,3071	0,7587
	β_6	-0,0014	0,0026	-0,5336	0,5936
	σ^2	0,5537	0,1176	4,7069	$< 10^{-4}$

Tabela 122: Teste da Razão de Verossimilhanças Generalizada para a função de ligação Complemento Log-Log.

Modelo de Regressão	Hipótese	$l(\theta)$	Graus de Liberdade	Estatística	Valor-P
L-logística	H_0	184,6608	4	4,2187	0,3772
	H_1	186,7702			
Beta	H_0	183,8706	4	1,1775	0,8818
	H_1	184,4593			
Simplex	H_0	181,0144	4	3,7453	0,0530
	H_1	182,8870			

Tabela 123: Teste tipo Wald com combinação do Algoritmo Genético com método Bootstrap para a função de ligação Complemento Log-Log e remoção das covariáveis x_4 , x_5 e x_6 .

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,5639	0,1556	-16,4748	$< 10^{-4}$
	β_1	2,2245	0,3104	7,1657	$< 10^{-4}$
	β_2	0,0032	0,0005	6,2483	$< 10^{-4}$
	β_3	0,0078	0,0013	6,2352	$< 10^{-4}$
	b	5,3122	0,3506	15,1531	$< 10^{-4}$
Beta	β_0	-2,4961	0,1394	-17,9056	$< 10^{-4}$
	β_1	2,1433	0,2818	7,6068	$< 10^{-4}$
	β_2	0,0033	0,0005	7,3979	$< 10^{-4}$
	β_3	0,0074	0,0011	6,5873	$< 10^{-4}$
	ϕ	42,9422	6,6706	6,4375	$< 10^{-4}$
Simplex	β_0	-2,6099	0,1461	-17,8693	$< 10^{-4}$
	β_1	2,3989	0,2864	8,3765	$< 10^{-4}$
	β_2	0,0028	0,0004	6,4995	$< 10^{-4}$
	β_3	0,0078	0,0013	6,2189	$< 10^{-4}$
	σ^2	0,5679	0,0807	7,0340	$< 10^{-4}$

Tabela 124: Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Bootstrap para a função de ligação Complemento Log-Log e remoção das covariáveis x_4 , x_5 e x_6 .

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,5639	0,1556	-16,4748	$< 10^{-4}$
	β_1	2,2245	0,3104	7,1657	$< 10^{-4}$
	β_2	0,0032	0,0005	6,2483	$< 10^{-4}$
	β_3	0,0078	0,0013	6,2352	$< 10^{-4}$
	b	5,3122	0,3506	15,1531	$< 10^{-4}$
Beta	β_0	-2,4961	0,1394	-17,9056	$< 10^{-4}$
	β_1	2,1433	0,2818	7,6068	$< 10^{-4}$
	β_2	0,0033	0,0005	7,3979	$< 10^{-4}$
	β_3	0,0074	0,0011	6,5873	$< 10^{-4}$
	ϕ	42,9422	6,6706	6,4375	$< 10^{-4}$
Simplex	β_0	-2,6099	0,1461	-17,8693	$< 10^{-4}$
	β_1	2,3989	0,2864	8,3765	$< 10^{-4}$
	β_2	0,0028	0,0004	6,4995	$< 10^{-4}$
	β_3	0,0078	0,0013	6,2189	$< 10^{-4}$
	σ^2	0,5679	0,0807	7,0340	$< 10^{-4}$

Tabela 125: Teste tipo Wald com combinação do Algoritmo Genético com método Jackknife para a função de ligação Complemento Log-Log e remoção das covariáveis x_4 , x_5 e x_6 .

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,5639	0,156	-16,4361	$< 10^{-4}$
	β_1	2,2245	0,3105	7,165	$< 10^{-4}$
	β_2	0,0032	0,0005	6,0999	$< 10^{-4}$
	β_3	0,0078	0,0012	6,3314	$< 10^{-4}$
	b	5,3122	0,3512	15,1242	$< 10^{-4}$
Beta	β_0	-2,4961	0,1393	-17,9246	$< 10^{-4}$
	β_1	2,1433	0,2815	7,6127	$< 10^{-4}$
	β_2	0,0033	0,0005	7,3395	$< 10^{-4}$
	β_3	0,0074	0,0011	6,4989	$< 10^{-4}$
	ϕ	42,9422	6,9739	6,1575	$< 10^{-4}$
Simplex	β_0	-2,6099	0,1469	-17,7659	$< 10^{-4}$
	β_1	2,3989	0,2895	8,2871	$< 10^{-4}$
	β_2	0,0028	0,0004	6,2945	$< 10^{-4}$
	β_3	0,0078	0,0013	6,0951	$< 10^{-4}$
	σ^2	0,5679	0,0856	6,6361	$< 10^{-4}$

Tabela 126: Teste tipo Wald com combinação do algoritmo Evolução Diferencial com método Jackknife para a função de ligação Complemento Log-Log e remoção das covariáveis x_4 , x_5 e x_6 .

