

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

Modelos Skellam Generalizados

Marina Gandolfi

Tese de Doutorado do Programa Interinstitucional de Pós-Graduação em Estatística (PIPGEs)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Marina Gandolfi

Modelos Skellam Generalizados

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP e ao Departamento de Estatística – DEs-UFSCar, como parte dos requisitos para obtenção do título de Doutora em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística.
VERSÃO REVISADA

Área de Concentração: Estatística

Orientadora: Profa. Dra. Katiane Silva Conceição

Coorientador: Prof. Dr. Carlos Alberto Ribeiro Diniz

USP – São Carlos
Agosto de 2024



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa Interinstitucional de Pós-Graduação em Estatística

Folha de Aprovação

Defesa de Tese de Doutorado da candidata Marina Gandolfi, realizada em 24/06/2024.

Comissão Julgadora:

Profa. Dra. Katiane Silva Conceição (USP)

Prof. Dr. Rui Manuel da Costa Martins (ULisboa)

Profa. Dra. Lilia Carolina Carneiro da Costa (UFBA)

Profa. Dra. Sandra Cristina de Oliveira (UNESP)

Profa. Dra. Vera Lucia Damasceno Tomazella (UFSCar)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa Interinstitucional de Pós-Graduação em Estatística.

Marina Gandolfi

Generalized Skellam Models

Thesis submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP and to the Department of Statistics – DEs-UFSCar – in accordance with the requirements of the Statistics Interagency Graduate Program, for the degree of Doctor in Statistics. *FINAL VERSION*

Concentration Area: Statistics

Advisor: Profa. Dra. Katiane Silva Conceição

Co-advisor: Prof. Dr. Carlos Alberto Ribeiro Diniz

USP – São Carlos

August 2024

AGRADECIMENTOS

Agradeço à minha família por todo apoio. Ao meu namorado Leonardo, pelo amor e companheirismo. À minha orientadora, Profa. Dra. Katiane Silva Conceição, pela compreensão, dedicação e disponibilidade. Aos meus amigos e colegas, pelos momentos compartilhados. À todos os professores e funcionários do programa, pelo compromisso com o ensino e a pesquisa.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Pesquisa desenvolvida com o auxílio dos recursos de HPC disponibilizados pela Superintendência de Tecnologia da Informação da Universidade de São Paulo.

RESUMO

GANDOLFI, M. **Modelos Skellam Generalizados**. 2024. 105 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2024.

Na estatística aplicada, frequentemente dados de contagem são observados em diferentes áreas de estudo. Devido a grande diversidade dos problemas que resultam nestes tipos dados, torna-se necessário a proposta de novos modelos. Neste trabalho, propomos generalizações da distribuição Skellam, cujo suporte consiste do conjunto formado pelos números inteiros (positivos e negativos), visando explorar também no contexto de modelos de regressão. Para o processo de estimação e inferência dos parâmetros dos modelos foram consideradas as abordagens clássica (método de máxima verossimilhança) e bayesiana (Monte Carlo em Cadeia de Markov) para fins de comparação. Especificamente sobre a abordagem bayesiana, que foi mais eficiente nas propostas aqui apresentadas, utilizamos uma variante do algoritmo Hamiltoniano Monte Carlo, que consiste na reformulação das equações de Hamilton ao introduzir uma componente estocástica na equação do gradiente, derivando o algoritmo Gradiente Estocástico Hamiltoniano Monte Carlo. Para ilustrações dos modelos propostos, apresentamos as análises de conjuntos de dados referentes a dois problemas reais (total de 3 conjuntos de dados): no primeiro problema foi considerado um conjunto de dados correspondente às observações da variação semanal da pontuação do Ibovespa, isto é, a diferença de preço, medida em *ticks* (centavos) do dia atual com relação ao dia anterior, no período entre janeiro de 2000 e dezembro de 2022. Os valores estimados para o parâmetro p , caracterizaram o conjunto de dados como inflacionado de observações -2 (*ticks*); no segundo problema, foram considerados dois conjuntos de dados, correspondentes aos valores das diferenças entre partidas ganhas e partidas perdidas pelos times na temporada regular de 2022-2023 da *National Basketball Association*, em cada conferência (Leste e Oeste). Os critérios de seleção indicaram para o modelo k -MS com $k = -12$ como o mais bem ajustado para a conferência Leste, enquanto para a conferência Oeste, o valor indicado foi $k = -38$. Diante dos bons resultados, tanto os modelos k -MS quanto os k -IS demonstraram ser boas alternativas para explicar o comportamento de dados com valores inteiros.

Palavras-chave: Distribuição k -Modificada; Inflação; Deflação; Abordagem Bayesiana; Gradiente Estocástico Hamiltoniano Monte Carlo.

ABSTRACT

GANDOLFI, M. **Generalized Skellam Models**. 2024. 105 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2024.

In applied statistics, counting data are often observed in different areas of study. Due to the great diversity of problems that result in these types of data, it is necessary to propose new models. In this work, we propose generalizations of the Skellam distribution, whose support consists of the set formed by integers (positive and negative), aiming to also explore in the context of regression models. For the process of estimating and inferring model parameters, the classical (maximum likelihood method) and Bayesian (Markov Chain Monte Carlo) approaches were considered for comparison purposes. Specifically regarding the Bayesian approach, which was more efficient in the proposals presented here, we used a variant of the Hamiltonian Monte Carlo algorithm, which consists of reformulating Hamilton's equations by introducing a stochastic component into the gradient equation, deriving the Stochastic Gradient Hamiltonian Monte Carlo algorithm. To illustrate the proposed models, we present the analyzes of data sets referring to two real problems (3 datasets in total): In the first problem, a set of data corresponding to observations of the weekly variation of the Ibovespa score was considered, that is, the price difference, measured in ticks (cents) of the current day in relation to the previous day, in the period between January 2000 and December 2022. The estimated values for the parameter p characterized the data set as inflated with observations -2 (*ticks*); In the second problem, two sets of data were considered, corresponding to the values of the differences between games won and games lost by teams in the 2022-2023 regular season of the National Basketball Association, in each conference (East and West). The selection criteria indicated the k -MS model with $k = -12$ as the best adjusted for the Eastern conference, while for the Western conference, the indicated value was $k = -38$. Given the good results, both the k -MS and the k -IS models proved to be good alternatives to explain the behavior of data with integer values.

Keywords: k -Modified Distribution; Inflation; Deflation; Bayesian approach; Stochastic Gradient Hamiltonian Monte Carlo.

LISTA DE ILUSTRAÇÕES

Figura 1	– Estimativas de p e μ_{k-MS} (juntamente com seus valores reais) obtidas a partir do ajuste do modelo k -IS, para $n = 200$	57
Figura 2	– Estimativas de p e μ_{k-MS} (juntamente com seus valores reais) obtidas a partir do ajuste do modelo k -DS, para $n = 200$	59
Figura 3	– Estimativas de p e μ_{k-MS} (juntamente com seus valores reais) obtidas a partir do ajuste do modelo k -MS, para valores de parâmetros do CD1 e tamanho de amostra $n = 100$	62
Figura 4	– Estimativas bayesianas de μ_{k-MS} (juntamente com seus valores reais) obtidas a partir do ajuste do modelo k -MS, para $n = 200$	75
Figura 5	– Estimativas bayesianas do parâmetro p obtidas a partir do ajuste do modelo k -MS, para $n = 200$	76
Figura 6	– Distribuição de frequências das quantidades de <i>ticks</i> registradas na variação semanal do Ibovespa no período de 2000-2022.	77
Figura 7	– Estimativas de p e de μ_{k-MS} obtidas a partir do modelo k -MS ajustado, com $k = -2$: (A) \hat{p} versus a covariável X ; (B) \mathbf{y} versus a covariável X , juntamente com a estimativa de μ_{k-MS}	80
Figura 8	– Ilustrações em gráficos de barras das pontuações (cinza claro) e também da covariável $X = J_G - J_P$ (cinza escuro), por conferência, na temporada regular 2022-2023 da NBA.	82
Figura 9	– Estimativas de p e de μ_{k-MS} , versus a covariável, X obtidas a partir do modelo k -MS ajustado, com: (A1) e (A2) para os dados da conferência Leste, modelo k -MS com $k = -12$; e (B1) e (B2) para os dados da conferência Oeste, modelo k -MS com $k = -38$	84
Figura 10	– Pontuação final da temporada regular da NBA 2023-2024 versus a diferença entre a quantidade de jogos ganhos com no mínimo 10 pontos de vantagem e jogos perdidos com no mínimo 10 pontos de desvantagem, para as conferências Leste e Oeste.	85
Figura 11	– Estimativas bayesianas de μ_{k-IS} (juntamente com seus valores reais) obtidas a partir do modelo k -IS ajustado, com $k_1 = -4$ e $k_2 = 1$, e tamanho amostral $n = 200$: \mathbf{y} versus a covariável X , juntamente com a estimativa de μ_{k-IS}	96
Figura 12	– Estimativas de p_1 , p_2 e de μ_{k-IS} obtidas a partir do modelo k -IS ajustado, com $k_1 = 2$ e $k_2 = 3$: (A) \hat{p}_1 versus a covariável X ; (B) \hat{p}_2 versus a covariável X ; (C) \mathbf{y} versus a covariável X , juntamente com a estimativa de μ_{k-MS}	98

LISTA DE ALGORITMOS

Algoritmo 1 – Hamiltoniano Monte Carlo (HMC)	66
Algoritmo 2 – Gradiente Estocástico Hamiltoniano Monte Carlo (SGHMC)	68

LISTA DE TABELAS

Tabela 1 – Derivadas das funções $g(\mathbf{x}_i\boldsymbol{\beta}, \phi)$ (bem como $g_\phi(\mathbf{x}_i\boldsymbol{\beta}, \phi)^k$) e $f(y_i; \mathbf{x}_i\boldsymbol{\beta}, \phi)$ do modelo k -MS.	49
Tabela 2 – Valores de k e dos parâmetros considerados no estudo de simulação.	53
Tabela 3 – Avaliação dos estimadores dos parâmetros do modelo k -IS via procedimento clássico, considerando diferentes valores de k e n	54
Tabela 4 – Avaliação dos estimadores dos parâmetros do modelo k -DS via procedimento clássico, considerando diferentes valores de k e n	55
Tabela 5 – Estimativas (e intervalos de confiança de 95%) dos parâmetros do modelo k -IS, via procedimento clássico.	56
Tabela 6 – Estimativas (e intervalos de confiança de 95%) dos parâmetros do modelo k -DS, via procedimento clássico.	58
Tabela 7 – Valores dos parâmetros para a geração de conjuntos de dados do modelo Skellam.	59
Tabela 8 – Estimativas (e intervalos de confiança de 95%) dos parâmetros do modelo k -MS ajustados aos dados gerados do modelo Skellam tradicional de tamanho $n = 100$	61
Tabela 9 – Estimativas (e intervalos de confiança de 95%) dos parâmetros do modelo k -MS ajustados aos dados gerados do modelo Skellam tradicional de tamanho $n = 200$	61
Tabela 10 – Avaliação dos estimadores dos parâmetros do modelo k -IS via procedimento bayesiano, considerando diferentes valores de k e n	72
Tabela 11 – Avaliação dos estimadores dos parâmetros do modelo k -DS via procedimento bayesiano, considerando diferentes valores de k e n	73
Tabela 12 – Estimativas (e intervalos HPD com 95% de credibilidade) dos parâmetros do modelo k -MS (k -IS e k -DS), via procedimento bayesiano.	74
Tabela 13 – Distribuição de frequências e algumas estatísticas descritivas das quantidades de <i>ticks</i> registradas na variação semanal do Ibovespa no período de 2000-2022.	78
Tabela 14 – Estimativas dos parâmetros e critérios de seleção dos modelos k -MS ajustados aos dados do Ibovespa, considerando diferentes valores de k	79
Tabela 15 – Estimativas e IC <i>bootstrap</i> (95%) dos parâmetros do modelo k -MS, com $k = -2$, ajustado aos dados do Ibovespa.	79