Modelo de Regressão	Parâmetro	Estimativa	Erro Padrão	Estatística	Valor-P
L-logística	β_0	-2,5639	0,156	-16,4361	$< 10^{-4}$
	β_1	2,2245	0,3105	7,165	$< 10^{-4}$
	β_2	0,0032	0,0005	6,0999	$< 10^{-4}$
	β_3	0,0078	0,0012	6,3314	$< 10^{-4}$
	b	5,3122	0,3512	15,1242	$< 10^{-4}$
Beta	β_0	-2,4961	0,1393	-17,9246	$< 10^{-4}$
	β_1	2,1433	0,2815	7,6127	$< 10^{-4}$
	β_2	0,0033	0,0005	7,3395	$< 10^{-4}$
	β_3	0,0074	0,0011	6,4989	$< 10^{-4}$
	ϕ	42,9422	6,9739	6,1575	$< 10^{-4}$
Simplex	β_0	-2,6099	0,1469	-17,7659	$< 10^{-4}$
	β_1	2,3989	0,2895	8,2871	$< 10^{-4}$
	β_2	0,0028	0,0004	6,2945	$< 10^{-4}$
	β_3	0,0078	0,0013	6,0951	$< 10^{-4}$
	σ^2	0,5679	0,0856	6,6361	$< 10^{-4}$

Tabela 127: Medidas de qualidade do ajuste para a função de ligação Complemento Log-Log.

Modelo de Regressão	$-l(\theta)$	AIC	BIC	R^2	R_a^2
L-logística	-184,6608	-359,3216	-319,4173	0,8615	0,8595
Beta	-183,8706	-357,7411	-317,8368	0,8618	0,8599
Simplex	-181,0144	-352,0287	-312,1244	0,8593	0,8573

Tabela 128: Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Complemento Log-Log.

Intervalo	Parâmetro	Limite Inferior	Limite Superior
Jackknife- <i>t</i>	β_0	-2,6944	-2,0994
	β_1	1,2912	2,4879
	β_2	0,0010	0,0038
	β_3	0,0052	0,0101
	β_4	-0,0614	0,8687
	β_5	$< 10^{-4}$	$< 10^{-4}$
	β_6	-0,0072	0,0042
	b	4,7936	6,0136
Pivotal	β_0	-3,7938	-2,0418
	β_1	1,2157	2,7791
	β_2	0,0002	0,0048
	β_3	0,0046	0,0153
	β_4	-0,0415	1,0961
	β_5	-0,0001	0,0002
	β_6	-0,0081	0,0019
	b	-0,1928	5,8579
Percentil	β_0	-2,7520	-1,0000
	β_1	1,0000	2,5634
	β_2	$< 10^{-4}$	0,0046
	β_3	$< 10^{-4}$	0,0107
	β_4	-0,2888	0,8489
	β_5	-0,0002	0,0001
	β_6	-0,0049	0,0051
	b	4,9494	11,0000
BC	β_0	-2,7450	-1,0000
	β_1	1,0000	2,5706
	β_2	$< 10^{-4}$	0,0045
	β_3	$< 10^{-4}$	0,0107
	β_4	-0,1312	0,8897
	β_5	-0,0002	0,0001
	β_6	-0,0087	0,0038
	b	4,6673	5,9593
BC _a	β_0	-2,7520	-1,0000
	β_1	1,0000	2,5840
	β_2	$< 10^{-4}$	0,0043
	β_3	$< 10^{-4}$	0,0107
	β_4	-0,0954	0,9131
	β_5	-0,0001	0,0001
	β_6	-0,0140	0,0026
	b	4,5358	5,9331

Tabela 129: Intervalos de Confiança com métodos de reamostragem para o modelo de regressão L-logística com função de ligação Complemento Log-Log e remoção das covariáveis x_4 , x_5 e x_6 .