Tabela 16 – Distribuição de frequências e estatísticas descritivas da diferença entre jogos ganhos e perdidos (a pontuação) das equipes de cada conferência da NBA na temporada regular de 2022-2023.	81
Tabela 17 – Estimativas (e HPD 95%) dos parâmetros dos modelos k -MS ajustados aos dados de pontuação da temporada regular 2022-20023 da NBA, considerando diferentes valores de k	83
Tabela 18 – Critérios de seleção dos modelos k -MS ajustados aos dados de pontuação da temporada regular 2022-20023 da NBA, considerando diferentes valores de k	83
Tabela 19 – Derivadas das funções $g(\mathbf{x}_i\boldsymbol{\beta}, \phi)$ e $f(z_i; \mathbf{x}_i\boldsymbol{\beta}, \phi)$ do modelo k -IS.	92
Tabela 20 – Avaliação dos estimadores dos parâmetros do modelo k -IS via procedimento bayesiano, considerando diferentes valores de n	94
Tabela 21 – Distribuição de frequências e estatísticas descritivas do conjunto de dados artificiais gerado do modelo k -IS, com $k_1 = -4$, $k_2 = 1$, considerando $n = 100$	94
Tabela 22 – Distribuição de frequências e estatísticas descritivas do conjunto de dados artificiais gerado do modelo k -IS, com $k_1 = -4$, $k_2 = 1$, considerando $n = 200$	94
Tabela 23 – Estimativas (e intervalos HPD com 95% de credibilidade) dos parâmetros dos modelos k -IS ajustados considerando diferentes valores de k , via procedimento bayesiano.	95
Tabela 24 – Valores dos critérios de seleção dos modelos k -IS ajustados ao considerar diferentes valores de k	96
Tabela 25 – Estimativas (e intervalo HPD com 95% de credibilidade) dos parâmetros dos modelos k -IS, considerando diferentes valores de k , ajustados aos dados do Ibovespa.	97
Tabela 26 – Valores dos critérios de seleção dos modelos k -IS ajustados aos dados do Ibovespa, ao considerar diferentes valores de k	98

LISTA DE ABREVIATURAS E SIGLAS

<i>k</i> -IS	Skellam k_1 e k_2 Inflacionada
<i>k</i> -SS	Skellam k_1 e k_2 Subtraída
\mathcal{B}	Viés Absoluto Médio
<i>k</i> -DS	Skellam <i>k</i> -Deflacionada
<i>k</i> -IS	Skellam <i>k</i> -Inflacionada
<i>k</i> -MS	Skellam <i>k</i> -Modificada
<i>k</i> -SS	Skellam <i>k</i> -Subtraída
AIC	Critério de Informação de Akaike
BIC	Critério de Informação Bayesiano
EAIC	Critério de Informação de Akaike Esperado
EBIC	Critério de Informação Bayesiano Esperado
EQM	Erro Quadrático Médio
FMP	Função de Massa de Probabilidade
HMC	Hamiltoniano Monte Carlo
HPD	<i>Highest Probability Density</i>
Ibovespa	Índice da Bolsa de Valores de São Paulo
MCMC	Monte Carlo em Cadeia de Markov
MH	Metropolis-Hastings
NBA	<i>National Basketball Association</i>
PC	Probabilidade de Cobertura
SGHMC	Gradiente Estocástico Hamiltoniano de Monte Carlo
Var	Variância do Estimador
WAIC	Critério de Informação Watanabe-Akaike
ZMS	Skellam Zero-Modificada
ZTN	Normal Zero-Truncada

SUMÁRIO

1	INTRODUÇÃO	21
1.0.1	<i>Objetivo</i>	24
1.0.2	<i>Organização do trabalho</i>	24
2	CONCEITOS E NOTAÇÕES PRELIMINARES	25
2.1	Distribuição Skellam	25
2.1.1	<i>Uma Reparametrização da Distribuição Skellam (θ_1, θ_2)</i>	26
2.1.2	<i>Propriedades da Distribuição Skellam</i>	27
2.2	Distribuição Skellam k -Subtraída	29
3	GENERALIZAÇÕES DA DISTRIBUIÇÃO SKELLAM	31
3.1	Distribuição Skellam k -Modificada	31
3.1.1	<i>Casos Particulares da Distribuição k-MS (μ, ϕ, p)</i>	33
3.1.2	<i>Propriedades Estatísticas da Distribuição k-MS (μ, ϕ, p)</i>	34
3.1.3	<i>Distribuição k-MS e sua Versão Hurdle</i>	37
3.2	Distribuição Skellam k_1 e k_2 Inflacionada	39
3.2.1	<i>Distribuição k-IS e sua Versão Hurdle</i>	42
4	MODELO DE REGRESSÃO SKELLAM k -MODIFICADO	45
4.1	O Modelo Skellam k -Modificado	45
4.1.1	<i>Modelo k-MS e sua Versão Hurdle</i>	46
4.1.2	<i>Função de Verossimilhança</i>	47
4.2	Estimação dos Parâmetros: Abordagem Clássica	50
4.2.1	<i>Algoritmo BOBYQA</i>	51
4.2.2	<i>Critérios de Seleção de Modelos</i>	52
4.2.2.1	<i>Critério de Informação de Akaike (AIC)</i>	52
4.2.2.2	<i>Critério de Informação Bayesiano (BIC)</i>	52
4.2.3	<i>Estudo de Simulação: Avaliação da Performance dos Estimadores Clássicos</i>	53
4.2.3.1	<i>Conjuntos de Dados k-Inflacionados</i>	54
4.2.3.2	<i>Conjuntos de Dados k-Deflacionados</i>	55
4.2.4	<i>Aplicação: Conjuntos de Dados Artificiais</i>	56
4.2.4.1	<i>Aplicação do Modelo Skellam k-Inflacionado</i>	56
4.2.4.2	<i>Aplicação do Modelo Skellam k-Deflacionado</i>	57

4.2.4.3	<i>Aplicação do Modelo Skellam</i>	59
4.3	Estimação dos Parâmetros: Abordagem Bayesiana	63
4.3.1	<i>Distribuição Conjunta a Posteriori</i>	63
4.3.2	<i>Algoritmo Hamiltoniano de Monte Carlo</i>	64
4.3.2.1	<i>Algoritmo Gradiente Estocástico Hamiltoniano de Monte Carlo</i>	67
4.3.3	<i>Critérios de Seleção de Modelos</i>	69
4.3.3.1	<i>Critério de Informação de Akaike Esperado (EAIC)</i>	69
4.3.3.2	<i>Critério de Informação Bayesiano Esperado (EBIC)</i>	69
4.3.3.3	<i>Critério de Informação Watanabe-Akaike (WAIC)</i>	69
4.3.4	<i>Estudo de Simulação: Avaliação da Performance dos Estimadores Bayesianos</i>	70
4.3.5	<i>Aplicação: Conjuntos de Dados Artificiais</i>	73
4.4	Aplicação: Conjunto de Dados Reais	76
4.4.1	<i>Dados do Índice da Bolsa de Valores de São Paulo</i>	76
4.4.2	<i>Dados da Temporada Regular 2022-2023 da NBA</i>	80
5	MODELO DE REGRESSÃO SKELLAM k_1 E k_2 INFLACIONADO	87
5.1	Modelo Skellam k_1 e k_2 Inflacionado	87
5.1.1	<i>Modelo k-IS e sua Versão Hurdle</i>	88
5.1.2	<i>Função de Verossimilhança</i>	88
5.1.3	<i>Função de Verossimilhança para a Versão Hurdle</i>	89
5.1.4	<i>Vetor Score</i>	91
5.2	Estimação dos Parâmetros: Abordagem Bayesiana	92
5.2.1	<i>Distribuição Conjunta a Posteriori</i>	92
5.3	Estudo de Simulação: Avaliação da Performance dos Estimadores Bayesianos	93
5.4	Aplicação: Conjuntos de Dados Artificiais	94
5.5	Aplicação: Conjunto de Dados Reais	97
5.5.1	<i>Dados do Índice da Bolsa de Valores de São Paulo</i>	97
6	COMENTÁRIOS FINAIS	101
	REFERÊNCIAS	103

INTRODUÇÃO

Estudos envolvendo dados de contagem são recorrentes em várias áreas de conhecimento e algumas distribuições discretas padrão são comumente usadas na modelagem desse tipo de dados. Porém, para determinados conjuntos de dados, algumas observações podem registrar frequências diferentes das frequências esperadas obtidas com as distribuições tradicionalmente usadas, tornando inadequado o uso destas. Explorar os dados sem levar em conta essa particularidade é inapropriado e pode acarretar em tomadas de decisões errôneas. Essa é uma problemática que tem sido discutida por muitos autores nas últimas décadas, gerando metodologias alternativas para esse cenário (SINGH, 1963), (PANDEY, 1964), (SINGH, 1965), (GERSTENKORN, 1979).

Trabalhos pioneiros da área concentram-se em abordagens para modelar a inflação de zeros. Quando um conjunto de dados possui a observação zero com uma frequência maior que a esperada ao considerar uma distribuição discreta padrão, é chamado de zero-inflacionado. Umbach (1981), um dos precursores no tema, propôs a distribuição de Poisson Zero-Inflacionada, assumindo uma mistura entre uma distribuição degenerada no zero e uma distribuição de Poisson. Em um contexto mais amplo, além da inflação de zero, Dietz e Böhning (2000) considerou o caso de deflação de zeros. Ou seja, quando a frequência de zero é menor que a esperada, sendo os conjuntos de dados referenciados como zero-deflacionados. Estes autores consideraram também o modelo de Poisson Zero-Modificado, um modelo geral, adequado para explicar o comportamento dos dados de contagem sem exigir qualquer conhecimento prévio sobre a discrepância (alta ou baixa) na frequência de zero. Além deste, o trabalho de Conceição (2013) é um exemplo de como as pesquisas na área progrediram ao longo dos anos. A ideia da zero-modificação é estendida para distribuições discretas pertencentes à família Série de Potência, que engloba várias distribuições com suporte nos números naturais, denominada de família de distribuições Série de Potência Zero-Modificada.