Intervalo	Parâmetro	Limite Inferior	Limite Superior
Jackknife- t	β_0	-2,8221	-2,3057
	β_1	1,7106	2,7384
	β_2	0,0024	0,0041
	β_3	0,0058	0,0099
	b	4,7308	5,8936
Pivotal	β_0	-2,8878	-2,2809
	β_1	1,7037	2,8829
	β_2	0,0021	0,0042
	β_3	0,0052	0,0101
	b	4,4491	5,8190
Percentil	β_0	-2,8470	-2,2400
	β_1	1,5661	2,7453
	β_2	0,0023	0,0044
	β_3	0,0055	0,0105
	b	4,8054	6,1753
BC	β_0	-2,8610	-2,2534
	β_1	1,5998	2,7944
	β_2	0,0023	0,0043
	β_3	0,0052	0,0102
	b	4,6252	5,8747
BC $_a$	β_0	-2,8685	-2,2632
	β_1	1,6113	2,8014
	β_2	0,0022	0,0042
	β_3	0,0052	0,0101
	b	4,5644	5,8545

Tabela 130: Estatísticas da distância do tipo Cook para todos os modelos e todas as funções de ligação.

Função de Ligação	Estatística	Modelo de Regressão		
		L-logística	Beta	Simplex
Logit	Mínimo	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
	Máximo	0,8184	0,6039	0,7505
	Média	0,0275	0,0285	0,0227
	Desvio Padrão	0,0780	0,0777	0,0796
	Mediana	0,0053	0,0035	0,0019
	IQR	0,0243	0,0244	0,0100
Probit	Mínimo	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
	Máximo	0,8154	0,6149	0,7327
	Média	0,0272	0,0283	0,0230
	Desvio Padrão	0,0781	0,0777	0,0795
	Mediana	0,0050	0,0035	0,0021
	IQR	0,0232	0,0240	0,0092
Complemento Log-log	Mínimo	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
	Máximo	0,5637	0,5056	0,4922
	Média	0,0305	0,0341	0,0266
	Desvio Padrão	0,0769	0,0754	0,0784
	Mediana	0,0046	0,0045	0,0025
	IQR	0,0244	0,0251	0,0151

Figura 21: Gráfico da distância do tipo Cook para os modelo de regressão com função de ligação Logit.

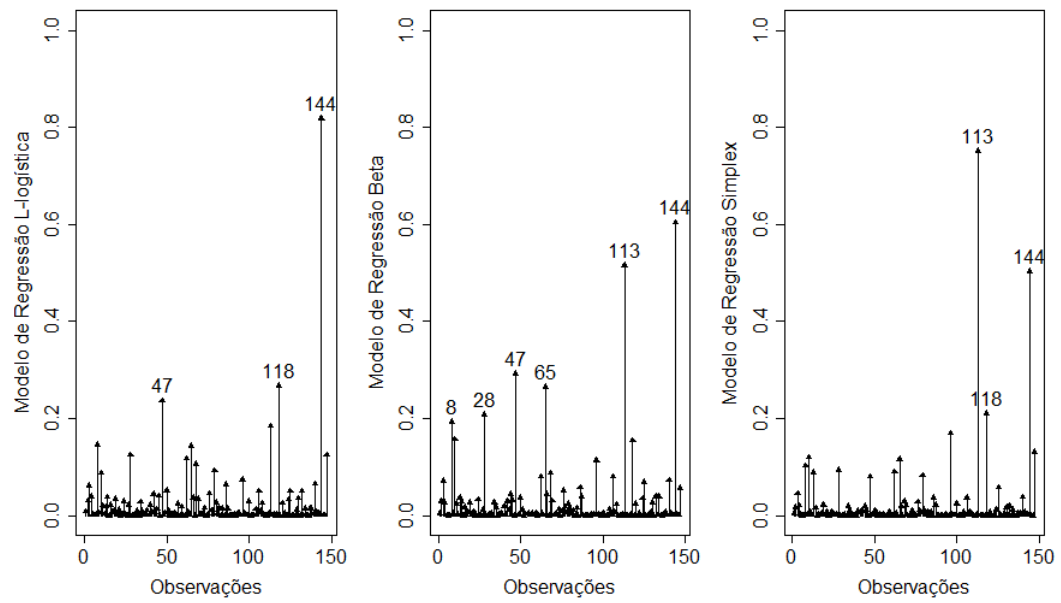


Figura 22: Gráfico da distância do tipo Cook para o modelo de regressão com função de ligação Probit.

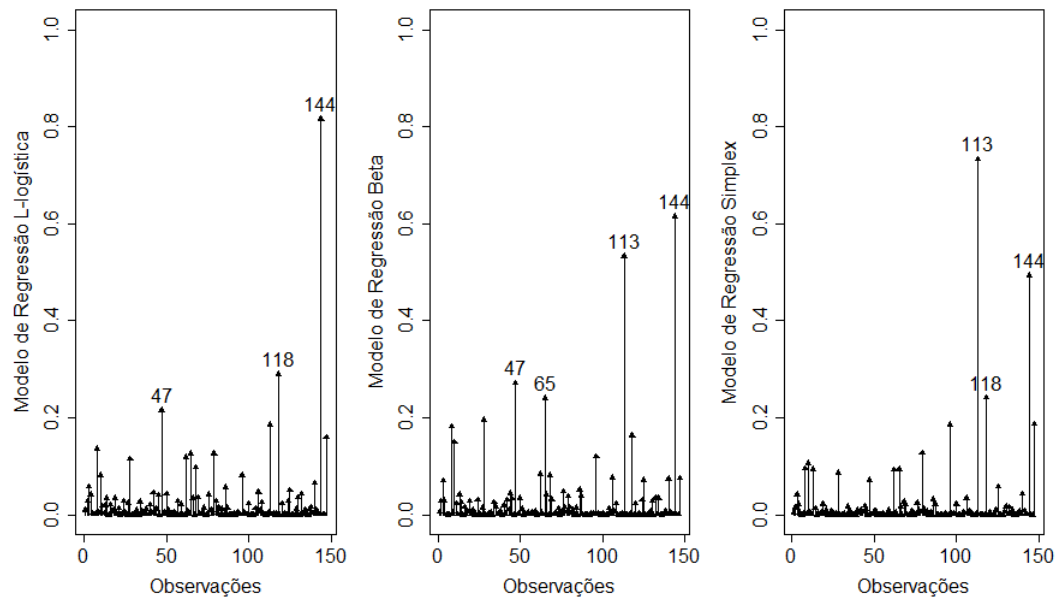
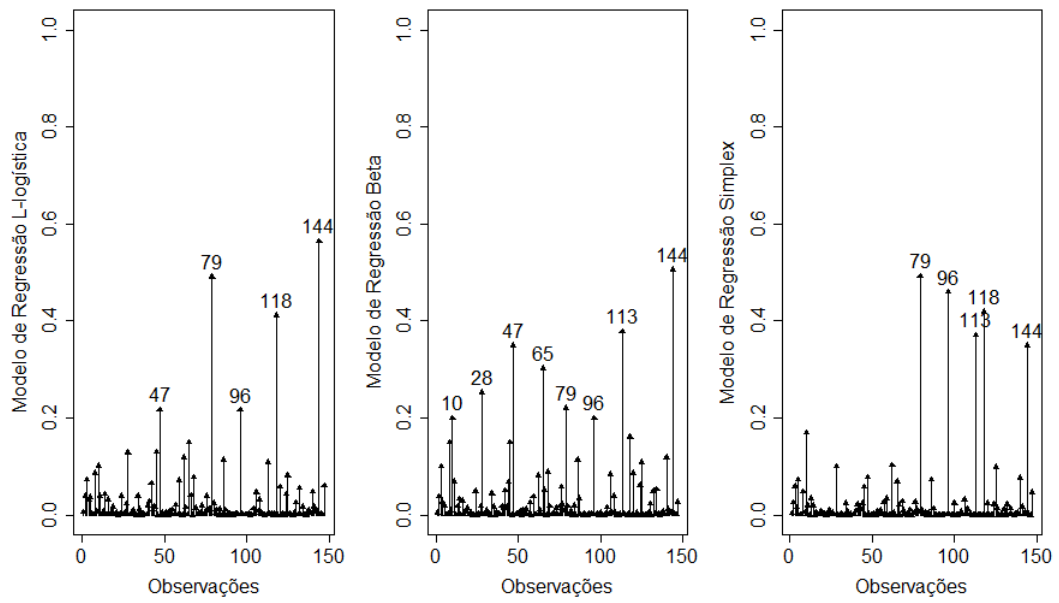


Figura 23: Gráfico da distância do tipo Cook para o modelo de regressão com função de ligação Complemento Log-log.



6 CONCLUSÕES

6.1 Considerações Finais

Ao longo desta dissertação estudamos a distribuição L-logística proposta em Paz et al. (2016) e seu respectivo modelo de regressão como opção aos tradicionais modelo de regressão Beta e Simplex. Este novo modelo possui aplicabilidade real na modelagem de variáveis com o suporte $[0, 1]$ (ou suporte limitado).