Em outra conjuntura, o que também é muito evidente em problemas práticos, são casos em que a alteração na frequência da observação ocorre para algum valor diferente de zero, dito

k . Como exemplo, um problema descrito por [Pandey \(1964\)](#) em um experimento para analisar a quantidade de flores da planta *Primula veris*. Verificou-se que o número de plantas com oito flores era muito superior ao das outras quantidades de flores observadas. Isso levou o autor ao uso da distribuição de Poisson inflacionada no ponto $k = 8$. Em [Murat e Szynal \(1998\)](#) é descrito o problema da ruína do jogador com dois jogadores, um dos quais (jogador A) é infinitamente rico, e o outro (jogador B) começa com k_0 unidades monetárias. Ao considerar o conjunto formado pelo número de jogos perdidos conjuntamente por estes dois apostadores contra um terceiro oponente infinitamente rico (jogador C), sem saber qual deles ganhou o jogo (jogador A ou jogador B), é inflacionado no ponto k_0 . Neste sentido, os autores estudaram a classe de distribuições Séries de Potências modificadas inflacionadas, com a inflação podendo ocorrer em qualquer um dos pontos do suporte (k inflação). Buscando uma família de distribuições mais geral, [Carvalho \(2017\)](#) estendeu esta ideia da zero-modificação para uma observação qualquer k , fornecendo uma ampla família de distribuições discretas, denominada Série de Potência k -Modificada, a qual inclui a família de distribuições Série de Potência Zero-Modificada como caso particular ($k = 0$), capaz de ser utilizada para modelar adequadamente conjuntos de dados k -inflacionados ou k -deflacionados.

Em contrapartida, pouco se é explorado sobre aspectos de modelagem/distribuição modificadas para dados que correspondem a valores inteiros positivos e negativos (o conjunto dos inteiros). No voleibol por exemplo, é comum as análises serem feitas a partir de conjuntos de dados que representam as diferenças entre as pontuações (número de *sets* vencidos) de uma equipe com seus adversários no decorrer de um torneio. Um exemplo real é descrito por [Ntzoufras, Palaskas e Drikos \(2019\)](#), em que o conjunto das diferenças entre as pontuações é $\{-3; -2; -1; 1; 2; 3\}$, correspondendo a pontuações $\{0 - 3; 1 - 3; 2 - 3; 3 - 2; 3 - 1; 3 - 0\}$, respectivamente. Ao analisar estes dados no contexto de k -modificação, poderíamos avaliar, por exemplo, o caso de deflação no ponto $k = 3$, que seria consequência de uma equipe com poucas partidas finalizadas com vitória nos três *sets*. Explorando ainda possíveis conjuntos de dados no voleibol, pode ser de interesse analisar a diferença entre os pontos obtidos pelas equipes em cada *set* e, por exemplo, podemos ter os seguintes resultados hipotéticos pelos times em um determinado confronto: 25×23 , 25×18 , 20×25 , 26×28 e 15×13 , que definem o conjunto de diferenças $\{2, 7, -5, -2, 2\}$. Para esta situação, temos a possibilidade de casos de inflação no(s) ponto(s) $k = -2$ e/ou $k = 2$, uma vez que a diferença mínima entre as pontuações para o encerramento do *set* precisa ser de dois pontos. Cabe ressaltar também que, para esta situação, não há chance de ocorrência das observações -1, 0 e 1 (probabilidades nulas).

No contexto de modelagem de variáveis aleatórias cujo suporte é formado por números inteiros (positivos e negativos), uma das distribuições de probabilidade promissoras é a distribuição Skellam, também conhecida na literatura como distribuição da diferença de duas Poisson. A distribuição Skellam foi proposta inicialmente por [Irwin \(1937\)](#) e posteriormente generalizada por [Skellam \(1946\)](#). [Irwin \(1937\)](#) introduziu a distribuição da diferença entre duas variáveis aleatórias independentes, cada uma com distribuição de Poisson, e considerando médias iguais.

Depois de uma década, [Skellam \(1946\)](#) desenvolveu a distribuição para o caso da diferença entre duas Poisson com médias diferentes. Embora a distribuição de Skellam tenha sido originalmente derivada como a diferença de duas variáveis aleatórias independentes de Poisson, também pode ser derivada como a diferença de distribuições que possuem uma estrutura variável latente trivariada específica. Isso implica que podemos derivar a distribuição Skellam como a diferença de outras distribuições, o que motiva seu uso em vários casos ([KARLIS; NTZOUFRAS, 2006](#)).

Publicações de estudos envolvendo a distribuição Skellam passaram a ser mais frequentes a partir do início deste milênio. No trabalho de [Hirakawa, Baqai e Wolfe \(2009\)](#), os autores trabalham com a distribuição Skellam em uma abordagem que consideram promissora para a redução de ruído em medidas correspondentes a dados vindos de comunicações digitais, processamento de sinais, astronomia e aplicações de imagem por ressonância magnética. [Alzaid e Omair \(2010\)](#) usam as propriedades da função *Bessel* modificada do primeiro tipo, para obter as estimativas de máxima verossimilhança dos parâmetros da distribuição Skellam. Também discutem sobre a propriedade de distribuição assintótica das estimativas de máxima verossimilhança e apresentam formas alternativas para a função de massa de probabilidade da distribuição Skellam. Tratam aplicações com dados da bolsa de valores da Arábia Saudita e do hospital Dallah. Em outro trabalho, [Alzaid e Omair \(2014\)](#), a distribuição Skellam é usada para modelar séries temporais com valores inteiros (com possibilidade de valores negativos) e correlações positivas ou negativas. Os autores introduziram o modelo autorregressivo com valores inteiros de diferença de Poisson, que pode ser utilizado como uma ferramenta para modelar dados de contagem não estacionários. Já em [Koopman, Lit e Lucas \(2017\)](#), um modelo dinâmico Skellam é usado para medir a volatilidade estocástica de mudanças discretas de alta frequência no preço de ações. Com foco no estudo de previsão da volatilidade intra-diária, mostram que o modelo dinâmico Skellam modificado fornece previsões mais precisas em comparação com abordagens alternativas de modelagem.

A primeira modificação da distribuição Skellam foi apresentada por [Karlis e Ntzoufras \(2006\)](#). Embora eles propusessem originalmente uma distribuição Skellam modificada com uma probabilidade mais alta (inflacionada em zero) de observar $Y = 0$, seu método pode ser facilmente adaptado para acomodar uma probabilidade mais baixa (deflacionada em zero) de observar $Y = 0$. Em [Karlis e Ntzoufras \(2008\)](#), os autores apontam que o tradicional uso da distribuição de Poisson como abordagem de modelagem para descrever o número de gols no futebol, pode ser questionável em casos de superdispersão nos dados observados. Como alternativa, trabalham diretamente com a diferença do número de gols, usando a distribuição Skellam numa metodologia bayesiana. Também abordam um modelo Skellam inflacionado de zero que possa acomodar o excesso de empates nos jogos. Um artigo recente, [Conceição, Suzuki e Andrade \(2020\)](#) apresentam a distribuição Skellam Zero-Modificada (ZMS), introduzindo também no contexto de modelo de regressão. Duas relevantes aplicações são apresentadas, uma delas com dados de registros de convulsões de pacientes epiléticos, visando analisar o efeito de uma nova droga ministrada em relação ao grupo placebo. Na outra aplicação, consideraram municípios do Estado da Bahia (Brasil) com Índice de Desenvolvimento Humano (IDH) superior

a 0,65, dos quais foram analisadas a diferença entre os números de notificação de casos de aids nos anos de 2015 e 2016, tendo o IDH como variável explicativa.

A relevância em promover o debate de modificações nas probabilidades de distribuições com suporte nos números inteiros, se dá diante do entendimento de que distribuições discretas adequadas para tais medidas ainda são pouco difundidas, visto que dados com valores inteiros ocorrem na medicina, na economia, na agropecuária, na indústria, no esporte e em várias outras áreas.

1.0.1 Objetivo

O objetivo deste trabalho é estender a ideia da modificação na probabilidade para além da zero-modificação na distribuição Skellam. Isso, introduzindo generalizações da distribuição Skellam, em que uma massa de probabilidade é transferida para um ponto k e, posteriormente, para dois pontos k_1 e k_2 . A partir das generalizações, se define uma etapa imprescindível, que é a construção de modelos de regressão para as generalizações da distribuição Skellam. O procedimento de estimação e inferência dos parâmetros dos modelos foram apresentados considerando duas abordagens para fins de comparação: a abordagem clássica, via método de máxima verossimilhança; e também a abordagem bayesiana, a partir da classe Monte Carlo em Cadeia de Markov, utilizando uma variante do algoritmo Hamiltoniano Monte Carlo que consiste na reformulação das equações de Hamilton ao introduzir uma componente estocástica na equação do gradiente, derivando o algoritmo Gradiente Estocástico Hamiltoniano Monte Carlo.

1.0.2 Organização do trabalho

Seis tópicos compõem a estrutura deste trabalho: no Capítulo 2 são apresentados conceitos preliminares necessários para o desenvolvimento do trabalho. Dentre eles, a distribuição Skellam com suas principais propriedades. No Capítulo 3, são introduzidas generalizações da distribuição Skellam, versão *hurdle* e suas propriedades estruturais. No Capítulo 4 é apresentada a construção do Modelo Skellam k -Modificado e sua versão *hurdle*. São utilizadas as abordagens clássica e bayesiana para o processo de estimação e inferência dos parâmetros, com as quais são realizados estudo de simulação, aplicação com dados artificiais e aplicações à dados reais. O Capítulo 5 é destinado ao Modelo Skellam k_1 e k_2 Inflacionado. É apresentado o processo de estimação e inferência dos parâmetros do modelo adotando uma abordagem bayesiana, com estudo de simulação, ilustrações da aplicação do modelo com dados artificiais e dados reais. No Capítulo 6 são feitos os comentários finais e apresentadas as propostas futuras.

CONCEITOS E NOTAÇÕES PRELIMINARES

Generalizações da distribuição Skellam podem ser construídas a partir da modificação das probabilidades da distribuição Skellam tradicional. Antes de introduzir as distribuições mais gerais, apresentamos algumas definições necessárias para desenvolvimento do trabalho.

2.1 Distribuição Skellam

Seja (X_1, X_2) um par de variáveis aleatórias discretas, em que $X_1 = W_1 + W_3$ e $X_2 = W_2 + W_3$ de modo que W_1 e W_2 são independentes e W_i , $i = 1, 2$, segue uma distribuição de Poisson com parâmetro $\theta_i > 0$ ($W_1 \sim P(\theta_1)$ e $W_2 \sim P(\theta_2)$), e W_3 segue qualquer distribuição. A variável aleatória $Y = X_1 - X_2$ tem distribuição Skellam, cuja função de massa de probabilidade (FMP) é dada por:

$$\pi_s(y; \theta_1, \theta_2) = e^{-(\theta_1 + \theta_2)} \left(\frac{\theta_1}{\theta_2} \right)^{\frac{y}{2}} \mathcal{J}_y \left(2\sqrt{\theta_1 \theta_2} \right), \quad y \in \mathbb{Z}, \quad (2.1)$$

em que $\mathbb{Z} = \{\dots, -1, 0, 1, \dots\}$ é o suporte da variável Y ; e $\mathcal{J}_y(\xi)$ é a função de Bessel modificada do primeiro tipo de ordem y e argumento ξ , dada por:

$$\mathcal{J}_y(\xi) = \left(\frac{\xi}{2} \right)^y \sum_{j=0}^{\infty} \frac{(\xi/2)^{2j}}{j!(y+j)!}, \quad \xi = 0, 1, 2, \dots$$

para todo $y \in \mathbb{Z}$, temos $\mathcal{J}_{-y}(\xi) = \mathcal{J}_y(\xi)$. Para mais detalhes ver [Arfken, Weber e Harris \(2013\)](#).

Notação: $Y \sim \text{Skellam}(\theta_1, \theta_2)$.

A média e a variância de Y são, respectivamente, $\mu = \theta_1 - \theta_2$ e $\phi = \sigma^2 = \theta_1 + \theta_2$. Usando a Desigualdade Triangular, temos $|\theta_1 - \theta_2| \leq \theta_1 + \theta_2$ e, portanto, $|\mu| \leq \phi$.

[Alzaid e Omaid \(2010\)](#) usam as propriedades da função Bessel modificada do primeiro tipo, para obter as estimativas de máxima verossimilhança dos parâmetros da distribuição Skellam.

Também discutem sobre a propriedade de distribuição assintótica dos estimadores de máxima verossimilhança e apresentam formas alternativas para a função de massa de probabilidade da distribuição Skellam.

2.1.1 Uma Reparametrização da Distribuição Skellam (θ_1, θ_2)

Como o objetivo do nosso estudo é também incluir o contexto de modelos de regressão, é conveniente reparametrizar a distribuição Skellam em termos da média e da variância. Sendo assim, a partir da média da variável aleatória com distribuição Skellam $(\mu = \theta_1 - \theta_2)$, podemos escrever

$$\begin{cases} \theta_1 = \mu + \theta_2 \\ \theta_2 = \theta_1 - \mu \end{cases},$$

e substituindo θ_1 e θ_2 nas equações dadas pela variância $(\phi = \theta_1 + \theta_2)$ temos:

$$\begin{aligned} \begin{cases} \theta_1 = \phi - \theta_2 \\ \theta_2 = \phi - \theta_1 \end{cases} &\Rightarrow \begin{cases} \theta_1 = \phi - \theta_1 + \mu \\ \theta_2 = \phi - \mu - \theta_2 \end{cases} \\ &\Rightarrow \begin{cases} \theta_1 = \frac{\phi + \mu}{2} \\ \theta_2 = \frac{\phi - \mu}{2}. \end{cases} \end{aligned} \quad (2.2)$$

Para reescrever a FMP dada em (2.1) em termos da média μ e da variância ϕ , iremos considerar θ_1 e θ_2 dado em (2.2) e substituir em (2.1). Assim, temos:

$$\begin{aligned} \pi_s(y; \mu, \phi) &= \exp \left\{ - \left(\frac{\phi + \mu}{2} \right) - \left(\frac{\phi - \mu}{2} \right) \right\} \left(\frac{\phi + \mu}{2} \right)^{\frac{y}{2}} \left(\frac{\phi - \mu}{2} \right)^{\frac{y}{2}} \mathcal{J}_y \left(2 \sqrt{\left(\frac{\phi + \mu}{2} \right) \left(\frac{\phi - \mu}{2} \right)} \right) \\ &= \exp \left\{ \frac{-\phi - \mu - \phi + \mu}{2} \right\} \left(\frac{\phi + \mu}{2} \right)^{\frac{y}{2}} \left(\frac{\phi - \mu}{2} \right)^{\frac{y}{2}} \mathcal{J}_y \left(2 \sqrt{\frac{\phi^2 - \phi\mu + \phi\mu - \mu^2}{4}} \right), \end{aligned}$$

e, portanto, a distribuição Skellam reparametrizada em termos da média μ e da variância ϕ é dada por:

$$\pi_s(y; \mu, \phi) = \exp \{-\phi\} \left(\frac{\phi + \mu}{\phi - \mu} \right)^{\frac{y}{2}} \mathcal{J}_y \left(\sqrt{\phi^2 - \mu^2} \right), \quad y \in \mathbb{Z}. \quad (2.3)$$

Notação: $Y \sim \text{Skellam}(\mu, \phi)$.

Observação 2.1. Se Y for uma variável aleatória com distribuição Skellam com média μ e variância $\phi > 0$, então a FMP de Y pode ser escrita por:

$$\pi_s(y; \mu, \phi) = g(\mu, \phi)^y f(y; \mu, \phi), \quad y \in \mathbb{Z},$$

em que $g(\mu, \phi) = \left(\frac{\phi + \mu}{\phi - \mu} \right)^{\frac{1}{2}}$ e $f(y; \mu, \phi) = \exp \{-\phi\} \mathcal{J}_y \left(\sqrt{\phi^2 - \mu^2} \right)$.

2.1.2 Propriedades da Distribuição Skellam

Propomos o uso da distribuição Skellam para desenvolver novas distribuições e construir modelos de regressão. Para isso, é fundamental estudar as propriedades da distribuição Skellam tradicional. Baseando-nos na literatura existente sobre essa distribuição, apresentamos a seguir algumas de suas propriedades mais relevantes. Para mais detalhes sobre essas propriedades, recomendamos consultar os trabalhos de [Karlis e Ntzoufras \(2008\)](#), [Alzaid e Omair \(2010\)](#).

Proposição 2.1 (Propriedade da Soma). Sejam Y_1 e Y_2 variáveis aleatórias independentes com distribuição Skellam, tais que $Y_1 \sim \text{Skellam}(\theta_1, \theta_2)$ e $Y_2 \sim \text{Skellam}(\theta_3, \theta_4)$. Então, a soma $Z = Y_1 + Y_2$ também terá distribuição Skellam, cujos parâmetros são $\theta_1 + \theta_3$ e $\theta_2 + \theta_4$, isto é, $Z \sim \text{Skellam}(\theta_1 + \theta_3, \theta_2 + \theta_4)$.

Demonstração. Considere $\theta_1^* = \theta_1 + \theta_3$ e $\theta_2^* = \theta_2 + \theta_4$. A FMP da variável aleatória $Z \sim \text{Skellam}(\theta_1^*, \theta_2^*)$ é dada por:

$$\begin{aligned} \pi_s(z; \boldsymbol{\theta}) &= \exp\left\{-\left(\theta_1^* + \theta_2^*\right)\right\} \left(\frac{\theta_1^*}{\theta_2^*}\right)^{\frac{z}{2}} \mathcal{J}_z\left(2\sqrt{\theta_1^* \theta_2^*}\right) \\ &= \exp\left\{-\left(\theta_1 + \theta_3\right) - \left(\theta_2 + \theta_4\right)\right\} \left(\frac{\theta_1 + \theta_3}{\theta_2 + \theta_4}\right)^{\frac{z}{2}} \mathcal{J}_z\left(2\sqrt{\left(\theta_1 + \theta_3\right)\left(\theta_2 + \theta_4\right)}\right), \end{aligned} \quad (2.4)$$

para todo $z \in \mathbb{Z}$, em que $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4)$ é o vetor de parâmetros associado a Z . Assim, temos que:

$$\mu^* = \theta_1^* - \theta_2^* = (\theta_1 + \theta_3) - (\theta_2 + \theta_4) \quad \text{e} \quad \phi^* = \theta_1^* + \theta_2^* = (\theta_1 + \theta_3) + (\theta_2 + \theta_4).$$

A partir de μ^* , podemos escrever $\begin{cases} (\theta_1 + \theta_3) = \mu^* + (\theta_2 + \theta_4) \\ (\theta_2 + \theta_4) = (\theta_1 + \theta_3) - \mu^* \end{cases}$ e substituindo na variância ϕ^* , temos:

$$\begin{aligned} \begin{cases} (\theta_1 + \theta_3) = \phi^* - (\theta_2 + \theta_4) \\ (\theta_2 + \theta_4) = \phi^* - (\theta_1 + \theta_3) \end{cases} &\Rightarrow \begin{cases} \theta_1 = \phi^* - (\theta_1 + \theta_3 - \mu^*) \\ \theta_2 = \phi^* - (\mu^* + \theta_2 + \theta_4) \end{cases} \\ &\Rightarrow \begin{cases} \theta_1 + \theta_3 = \frac{\phi^* + \mu^*}{2} \\ \theta_2 + \theta_4 = \frac{\phi^* - \mu^*}{2} \end{cases}. \end{aligned}$$

Reescrevendo a (2.4) em termos da média μ^* e da variância ϕ^* , temos:

$$\begin{aligned} \pi_s(z; \mu^*, \phi^*) &= \exp\left\{-\left(\frac{\phi^* + \mu^*}{2}\right) - \left(\frac{\phi^* - \mu^*}{2}\right)\right\} \left(\frac{\frac{\phi^* + \mu^*}{2}}{\frac{\phi^* - \mu^*}{2}}\right)^{\frac{z}{2}} \\ &\quad \mathcal{J}_z\left(2\sqrt{\left(\frac{\phi^* + \mu^*}{2}\right)\left(\frac{\phi^* - \mu^*}{2}\right)}\right) \\ &= \exp\{-\phi^*\} \left(\frac{\phi^* + \mu^*}{\phi^* - \mu^*}\right)^{\frac{z}{2}} \mathcal{J}_z\left(\sqrt{(\phi^*)^2 - (\mu^*)^2}\right), \quad z \in \mathbb{Z}. \end{aligned} \quad (2.5)$$

Observe que (2.5) é a FMP e é equivalente a função dada em (2.3). \square

Nota: Ao considerar uma amostra aleatória de n de variáveis aleatórias independentes de $Y \sim \text{Skellam}(\theta_1, \theta_2)$, ou seja $Y_i \sim \text{Skellam}(\theta_1, \theta_2)$, $i = 1, 2, \dots, n$, temos que $Z_n = \sum_{i=1}^n Y_i \sim \text{Skellam}(n\theta_1, n\theta_2)$ e se aproxima de uma distribuição normal rapidamente (KARLIS; NTZOUFRAS, 2006).

Proposição 2.2 (Propriedade da Diferença). Sejam Y_1 e Y_2 variáveis aleatórias independentes com distribuição Skellam, tais que $Y_1 \sim \text{Skellam}(\theta_1, \theta_2)$ e $Y_2 \sim \text{Skellam}(\theta_3, \theta_4)$. Então, a diferença $Z = Y_1 - Y_2$ também terá distribuição Skellam, cujos parâmetros são $\theta_1 + \theta_4$ e $\theta_2 + \theta_3$, isto é, $Z = Y_1 - Y_2 \sim \text{Skellam}(\theta_1 + \theta_4, \theta_2 + \theta_3)$.