Este trabalho apresentou o enfoque frequentista, por meio do método de estimação por máxima verossimilhança. As estimações pontuais foram proporcionadas pelas metaheurísticas Algoritmo Genético e Evolução Diferencial, enquanto que as estimativas do erro-padrão foram obtidas pelos métodos de reamostragem Bootstrap e Jackknife. A eficiência da combinação das técnicas foi constatada por meio de um estudo de simulação, no qual verificou-se que as estimativas pontuais estavam próximas dos valores reais dos parâmetros e que as estimativas do erro-padrão forneciam consistentes intervalos de confiança com base na probabilidade de cobertura.

A aplicabilidade do modelo de regressão L-logística foi evidenciada por meio da modelagem de dois conjuntos de dados reais e seus resultados foram comparados com os tradicionais modelo de regressão Beta e Simplex. Em virtude da recente proposta de Paz et al. (2016), utilizou-se somente como diagnóstico uma medida de distância do tipo Cook para detectar observações atípicas.

Finalmente, ao leitor interessado na modelagem de variáveis com o suporte $[0, 1]$ (ou suporte limitado), sugere-se que considere o modelo de regressão L-logística como opção de modelagem e utilize a proposta da combinação de técnicas.

6.2 Estudos futuros

Como melhoria para esta pesquisa, propõe-se os seguintes pontos:

- Estudo do desempenho do modelo de regressão L-logística em outros conjuntos de dados.
- Estudo de outros algoritmos de otimização e reamostragem.
- Estudo comparativo para testar a performance do modelo de regressão L-logística contra a performance de modelos livres de distribuição (não-paramétricos), regularmente visto na área de aprendizado estatístico, tais como Redes Neurais Artificiais (*Artificial Neural Networks*), Árvores de Decisão (Decision Trees) e Máquinas de Vetor de Suporte (Support Vector Machines).
- Estudo de métodos para melhorar a performance computacional da pesquisa (redução de custo computacional).
- Estudo de novos métodos de diagnóstico (análise de resíduos, análise de influência local, indicadores de adequação da função de ligação etc).

REFERÊNCIAS

- [1] Arnold, B. C., Groeneveld, R. A. (1995). Measuring Skewness with Respect to the Mode. *The American Statistician*. **49**(1), 34-38.
- [2] Akman, F., Akman, O., Hallam, J. W. (2011). Maximum Likelihood Estimation Using Accelerated Genetic Algorithms. *Journal of Applied Statistical Science*. **18**(1), 1-7.
- [3] Ari ,Ç., Aksoy, S., Ari kan, O. (2012). Maximum Likelihood estimation of Gaussian mixture models using stochastic search. *Pattern Recognition*. **45**(1). 2804-2816.
- [4] Bäck, T. (1996). Evolutionary Algorithms in Theory and Praticce. **Oxford University Press, Inc.**
- [5] Blattberg, R., Sargent, T. (1971). Regression With Non-Gaussian Stable Disturbances: Some Sampling Results. *Econometrica*. **39**(3). 501-510.
- [6] Bassett, G. Jr., Koenker, R. (1978). Asymptotic Theory of Least Absolute Error Regression. *Journal of the American Statistical Association*. **73**(363). 618-622.
- [7] Basu, A., Ghosh, A., Mandal, A., Martin, N., Pardo, L. 2017. A Wald-type test statistic for testing linear hypothesis in logistic regression models based on minimum density power divergence estimator. *Electronic Journal of Statistics*. **11**(2). 2741-2772.
- [8] Belsley, D.A., Kuh, E. and Welsch, Jt.E. (1980) Regression Diagnostics. J. Wiley e Sons, New York.
- [9] Bolfarini, H., Sandoval, M. C. (2010). Introdução à Inferência Estatística, 2ª edição. **SBM**.
- [10] Brindle, A. (1981). Genetic algorithms for function optimization (Doctoral dissertation and Technical Report TR81-2). **Edmonton: University of Alberta, Department of Computer Science.**
- [11] Brys, G., Hubert, M., Struyf, A. (2003). A comparison of some new measures of skewness. In: Dutter, R., Filzmoser, P., Gather, U., Rousseeuw, P.J. (Eds.), *Developments in Robust Statistics, ICORS 2001*. Springer, Heidelberg, 98–113.
- [12] Cordeiro, G. M., Lemonte, A. J. (2012). The McDonald inverted beta distribution. *Journal of the Franklin Institute*. **349**(1), 1174-1197.
- [13] Coley, D. A. (1999). An Introduction to Genetic Algorithms for Scientists and Engineers. **Wold Scientific Publishing Co. Pte. Ltd.**
- [14] Darwin, C. (1859). On the origin of species by means of natural selection, or, the preservation of favoured races in the struggle for life. **London: J. Murray.**
- [15] De Jong, K. A. (1975). An analysis of the behavior of a class of genetic adaptive systems (Doctoral dissertation, University of Michigan). *Dissertation Abstracts International*. **36**(10), 5140B. (University Microfilms No. 76-9381)
- [16] Dufour, J.-M. (1997). Some Impossibility Theormes in Econometrics with Applications to Structural and Dynamic Models. *Econometrica*. **61**(1), 1365-1388.
- [17] Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*. **7**(1), 1-26.

- [18] Efron, B. (1981). Nonparametric standard errors and confidence intervals. *The Canadian Journal of Statistics*. **9**(2), 139-172.
- [19] Efron, B. (1982). The Jackknife, the Bootstrap and Others Resampling Plans. **Society for Industrial and Applied Mathematics**.
- [20] Efron, B. (1987). Better Bootstrap Confidence Intervals. *Journal of the American Statistical Association*. **82**(1), 171-185.
- [21] Efron, B., Tibshirani, R. J. (1993). An Introduction to the Bootstrap. **Springer Science + Business Media Dordrecht**.
- [22] Elsheman, L. J., Schaffer, J. D. (1993). Real-coded genetic algorithms and intervalschema. *Foundation of Genetic Algorithms 2, Morgan Kaufmann Publishers*, 187-202.
- [23] Engelbrecht, A. P. (2007). Computational Intelligence: An introduction. **John Wiley Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England**.
- [24] Eshelman, L. J., Schaffer, J. D. (1993). Real-Coded Genetic Algorithms and Interval-Schemata. *Foundations of Genetic Algorithms*. **2**(1), 187-202.
- [25] Fay, M. P., Graubard, B. I. (2001). Small-Sample Adjustments for Wald-Type Tests Using Sandwich Estimators. *Biometrics*. **57**(4), 1198-1206.
- [26] Ferrari, S. L. P., Cribari-Neto, F. (2004). Beta Regression for Modelling Rates and Proportions. *Journal of Applied Statistics*. **31**(7), 799-815.
- [27] Feoktistov, V. (2006). Differential Evolution: In Search of Solution. **Springer Science+Business Media, LLC**.
- [28] Gen, M., Cheng, R. (2000). Genetic Algorithms and Engineering Optimization. **John Wiley & Sons, Inc**
- [29] Goldberg, D. E. (1989). Genetic Algorithms in Search, Optimization, and Machine Learning. **Addison-Wesley Publishing Company, Inc**.
- [30] Goldberg, D. E. (1991). Real-coded Genetic Algorithms, Virtual Alphabets, and Blocks. *Complex Systems*. **5**(1), 139-167.
- [31] Groeneveld, R. A., Meeden, G. (1984). Measuring Skewness and Kurtosis. *Journal of the Royal Statistical Society, Series D*. **33**(4), 391-399.
- [32] Hahn, E. D. (2008). Mixture densities for project management activity times: A robust approach to PERT. *European Journal of Operational Research*. **188**(1), 450-459.
- [33] Haupt, R. L., Haupt, S. E. (2004). Practical Genetic Algorithms. **John Wiley & Sons, Inc**.
- [34] Hinkley, D. V. (1975). On Power Transformations to Symmetry. *Biometrika*. **62**(1), 101-111.
- [35] Hinterding, R. (1995). Gaussian Mutation and Self-Adaptation for Numeric Genetic Algorithms. *IEEE International Conference on Evolutionary Computation*. **1**(1), 384-389.
- [36] Hoaglin, D.C. and Welsch, R.E. (1978). The hat matrix in regression and ANOVA. *Am. Statistician* **32**, 17-22 (Corr. **32**, 146).