Demonstração. Análoga a demonstração da Propriedade da Soma. \square

Proposição 2.3 (Coeficiente de Assimetria de Fisher). Para uma variável aleatória $Y \sim \text{Skellam}(\theta_1, \theta_2)$, o coeficiente de assimetria de Fisher é dado por $A_F = \frac{\theta_1 - \theta_2}{(\theta_1 + \theta_2)^{\frac{3}{2}}}$.

Demonstração. Por definição, o coeficiente de assimetria de Fisher é dado por $A_F = \frac{M_3}{(M_2)^{\frac{3}{2}}}$, em que M_r denota o r -ésimo momento central (obtido a partir da r -ésima derivada sucessiva da função geradora de momentos $M_Y(t)$ e fazendo $t = 0$).

A $M_Y(t)$ da distribuição Skellam tem a seguinte forma:

$$M_Y(t) = \exp[-(\theta_1 + \theta_2) + \theta_1 e^t + \theta_2 e^{-t}].$$

Calculando a segunda e a terceira derivada de $M_Y(t)$ e substituindo na equação para A_F , obtemos

$$A_F = \frac{\theta_1 - \theta_2}{(\theta_1 + \theta_2)^{\frac{3}{2}}}.$$

\square

Nota 1: A assimetria é determinada pelo sinal de $\theta_1 - \theta_2$. Portanto, para todo $\theta_1 > \theta_2$, a assimetria é positiva ou à direita; quando $\theta_1 = \theta_2$, tem-se simetria; para todo $\theta_1 < \theta_2$, a assimetria é negativa ou à esquerda.

Nota 2: Para grandes valores de $\theta_1 + \theta_2$, a distribuição pode ser suficientemente aproximada pela distribuição normal. Se o parâmetro θ_2 for muito próximo de 0, então a distribuição tende a uma distribuição de Poisson. No caso em que o parâmetro θ_1 se aproxima de 0, existirá uma distribuição de Poisson no eixo negativo.

Uma outra propriedade interessante sobre simetria é dada por $\pi_s(y; \theta_1, \theta_2) = \pi_s(-y; \theta_2, \theta_1)$ (ALZAID; OMAIR, 2010).

Proposição 2.4 (Coeficiente Momento de Curtose). Para uma variável aleatória $Y \sim \text{Skellam}(\theta_1, \theta_2)$, o coeficiente momento de curtose é $C_m = 3 + \frac{1}{\theta_1 + \theta_2}$.

Demonstração. Por definição, o coeficiente momento de curtose é dado por $C_m = \frac{M_4}{(M_2)^2}$, em que M_r denota o r -ésimo momento central. Sendo a função geradora de momentos da distribuição Skellam $M_Y(t) = \exp[-(\theta_1 + \theta_2) + \theta_1 e^t + \theta_2 e^{-t}]$, calculando a segunda e a quarta derivada de $M_Y(t)$ e substituindo na equação da definição do C_m , obtemos:

$$C_m = \frac{1 + 3(\theta_1 + \theta_2)}{\theta_1 + \theta_2} = 3 + \frac{1}{\theta_1 + \theta_2}.$$

□

Nota: Para valores grandes de θ_1 e θ_2 , isto é, $\theta_1 \rightarrow +\infty$ e $\theta_2 \rightarrow +\infty$, o coeficiente momento de curtose tende para 3 (ALZAID; OMAIR, 2010).

2.2 Distribuição Skellam k -Subtraída

A partir da distribuição Skellam, podemos definir uma distribuição para uma variável aleatória em que um determinado valor k , $k \in \mathbb{Z}$, tem probabilidade zero de ser observado. Para esta nova variável aleatória, denotamos o suporte e a FMP por $\mathbb{Z}_{\{-k\}}$ e $\pi_{k-SS}(y; \mu, \phi)$, respectivamente.

Teorema 2.1. Considere Y uma variável aleatória cujo suporte é definido pelos números inteiros, \mathbb{Z} . A função $\pi_{k-SS}(y; \mu, \phi)$, construída ao remover do suporte \mathbb{Z} o inteiro k ($k \in \mathbb{Z}$), é dada por:

$$\pi_{k-SS}(y; \mu, \phi) = \frac{\pi_s(y; \mu, \phi)}{1 - \pi_s(k; \mu, \phi)} (1 - I_{\{k\}}(y)), \quad y \in \mathbb{Z},$$

em que $I_{\{k\}}(y)$ é uma função indicadora, com $I_{\{k\}}(y) = 1$, se $y = k$ e $I_{\{k\}}(y) = 0$, se $y \neq k$, e pode ser escrita de forma equivalente por:

$$\pi_{k-SS}(y; \mu, \phi) = \frac{\pi_s(y; \mu, \phi)}{1 - \pi_s(k; \mu, \phi)}, \quad y \in \mathbb{Z}_{\{-k\}}. \quad (2.6)$$

Então, $\pi_{k-SS}(y; \mu, \phi)$ é uma FMP.

Demonstração. A Equação (2.6) é uma FMP se

(i) $\pi_{k-SS}(y; \mu, \phi) > 0$, para todo $y \in \mathbb{Z}_{\{-k\}}$; e

(ii) $\sum_{y \in \mathbb{Z}_{\{-k\}}} \pi_{k-SS}(y; \mu, \phi) = 1$.

Para provar (i), é suficiente notar que

$$\frac{\pi_s(y; \mu, \phi)}{1 - \pi_s(k; \mu, \phi)} > 0$$

uma vez que $0 < \pi_s(y; \mu, \phi) < 1$ para todo $y \in \mathbb{Z}$ (incluindo $y = k$, usado no denominador da fração). Portanto, $\pi_{k-SS}(y; \mu, \phi) > 0$, para todo $y \in \mathbb{Z}_{\{-k\}}$.

A fim de provar (ii), temos:

$$\begin{aligned} \sum_{y \in \mathbb{Z}_{\{-k\}}} \pi_{k-SS}(y; \mu, \phi) &= \frac{1}{1 - \pi_s(k; \mu, \phi)} \sum_{y \in \mathbb{Z}_{\{-k\}}} \pi_s(y; \mu, \phi) \\ &= \frac{1}{1 - \pi_s(k; \mu, \phi)} (1 - \pi_s(k; \mu, \phi)) = 1. \end{aligned}$$

Portanto, $\pi_{k-SS}(y; \mu, \phi)$ é uma FMP. □

Definição 2.1. A função $\pi_{k-SS}(y; \mu, \phi)$ dada em (2.6) é denominada distribuição Skellam k -Subtraída (k -SS).

Notação: $Y \sim k\text{-SS}(\mu, \phi)$.

Vale ressaltar que o termo k -subtraído é utilizado para representar a retirada do valor k do suporte da variável aleatória com distribuição Skellam. Ou seja, se $Y \sim k\text{-SS}(\mu, \phi)$, a probabilidade do evento $Y = k$ é zero. Este é um conceito diferente daquele empregado para distribuições truncadas (por exemplo, distribuição truncada em k), onde a probabilidade é nula para todos os valores menores ou iguais a k .

GENERALIZAÇÕES DA DISTRIBUIÇÃO SKELLAM

3.1 Distribuição Skellam k -Modificada

Considere a situação em que a frequência da observação k (conhecida) em um conjunto de dados do processo Skellam é maior (ou menor) do que o esperado pelo processo. Esse tipo de conjunto de dados é considerado k -inflacionado (ou k -deflacionado) e, em tais situações, a suposição da distribuição Skellam tradicional não é adequada. Nesse sentido, é possível construir distribuições de probabilidade com modificações principalmente na probabilidade da observação k ($k \in \mathbb{Z}$).

Desse modo, considere Y uma variável aleatória definida nos inteiros (positivos e negativos, \mathbb{Z}). A distribuição Skellam k -Modificada pode ser construída a partir da distribuição Skellam tradicional, como apresentado no Teorema 3.1.

Teorema 3.1. Seja $\pi_{k-MS}(y; \mu, \phi, p)$ uma função definida por:

$$\pi_{k-MS}(y; \mu, \phi, p) = (1 - p)I_{\{k\}}(y) + p\pi_S(y; \mu, \phi), \quad y \in \mathbb{Z}, \quad (3.1)$$

em que p é o parâmetro responsável por modificar as probabilidades da distribuição Skellam tradicional, de modo que

$$0 \leq p \leq \frac{1}{1 - \pi_S(k; \mu, \phi)}. \quad (3.2)$$

Então, $\pi_{k-MS}(y; \mu, \phi, p)$ é uma FMP.

Demonstração. Se $\pi_{k-MS}(y; \mu, \phi, p)$ é uma FMP, as seguintes propriedades devem ser satisfeitas:

- (i) $\pi_{k-MS}(y; \mu, \phi, p) \geq 0$, para todo $y \in \mathbb{Z}$; e

$$(ii) \sum_{y \in \mathbb{Z}} \pi_{k-MS}(y; \mu, \phi, p) = 1.$$

Caso (i):

Para que (i) possa ser comprovada duas situações devem ser consideradas: a primeira, se $y = k$; e a segunda, se $y \neq k$.

(i.1) Se $y = k$, então $I_{\{k\}}(y) = 1$ e

$$\pi_{k-MS}(k; \mu, \phi, p) = 1 - p + p\pi_s(k; \mu, \phi) = 1 - p(1 - \pi_s(k; \mu, \phi)).$$

Como $0 \leq p \leq \frac{1}{1 - \pi_s(k; \mu, \phi)}$, segue que $0 \leq p(1 - \pi_s(k; \mu, \phi)) \leq 1$. Logo,

$$0 \leq \pi_{k-MS}(k; \mu, \phi, p) \leq 1.$$

(i.2) Para todo $y \neq k$, temos que $I_{\{k\}}(y) = 0$ e

$$\pi_{k-MS}(y; \mu, \phi, p) = p\pi_s(y; \mu, \phi).$$

Uma vez que $\pi_s(y; \mu, \phi) \geq 0$, temos que $\pi_{k-MS}(y; \mu, \phi, p) \geq 0$.

Desta forma, de (i.1) e (i.2) segue que $\pi_{k-MS}(y; \mu, \phi, p) \geq 0$ para todo $y \in \mathbb{Z}$.

Caso (ii):

Para provar (ii), é suficiente notar que

$$\begin{aligned} \sum_{y \in \mathbb{Z}} \pi_{k-MS}(y; \mu, \phi, p) &= (1 - p) + p \sum_{y \in \mathbb{Z}} \pi_s(y; \mu, \phi) \\ &= (1 - p) + p = 1. \end{aligned}$$

Portanto, $\pi_{k-MS}(y; \mu, \phi, p)$ é uma FMP. □

Vale ressaltar que o parâmetro p na função $\pi_{k-MS}(y; \mu, \phi, p)$ pode assumir valores maiores que 1, o que a torna diferente das distribuições de mistura comumente encontradas na literatura.