- [37] Holland, J. (1975). *Adaptation in Natural and Artificial Systems*. **Ann Arbor: University of Michigan Press**.
- [38] Janikow, C. Z., Michalewicz, Z. (1991). An Experimental Comparison of Binary and Floating Point Representations in Genetic Algorithms. **Departament of Mathematics and Computer Science, University of Missouri, St. Louis, Missouri**.
- [39] Johnson, W. D., Romer, J. E. (2016). Hypothesis Testing of Population Percentiles via the Wald Test with Bootstrap Variance Estimates. *Open journal of statistics*. **6**(1), 14-24.
- [40] Krzanowski, W. J. (2000). *Principles of multivariate analysis: a user's perspective*. 2nd. ed. **New York: Oxford University Press**.
- [41] Katharine Mullen, David Ardia, David Gil, Donald Windover, James Cline (2011). 'DEoptim': An R Package for Global Optimization by Differential Evolution. *Journal of Statistical Software*, **40**(6), 1-26. URL <http://www.jstatsoft.org/v40/i06/>.
- [42] Luca Scrucca (2013). GA: A Package for Genetic Algorithms in R. *Journal of Statistical Software*, **53**(4), 1-37. URL <http://www.jstatsoft.org/v53/i04/>
- [43] Man, K. F., Tang K. S., Kwong S. (1996). Genetic Algorithms: Concepts and Applications. *IEEE Transactions on Industrial Electronics*. **43**(5), 519-534
- [44] McDonald, J. B. (1984). Some generalized functions for the size distribution of income. *Econometrica*. **52**(1), 647-664.
- [45] Melanie, M. (1996). *An Introduction to Genetic Algorithms*. **Massachusetts Institute of Technology**.
- [46] Miller, R. G. (1974). The Jackknife - A Review. *Biometrika*. **61**(1), 1-15.
- [47] Michalewicz, Z. (1996). *Genetic Algorithms + Data Structures = Evolution Programs*. **Springer-Verlag Berlin Heidelberg**.
- [48] Mitnik, P. A., Baek, S. (2013). The Kumaraswamy distribution: median-dispersion re-parameterizations for regression modeling and simulation-based estimation. *Statistical Papers*. **54**(1), 177-192.
- [49] Miyashiro, E. S. (2008). *Modelos de regressão Beta e Simplex para análise de proporções*. 76f. Dissertação (Mestrado), Universidade de São Paulo, São Paulo, 2008.
- [50] Moors, J. J. A. (1988). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society, Series B*. **37**(1), 25-32.
- [51] Mülenbein, H. (1992). How genetic algorithms really work I. Mutation and Hillclimbing. *Proceedings of the Parallel Problem Solving from Nature*. **2**(1), 15-26.
- [52] Mullen, K.M, Ardia, D., Gil, D., Windover, D., Cline, J. (2011). DEoptim: An R Package for Global Optimization by Differential Evolution. *Journal of Statistical Software*. **40**(6), 1-26.
- [53] Narula, S. C., Wellington, J. F. (1982). The Minimum Sum of Absolute Errors Regression: A State of the Art Survey. *InternationaI Statistical Review*. **50**(3), 317-326.
- [54] Nelder, A., Wedderburn, R. W. M. (1972) Generalized Linear Models. *Journal of the Royal*

Statistical Society. Series A (General). **135**(3), 370-384.

- [55] McCullagh, P., Nelder, J. A. (1989). *Generalized Linear Models*. **London: Chapman and Hall**.
- [56] Nocedal, J., Wright, S. J. (2006). *Numerical Optimization* 2nd edition. Springer.
- [57] Paz, R. F. da, Balakrishnan, N., Bazán, J. L. (2016). L-logistic Distribution: Properties, Inference and an Application to study poverty and inequality in Brazil. *Teoria e Métodos – Série A*. Abril/2016, 262. Programa Interinstitucional de Pós-Graduação em Estatística, UFSCar-USP.
- [58] Paz, R. F. da, Balakrishnan, N., Bazán, J. L. (2018). L-Logistic regression models: Prior sensitivity analysis, robustness to outliers and applications. *Brazilian Journal of Probability and Statistics*.
- [59] Price, K. (1996). Differential Evolution: A Fast and Simple Numerical Optimizer. *NAFIPS'96*. **1**(1), 524-527.
- [60] Quenouille, M. H. (1949). Approximate tests of correlation in time series. *Journal of the Royal Statistical Society, Series B*. **11**(1), 18-44.
- [61] Quenouille, M.H. (1956). Notes on Bias in Estimation. *Biometrika*. **43**(3/4), 353-360.
- [62] Schaffer, J. D., Caruana, R., Eshelman, L. J., Das, R. (1989). A Study of Control Parameters Affecting Online Performance of Genetic Algorithms for Function Optimization. Third international Conference on Genetic Algorithms, January, 1989. 51-60.
- [63] Scrucca L. (2013). GA: A Package for Genetic Algorithms in R. *Journal of Statistical Software*. **53**(4), 1-37.
- [64] Shao, J., WU, C. F. J. (1989). A General Theory for Jackknife Variance Estimation. *The Annals of Statistics*. **17**(3), 1176-1197.
- [65] Silva, N. A. (2008). *O uso do Algoritmo Genético na Construção de Mapas de Perfusão Cerebral e sua Aplicação em pacientes com Anemia Falciforme*. 123f. Tese (Doutorado), Universidade de São Paulo, São Paulo, 2008.
- [66] Silvapulle, M. J., Silvapulle, P. (1995). A Score Test against One-Sided Alternatives. *Journal of the American Statistical Association*. **90**(429), 342-349.
- [67] Spears, W. M. (1998). *The Role of Mutation and Recombination in Evolutionary Algorithms*. PhD thesis, George Mason University, Fairfax, Virginia.
- [68] Song, P. X. K., Tan, M. (2000). Marginal Models for Longitudinal Continuous Proportional Data. *Biometrics*. **56**(1), 496-502.
- [69] Srinivas, M., Patnaik, M. L. (1994). Genetic Algorithms: A survey. *Computer*. **27**(6), 17-16.
- [70] Storn, R. (1996). Differential Evolution Design of an IIR-Filter. *In Proceedings of the IEEE International Conference on Evolutionary Computation*. **1**(1), 268-273.
- [71] Storn, R. and Price, K. (1995) Differential evolution — A simple and efficient adaptive scheme for global optimization over continuous spaces. **International Computer Science Institute, Berkley**.
- [72] Storn R., Price, K. V. (1996). Minimizing the real functions of the ICEC'96 contest by differential