Definição 3.1. A função $\pi_{k-MS}(y; \mu, \phi, p)$ dada em (3.1) é denominada distribuição Skellam k -Modificada (k -MS).

Notação: $Y \sim k\text{-MS}(\mu, \phi, p)$.

Cabe mencionar que a nova distribuição k -MS é uma generalização da distribuição ZMS proposta por [Conceição, Suzuki e Andrade \(2020\)](#), sendo essas equivalentes quando $k = 0$.

3.1.1 Casos Particulares da Distribuição k -MS (μ, ϕ, p)

Diferentes valores de p (o qual controla a probabilidade da observação k) podem levar a diferentes caracterizações da distribuição k -MS. Em termos práticos, isso acarretará na proporção de observações k presentes nos dados amostrais.

A proporção adicional ou de escassez da observação k é dada por:

$$\begin{aligned}\pi_{k-MS}(k; \mu, \phi, p) - \pi_S(k; \mu, \phi) &= 1 - p + p\pi_S(k; \mu, \phi) - \pi_S(k; \mu, \phi) \\ &= 1 - p - \pi_S(k; \mu, \phi)(1 - p) \\ &= (1 - p)(1 - \pi_S(k; \mu, \phi)).\end{aligned}\quad (3.3)$$

Assim, ao considerar diferentes valores de p ($0 \leq p \leq 1/(1 - \pi_S(k; \mu, \phi))$), temos diferentes casos particulares da distribuição k -MS, como descrito nos Corolários abaixo.

Corolário 3.1. Quando $p = 0$, $\pi_{k-MS}(y; \mu, \phi, p)$ é uma distribuição degenerada com toda massa no ponto k .

Demonstração. Para $p = 0$ em (3.3), $\pi_{k-MS}(k; \mu, \phi, p) - \pi_S(k; \mu, \phi) = 1 - \pi_S(k; \mu, \phi)$. Logo, $\pi_{k-MS}(k; \mu, \phi, p) = 1$. \square

Corolário 3.2. Para todo $0 < p < 1$, $\pi_{k-MS}(y; \mu, \phi, p)$ é a distribuição Skellam k -Inflacionada (k -IS), a qual tem uma proporção adicional de k .

Demonstração. Para todo $0 < p < 1$ em (3.3), temos que $(1 - p)(1 - \pi_S(k; \mu, \phi)) > 0$. Portanto, $\pi_{k-MS}(k; \mu, \phi, p) > \pi_S(k; \mu, \phi)$. \square

Corolário 3.3. Quando $p = 1$, $\pi_{k-MS}(y; \mu, \phi, p)$ é a distribuição Skellam usual.

Demonstração. Para $p = 1$ em (3.3), $\pi_{k-MS}(k; \mu, \phi, p) - \pi_S(k; \mu, \phi) = 0$. Logo, $\pi_{k-MS}(k; \mu, \phi, p) = \pi_S(k; \mu, \phi)$. \square

Corolário 3.4. Para todo $1 < p < \frac{1}{1 - \pi_S(k; \mu, \phi)}$, $\pi_{k-MS}(y; \mu, \phi, p)$ é a distribuição Skellam k -Deflacionada (k -DS), a qual tem uma proporção de escassez de k .

Demonstração. Para todo $1 < p < \frac{1}{1 - \pi_S(k; \mu, \phi)}$ em (3.3), temos que $(1 - p)(1 - \pi_S(k; \mu, \phi)) < 0$. Portanto, $\pi_{k-MS}(k; \mu, \phi, p) < \pi_S(k; \mu, \phi)$. \square

Corolário 3.5. Quando $p = \frac{1}{1 - \pi_S(k; \mu, \phi)}$, $\pi_{k-MS}(y; \mu, \phi, p)$ é uma distribuição k -SS.

Demonstração. Para $p = \frac{1}{1 - \pi_S(k; \mu, \phi)}$ em (3.3), $\pi_{k-MS}(k; \mu, \phi, p) - \pi_S(k; \mu, \phi) = -\pi_S(k; \mu, \phi)$. O que implica em $\pi_{k-MS}(k; \mu, \phi, p) = 0$. \square

Tendo conhecimento dos Corolários 3.1 - 3.5, ao ajustar a distribuição k -MS aos dados amostrais, a estimativa do parâmetro p possibilita identificar em qual dos casos particulares o conjunto de dados se enquadra e, conseqüentemente, a sua caracterização com relação ao tipo de k -modificação.

3.1.2 Propriedades Estatísticas da Distribuição k -MS (μ, ϕ, p)

No estudo de uma distribuição de probabilidade é essencial compreender as propriedades estatísticas. Dessa forma, algumas propriedades importantes da distribuição k -MS são descritas a seguir.

Formalizando, considere Y uma variável aleatória definida sobre \mathbb{Z} , com distribuição de probabilidade dada em (3.1). É interessante listar as propriedades estatísticas de Y em função da distribuição Skellam tradicional.

Corolário 3.6 (Função de Distribuição). A função de distribuição de Y , F_Y , é dada por:

$$F_Y(b) = \begin{cases} pF(b), & \text{se } b < k \\ 1 - p(1 - F(b)), & \text{se } b \geq k, \end{cases},$$

em que $F(b)$ é a função de distribuição da Skellam tradicional no ponto b , para todo $b \in \mathbb{Z}$.

Demonstração. Por definição de função de distribuição, temos:

$$\begin{aligned} F_Y(b) = P(Y \leq b) &= \sum_{y:y \leq b} \left\{ (1-p)I_{\{k\}}(y) + p\pi_s(y; \mu, \phi) \right\} \\ &= \sum_{y:y \leq b} (1-p)I_{\{k\}}(y) + \sum_{y:y \leq b} p\pi_s(y; \mu, \phi) \\ &= (1-p) \sum_{y:y \leq b} I_{\{k\}}(y) + pF(b), \end{aligned}$$

$\forall b \in \mathbb{Z}$. Portanto,

$$F_Y(b) = \begin{cases} pF(b), & \text{se } b < k \\ 1 - p(1 - F(b)), & \text{se } b \geq k \end{cases}.$$

□

Corolário 3.7 (Função Geradora de Probabilidade). A função geradora de probabilidade de Y , G_Y , é dada por:

$$G_Y(z) = (1-p)z^k + pG(z),$$

para todo z tal que $G(z)$ existe, em que $G(z)$ é a função geradora de probabilidade da distribuição Skellam tradicional.

Demonstração. Por definição de função geradora de probabilidade, temos:

$$\begin{aligned}
 G_Y(z) = E(z^Y) &= \sum_{y \in \mathbb{Z}} z^y \left\{ (1-p)I_{\{k\}}(y) + p\pi_s(y; \mu, \phi) \right\} \\
 &= \sum_{y \in \mathbb{Z}} z^y (1-p)I_{\{k\}}(y) + \sum_{y \in \mathbb{Z}} z^y p\pi_s(y; \mu, \phi) \\
 &= (1-p)z^k + p \sum_{y \in \mathbb{Z}} z^y p\pi_s(y; \mu, \phi) \\
 &= (1-p)z^k + pG(z).
 \end{aligned}$$

□

Corolário 3.8 (Função Característica). A função característica de Y , φ_Y , é dada por:

$$\varphi_Y(t) = (1-p)e^{itk} + p\varphi(t),$$

em que $\varphi(t)$ é a função característica da distribuição Skellam tradicional.

Demonstração. Por definição de função característica, temos:

$$\begin{aligned}
 \varphi_Y(t) = E(e^{itY}) &= \sum_{y \in \mathbb{Z}} e^{ity} \left\{ (1-p)I_{\{k\}}(y) + p\pi_s(y; \mu, \phi) \right\} \\
 &= \sum_{y \in \mathbb{Z}} e^{ity} (1-p)I_{\{k\}}(y) + p \sum_{y \in \mathbb{Z}} e^{ity} \pi_s(y; \mu, \phi) \\
 &= e^{itk}(1-p) + p\varphi(t).
 \end{aligned}$$

□

Corolário 3.9 (Função Geradora de Momentos). A função geradora de momentos de Y , M_Y , é dada por:

$$M_Y(t) = e^{tk}(1-p) + pM(t),$$

para todo t tal que $M(t)$ existe, em que $M(t)$ é a função geradora de momentos da distribuição Skellam tradicional.

Demonstração. Por definição de função geradora de momentos, temos:

$$\begin{aligned}
 M_Y(t) = E(e^{tY}) &= \sum_{y \in \mathbb{Z}} e^{ty} \left\{ (1-p)I_{\{k\}}(y) + p\pi_s(y; \mu, \phi) \right\} \\
 &= \sum_{y \in \mathbb{Z}} e^{ty} (1-p)I_{\{k\}}(y) + p \sum_{y \in \mathbb{Z}} e^{ty} \pi_s(y; \mu, \phi) \\
 &= e^{tk}(1-p) + pM(t).
 \end{aligned}$$

□

Ao derivar sucessivamente $M_Y(t)$ em relação a t , obtemos

$$M_Y^{(r)}(t) = \frac{d^r M_Y(t)}{dt^r} = (1-p)k^r e^{tk} + pM^{(r)}(t), \quad (3.4)$$

em que (r) indica a derivada de r -ésima ordem de $M_Y(t)$ em relação a t .

Assim, o r -ésimo momento da distribuição k -MS, denotado por $m_Y^{(r)} = E(Y^r)$, é obtido por fazer $t = 0$ em (3.4), isto é

$$m_Y^{(r)} = (1-p)k^r + pm^{(r)}, \quad (3.5)$$

em que $m^{(r)}$ é o r -ésimo momento da distribuição Skellam tradicional.

Proposição 3.1. (Média e Variância). Se Y é uma variável aleatória com distribuição k -MS(μ, ϕ, p), então sua média μ_{k-MS} e variância σ_{k-MS}^2 são dadas, respectivamente, por:

$$\mu_{k-MS} = (1-p)k + p\mu \quad \text{e} \quad \sigma_{k-MS}^2 = p\{(1-p)(k-\mu)^2 + \phi\}$$

em que μ a média e $\phi = \sigma^2$ a variância da distribuição Skellam tradicional.

Demonstração. Prova imediata usando a equação (3.5). □

Nota: Uma vez que o parâmetro p pode assumir valores maiores que 1, asseguramos que $\sigma_{k-MS}^2 = m_Y^{(2)} - \{m_Y^{(1)}\}^2 \geq 0$ a partir da Desigualdade de Jensen, a qual diz que $m_Y^{(2)} \geq \{m_Y^{(1)}\}^2$.

Proposição 3.2 (Função Geradora de Cumulantes). Seja M_Y a função geradora de momentos de Y . A função geradora de cumulantes de Y , Ψ_Y , é dada por:

$$\begin{aligned} \Psi_Y(t) &= \log(M_Y(t)) \\ &= \log(e^{tk}(1-p) + pM(t)), \end{aligned}$$

e o r -ésimo cumulante da distribuição k -MS é obtido avaliando a r -ésima derivada de $\Psi_Y(t)$ no ponto $t = 0$.

Demonstração. Prova imediata a partir da definição da função geradora de cumulantes. □

Proposição 3.3 (Coeficiente de Dispersão). O coeficiente de dispersão da distribuição k -MS é dada por

$$d_{k-MS} = \frac{p\{(1-p)(k-\mu)^2 + \phi\}}{|(1-p)k + p\mu|},$$

a qual é uma medida que fornece o grau de dispersão existente na distribuição.

Demonstração. A prova é imediata a partir da definição do coeficiente de dispersão, dada por

$$d_{k-MS} = \frac{\sigma_{k-MS}^2}{|\mu_{k-MS}|}.$$

Pela Proposição (3.1), temos que $\mu_{k-MS} = (1-p)k + p\mu$ e $\sigma_{k-MS}^2 = p\{(1-p)(k-\mu)^2 + \phi\}$. Portanto,

$$d_{k-MS} = \frac{p\{(1-p)(k-\mu)^2 + \phi\}}{|(1-p)k + p\mu|}.$$

□

A construção deste coeficiente vem da propriedade de equidispersão ($\mu = \sigma^2$) da distribuição de Poisson. A partir desta medida, podemos concluir sobre a dispersão:

- Se $0 < d_{k-MS} < 1$, temos que $\sigma_{k-MS}^2 < |\mu_{k-MS}|$ e dizemos que a distribuição k -MS é subdispersa;
- Se $d_{k-MS} = 1$, temos que $\sigma_{k-MS}^2 = |\mu_{k-MS}|$ e dizemos que a distribuição k -MS é equidispersa;
- Se $d_{k-MS} > 1$, temos que $\sigma_{k-MS}^2 > |\mu_{k-MS}|$ e dizemos que a distribuição k -MS é sobredispersa.

3.1.3 Distribuição k -MS e sua Versão Hurdle

Apesar dos modelos *hurdle* serem conceitualmente mais amplos (ver [Mullahy \(1986\)](#) e [King \(1989\)](#)), uma reparametrização da distribuição k -MS fornece uma estrutura similar. Assim, a versão *hurdle* da distribuição k -MS é uma forma de trabalhar com uma distribuição de duas partes: uma que explica a ocorrência da observação k ; e outra que explica a ocorrência das observações diferentes de k .

Corolário 3.10. A FMP dada em (3.1) pode ser escrita da seguinte forma:

$$\pi_{k-MS}(y; \mu, \phi, \omega) = (1 - \omega)I_{\{k\}}(y) + \omega\pi_{k-SS}(y; \mu, \phi), \quad y \in \mathbb{Z}, \quad (3.6)$$

em que $\omega \in \Omega = [0, 1]$ é o parâmetro de modificação das probabilidades da distribuição k -SS.

Demonstração. Da Equação (3.1), podemos reescrever $\pi_{k-MS}(y; \mu, \phi, p)$ usando a função indicadora $I_{\{k\}}(y)$:

$$\begin{aligned} \pi_{k-MS}(y; \mu, \phi, p) &= \pi_{k-MS}(k; \mu, \phi, p)I_{\{k\}}(y) + \pi_{k-MS}(y; \mu, \phi, p)(1 - I_{\{k\}}(y)) \\ &= \{1 - p(1 - \pi_s(k; \mu, \phi))\}I_{\{k\}}(y) + \\ &\quad p(1 - \pi_s(k; \mu, \phi)) \left\{ \frac{\pi_s(y; \mu, \phi)}{1 - \pi_s(k; \mu, \phi)} \right\} (1 - I_{\{k\}}(y)) \\ &= \{1 - p(1 - \pi_s(k; \mu, \phi))\}I_{\{k\}}(y) + p(1 - \pi_s(k; \mu, \phi))\pi_{k-SS}(y; \mu, \phi), \quad \forall y \in \mathbb{Z}. \end{aligned}$$

Uma vez que $0 \leq p \leq \frac{1}{1 - \pi_s(k; \mu, \phi)}$, então $0 \leq p(1 - \pi_s(k; \mu, \phi)) \leq 1$. Definindo $\omega = p(1 - \pi_s(k; \mu, \phi))$, temos:

$$\pi_{k-MS}(y; \mu, \phi, \omega) = (1 - \omega)I_{\{k\}}(y) + \omega\pi_{k-SS}(y; \mu, \phi), \forall y \in \mathbb{Z}.$$

em que $0 \leq \omega \leq 1$. □

Do Corolário 3.10, $\pi_{k-MS}(y; \mu, \phi, \omega)$ pode ser interpretada como uma distribuição de mistura, sendo $1 - \omega$ a probabilidade do evento $y = k$ e $\omega\pi_{k-SS}(y; \mu, \phi)$ a probabilidade de uma observação y , $y \neq k$. Desta forma, esta representação da distribuição k -MS pode ser interpretada como uma superposição de dois processos, isto é, um que produz apenas a observação k e outro que produz observações diferentes de k a partir de uma distribuição k -SS. Para a distribuição k -MS parametrizada por ω , denotaremos por $k-MS(\mu, \phi, \omega)$.

Vale frisar ainda que a versão *hurdle* da distribuição k -MS contém a distribuição k -SS como uma mistura, que difere da representação de mistura tradicional de distribuições inflacionadas.

De maneira similar, diferentes valores de ω podem levar a diferentes caracterizações da distribuição $k-MS(\mu, \phi, \omega)$. Assim, a proporção adicional ou de escassez da observação k é dada por:

$$\pi_{k-MS}(k; \mu, \phi, \omega) - \pi_s(k; \mu, \phi) = 1 - \omega - \pi_s(k; \mu, \phi). \quad (3.7)$$

Ao considerar diferentes valores de ω ($0 \leq \omega \leq 1$) em (3.7), temos os casos particulares descritos nos Corolários a seguir.

Corolário 3.11. Quando $\omega = 0$, $\pi_{k-MS}(y; \mu, \phi, \omega)$ é uma distribuição degenerada com toda massa no ponto k .

Demonstração. Para $\omega = 0$ em (3.7), $\pi_{k-MS}(k; \mu, \phi, \omega) - \pi_s(k; \mu, \phi) = 1 - \pi_s(k; \mu, \phi)$. Logo, $\pi_{k-MS}(k; \mu, \phi, \omega) = 1$. □

Corolário 3.12. Para todo $0 < \omega < (1 - \pi_s(k; \mu, \phi))$, $\pi_{k-MS}(y; \mu, \phi, \omega)$ é a distribuição k -IS.

Demonstração. Para todo $0 < \omega < (1 - \pi_s(k; \mu, \phi))$ em (3.7), temos que $(1 - \pi_s(k; \mu, \phi)) - \omega > 0$. Portanto, $\pi_{k-MS}(k; \mu, \phi, \omega) > \pi_s(k; \mu, \phi)$. □

Corolário 3.13. Quando $\omega = 1 - \pi_s(k; \mu, \phi)$, $\pi_{k-MS}(y; \mu, \phi, \omega)$ é a distribuição Skellam tradicional.

Demonstração. Para $\omega = 1 - \pi_s(k; \mu, \phi)$ em (3.7), $\pi_{k-MS}(k; \mu, \phi, \omega) - \pi_s(k; \mu, \phi) = 0$. Logo, $\pi_{k-MS}(k; \mu, \phi, \omega) = \pi_s(k; \mu, \phi)$. □

Corolário 3.14. Para todo $(1 - \pi_S(k; \mu, \phi)) < \omega < 1$, $\pi_{k-MS}(y; \mu, \phi, \omega)$ é a distribuição k -DS.

Demonstração. Para todo $(1 - \pi_S(k; \mu, \phi)) < \omega < 1$ em (3.7), temos que $(1 - \pi_S(k; \mu, \phi)) - \omega < 0$. Portanto, $\pi_{k-MS}(k; \mu, \phi, \omega) < \pi_S(k; \mu, \phi)$. \square

Corolário 3.15. Quando $\omega = 1$, $\pi_{k-MS}(y; \mu, \phi, \omega)$ é uma distribuição k -SS.

Demonstração. Para $\omega = 1$ em (3.7), $\pi_{k-MS}(k; \mu, \phi, \omega) - \pi_S(k; \mu, \phi) = -\pi_S(k; \mu, \phi)$ e implica em $\pi_{k-MS}(k; \mu, \phi, \omega) = 0$. \square

3.2 Distribuição Skellam k_1 e k_2 Inflacionada

Considere agora a situação em que um conjunto de dados do processo Skellam apresenta altas frequências (inflações) em duas observações, ditas k_1 e k_2 (conhecidas). Em problemas reais, é muito comum a discrepância das frequências ocorrer nas observações zero e um ($k_1 = 0$ e $k_2 = 1$), por exemplo.

Para um contexto mais geral, em que k_1 e k_2 podem assumir quaisquer valores inteiros ($k_1, k_2 \in \mathbb{Z}$, tal que $k_1 \neq k_2$), propomos nesta seção a distribuição Skellam k_1 e k_2 Inflacionada. A ideia com essa nova proposta é explicar adequadamente o comportamento dos dados que apresentam tal característica.

Assim, considere agora uma variável aleatória Z definida sobre os inteiros (positivos e negativos), isto é, Z assume valores em \mathbb{Z} .

Teorema 3.2. Seja a função $\pi_{k-IS}(z; \mu, \phi, \mathbf{p})$ dada por:

$$\pi_{k-IS}(z; \mu, \phi, \mathbf{p}) = (1 - p_1)I_{\{k_1\}}(z) + (1 - p_2)I_{\{k_2\}}(z) + p_0\pi_S(z; \mu, \phi), \quad \forall z \in \mathbb{Z}, \quad (3.8)$$

em que $\mathbf{k} = (k_1, k_2)$ é um vetor correspondente aos pontos de modificações ($k_1, k_2 \in \mathbb{Z}$, tal que $k_1 \neq k_2$); $I_{\{k_j\}}(z)$, com $j = 1, 2$, é uma função indicadora, tal que:

$$I_{\{k_j\}}(z) = \begin{cases} 1, & \text{se } z = k_j \\ 0, & \text{se } z \neq k_j \end{cases};$$

e $\mathbf{p} = (p_0, p_1, p_2)$ é o vetor de parâmetros responsáveis pela modificação das probabilidades (principalmente das observações k_1 e k_2) em relação à distribuição Skellam tradicional, sob as condições:

C1. $p_t \in \mathcal{P} = (0, 1)$, $\forall t = 0, 1, 2$;

C2. $p_0 = p_1 + p_2 - 1$, tal que $p_0 + \sum_{t=1}^2 (1 - p_t) = 1$.

Então, a função $\pi_{k-IS}(z; \mu, \phi, \mathbf{p})$ é uma FMP.