- evolution. *IEEE international conference on evolutionary computation*. **1**(1), 842–844.
- [73] Storn, R., Price, K. (1997). Differential Evolution – A simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global Optimization*. **11**(1), 341-359.
- [74] Tadikamalla, P. R., Johnson, N. L. (1982). Systems of Frequency Curves Generated by Transformations of Logistic Variables. *Biometrika*. **69**(2), 461-465.
- [75] Tibshirani, R. (1984). Bootstrap Confidence Intervals. **Technical Reports No. 3. Department of Statistics, Stanford University, California, EUA.**
- [76] Tukey, J. W. (1958). Bias and confidence in not quite large samples. (*Abstract*) *The Annals of Mathematical and Statistics*. **29**(1), 614.
- [77] van der Vaart, W. A., Wellner, J. A. (1996) Weak Convergence and Empirical Processes. Springer-Verlag New York, Inc.
- [78] Wang, M., Rennolls, K. (2005). Tree diameter distribution modelling: introducing the logit-logistic distribution. *Canadian Journal of Forest Research*. **35**(6), 1305-1313.
- [79] Wilcox, R. R. (1995). Comparing Two Independent Groups Via Multiple Quantiles. *Journal of the Royal Statistical Society. Series D (The Statistician)*. **44**(1), 91-99.
- [80] Yang, X. (2010). Engineering Optimization: An Introduction with Metaheuristic Applications. **John Wiley & Sons, Inc., Hoboken, New Jersey.**
- [81] Yang, X. (2014). Nature-Inspired Optimization Algorithms. **Elsevier Inc.**
- [82] Zhang, Z., Lin, J., Shi, Y. (2013). Application of Artificial Bee Colony Algorithm to Maximum Likelihood DOA Estimation. *Journal of Bionic Engineering*. **10**(1), 100-109.
- [83] Zäpfel, G., Braune, R., Bögl, M. (2010). Metaheuristic Search Concepts: A Tutorial with Applications to Production and Logistics. **Springer-Verlag Berlin Heidelberg.**

APÊNDICE A – DERIVADAS PARCIAIS

$$\begin{aligned}
 \frac{\partial l(m, b; \mathbf{y})}{\partial m} = & \sum_i^n \left(\frac{1}{b[m(1-m)]^b [y_i(1-y_i)]^{b-1}} \times \right. \\
 & \left[\left(\frac{-b}{1-m} \frac{b(1-m)^b m^b y_i^{b-1} (1-y_i)^{b-1}}{[(1-m)^b y_i^b + m^b (1-y_i)^b]^2} + \right. \right. \\
 & \left. \frac{b}{m} \frac{b(1-m)^b m^b y_i^{b-1} (1-y_i)^{b-1}}{[(1-m)^b y_i^b + m^b (1-y_i)^b]^2} - \right. \\
 & \left. \frac{b(1-m)^b m^b y_i^{b-1} (1-y_i)^{b-1}}{[(1-m)^b y_i^b + m^b (1-y_i)^b]^3} \times \right. \\
 & \left. 2 \left(\frac{b [m(1-y_i)]^b}{m} - \frac{b [(1-m)y_i]^b}{1-m} \right) \right) \times \\
 & \left. [(1-m)^b y_i^b + m^b (1-y_i)^b]^2 \right] \quad (52)
 \end{aligned}$$