Demonstração. Para provar que (3.8) é uma FMP, precisamos mostrar:

$$(i) \pi_{k-IS}(z; \mu, \phi, \mathbf{p}) > 0, \forall z \in \mathbb{Z};$$

$$(ii) \sum_{z: z \in \mathcal{A}_S} \pi_{k-IS}(z; \mu, \phi, \mathbf{p}) = 1.$$

Caso (i):

(i.1) Se $z = k_1$, temos:

$$\pi_{k-IS}(k_1; \mu, \phi, \mathbf{p}) = (1 - p_1) + p_0 \pi_S(k_1; \mu, \phi).$$

Uma vez que $0 < p_1 < 1$ e $0 < p_0 < 1$, então $\pi_{k-IS}(k_1; \mu, \phi, \mathbf{p}) > 0$.

(i.2) Se $z = k_2$, temos:

$$\pi_{k-IS}(k_2; \mu, \phi, \mathbf{p}) = (1 - p_2) + p_0 \pi_S(k_2; \mu, \phi).$$

Novamente, uma vez que $0 < p_2 < 1$ e $0 < p_0 < 1$, então $\pi_{k-IS}(k_2; \mu, \phi, \mathbf{p}) > 0$.

(i.3) Para todo $z \in \mathbb{Z}$, em que $z \neq k_1$ e $z \neq k_2$, temos:

$$\pi_{k-IS}(z; \mu, \phi, \mathbf{p}) = p_0 \pi_S(z; \mu, \phi).$$

Dado que $0 < p_0 < 1$, então $\pi_{k-IS}(z; \mu, \phi, \mathbf{p}) > 0$.

Logo, $\pi_{k-IS}(z; \mu, \phi, \mathbf{p}) > 0, \forall z \in \mathbb{Z}$.

Caso (ii):

$$\begin{aligned} \sum_{z \in \mathbb{Z}} \pi_{k-IS}(z; \mu, \phi, \mathbf{p}) &= \sum_{z \in \mathbb{Z}} \left\{ (1 - p_1) I_{\{k_1\}}(z) + (1 - p_2) I_{\{k_2\}}(z) + p_0 \pi_S(z; \mu, \phi) \right\} \\ &= (1 - p_1) + (1 - p_2) + p_0 \sum_{z \in \mathbb{Z}} \pi_S(z; \mu, \phi) \\ &= (1 - p_1) + (1 - p_2) + p_0 \\ &= (1 - p_1) + (1 - p_2) + p_1 + p_2 - 1 \\ &= 1. \end{aligned}$$

Uma vez que os itens (i) e (ii) foram verificados, temos, portanto, que $\pi_{k-IS}(z; \mu, \phi, \mathbf{p})$ é uma FMP. \square

Definição 3.2. A função $\pi_{k-IS}(z; \mu, \phi, \mathbf{p})$ dada pela Equação (3.8) é denominada distribuição Skellam k_1 e k_2 Inflacionada (\mathbf{k} -IS).

Notação: \mathbf{k} -IS(μ, ϕ, \mathbf{p}).

A FMP (3.8) pode ainda ser escrita da seguinte forma:

$$\pi_{k-IS}(z; \mu, \phi, \mathbf{p}) = \begin{cases} (1 - p_1) + (p_1 + p_2 - 1)\pi_s(k_1; \mu, \phi), & \text{se } z = k_1 \\ (1 - p_2) + (p_1 + p_2 - 1)\pi_s(k_2; \mu, \phi), & \text{se } z = k_2 \\ (p_1 + p_2 - 1)\pi_s(z; \mu, \phi), & \forall z \neq k_1 \text{ e } z \neq k_2, \text{ com } z \in \mathbb{Z} \\ 0, & \text{caso contrário} \end{cases},$$

uma vez que $p_0 = p_1 + p_2 - 1$.

Corolário 3.16 (Função Geradora de Momentos). A função geradora de momentos de Z , M_Z , é dada por:

$$M_Z(t) = (1 - p_1)e^{tk_1} + (1 - p_2)e^{tk_2} + p_0M(t),$$

para todo t tal que $M(t)$ existe, em que $M(t)$ é a função geradora de momentos da distribuição Skellam tradicional.

Demonstração. Por definição de função geradora de momentos, temos:

$$\begin{aligned} M_Z(t) = E(e^{tZ}) &= \sum_{z \in \mathbb{Z}} e^{tz} \left\{ (1 - p_1)I_{\{k_1\}}(z) + (1 - p_2)I_{\{k_2\}}(z) + p_0\pi_s(z; \mu, \phi) \right\} \\ &= \sum_{z \in \mathbb{Z}} e^{tz}(1 - p_1)I_{\{k_1\}}(z) + \sum_{z \in \mathbb{Z}} e^{tz}(1 - p_2)I_{\{k_2\}}(z) + p_0 \sum_{y \in \mathbb{Z}} e^{ty}\pi_s(y; \mu, \phi) \\ &= (1 - p_1)e^{tk_1} + (1 - p_2)e^{tk_2} + p_0M(t). \end{aligned}$$

□

Ao derivar sucessivamente $M_Z(t)$ em relação a t , obtemos

$$M_Z^{(r)}(t) = \frac{d^r M_Z(t)}{dt^r} = (1 - p_1)k_1^r e^{tk_1} + (1 - p_2)k_2^r e^{tk_2} + p_0M^{(r)}(t), \quad (3.9)$$

em que (r) indica a derivada de r – ésima ordem de $M_Z(t)$ em relação a t .

Assim, o r – ésimo momento da distribuição \mathbf{k} -IS, denotado por $m_Z^{(r)} = E(Z^r)$, é obtido por fazer $t = 0$ em (3.9), isto é

$$m_Z^{(r)} = (1 - p_1)k_1^r + (1 - p_2)k_2^r + p_0m^{(r)}, \quad (3.10)$$

em que $m^{(r)}$ é o r – ésimo momento da distribuição Skellam tradicional.

Proposição 3.4. (Média e Variância). Se Z é uma variável aleatória com distribuição \mathbf{k} -IS(μ, ϕ, \mathbf{p}), então sua média $\mu_{\mathbf{k-IS}}$ e variância $\sigma^2_{\mathbf{k-IS}}$ são dadas, respectivamente, por:

$$\mu_{\mathbf{k-IS}} = (1 - p_1)k_1 + (1 - p_2)k_2 + p_0\mu \quad \text{e} \quad \sigma^2_{\mathbf{k-IS}} = (1 - p_1)k_1^2 + (1 - p_2)k_2^2 + p_0(\sigma^2 + \mu^2)$$

em que μ a média e $\phi = \sigma^2$ a variância da distribuição Skellam tradicional.

Demonstração. Prova imediata usando a equação (3.10). □

É possível destacar algumas especificidades da distribuição \mathbf{k} -IS:

- Quando $p_1 \rightarrow 1$ (ou $p_2 \rightarrow 1$) temos a distribuição k_2 -IS(μ, ϕ, p_2) (ou k_1 -IS(μ, ϕ, p_1)), ou seja, a modificação existe em um único ponto do suporte (k -inflação, com $k = k_1$ ou $k = k_2$);
- Quando $p_1 \rightarrow 1$ e $p_2 \rightarrow 1$ temos a distribuição Skellam tradicional.
- Quando $p_1 = 0,5$ e $p_2 = 0,5$ temos $p_0 = 0$, o que implica em $\pi_{\mathbf{k-IS}}(z; \mu, \phi, \mathbf{p}) = (1 - p_1)I_{\{k_1\}}(z) + (1 - p_2)I_{\{k_2\}}(z)$. Dessa forma, a probabilidade de observação de k_1 e k_2 pode ser vista como um ensaio de Bernoulli com possibilidade de ocorrência de sucesso ou fracasso.

3.2.1 Distribuição \mathbf{k} -IS e sua Versão Hurdle

A partir da função $\pi_{\mathbf{k-IS}}(z; \mu, \phi, \mathbf{p})$, que corresponde a distribuição \mathbf{k} -IS, é possível definir a versão *hurdle*, a qual é vista como a superposição de três processos: um que explica a probabilidade de k_1 ; outro que explica a probabilidade de k_2 ; e um último que explica as probabilidades das observações diferentes de k_1 e de k_2 . A seguir é apresentada sua construção.

Corolário 3.17. A FMP dada em (3.8) pode ser escrita como:

$$\pi_{\mathbf{k-IS}}(z; \mu, \phi, \boldsymbol{\omega}) = (1 - \omega_1)I_{\{k_1\}}(z) + (1 - \omega_2)I_{\{k_2\}}(z) + \omega_0\pi_{\mathbf{k-SS}}(z; \mu, \phi), \quad z \in \mathbb{Z}, \quad (3.11)$$

em que $\pi_{\mathbf{k-SS}}(z; \mu, \phi)$ é a FMP da distribuição Skellam k_1 e k_2 Subtraída (\mathbf{k} -SS), dada por:

$$\pi_{\mathbf{k-SS}}(z; \mu, \phi) = \left(\frac{\pi_s(z; \mu, \phi)}{1 - \pi_s(k_1; \mu, \phi) - \pi_s(k_2; \mu, \phi)} \right) (1 - I_{\{k_1\}}(z)) (1 - I_{\{k_2\}}(z)), \quad z \in \mathbb{Z};$$

$\boldsymbol{\omega} = (\omega_0, \omega_1, \omega_2)$ é o vetor de parâmetros responsáveis pela modificação das probabilidades, sendo $\omega_0 = \omega_1 + \omega_2 - 1$ e satisfaz a condição $\omega_0 + \sum_{i=1}^2 (1 - \omega_i) = 1$. Além disso, segue que ω_1 e ω_2 tem espaços paramétricos condicionados aos parâmetros p_1 e p_2 :

$$p_1 - \pi_s(k_1; \mu, \phi) < \omega_1 < p_1 \quad \text{e} \quad p_2 - \pi_s(k_2; \mu, \phi) < \omega_2 < p_2.$$

Demonstração. A função $\pi_{k-IS}(z; \mu, \phi, \mathbf{p})$ dada na Equação (3.8) pode ser reescrita por meio da função indicadora $I_{\{k_j\}}(z)$, $j = 1, 2$:

$$\begin{aligned}
\pi_{k-IS}(z; \mu, \phi, \mathbf{p}) &= \pi_{k-IS}(z; \mu, \phi, \mathbf{p})I_{\{k_1\}}(z) + \pi_{k-IS}(z; \mu, \phi, \mathbf{p})I_{\{k_2\}}(z) \\
&+ \pi_{k-IS}(z; \mu, \phi, \mathbf{p})(1 - I_{\{k_1\}}(z))(1 - I_{\{k_2\}}(z)) \\
&= \left((1 - p_1) + p_0\pi_s(k_1; \mu, \phi) \right) I_{\{k_1\}}(z) + \left((1 - p_2) + p_0\pi_s(k_2; \mu, \phi) \right) I_{\{k_2\}}(z) \\
&+ p_0\pi_s(z; \mu, \phi)(1 - I_{\{k_1\}}(z))(1 - I_{\{k_2\}}(z)) \\
&= \left((1 - p_1) + p_0\pi_s(k_1; \mu, \phi) \right) I_{\{k_1\}}(z) + \left((1 - p_2) + p_0\pi_s(k_2; \mu, \phi) \right) I_{\{k_2\}}(z) \\
&+ p_0 \left(\frac{1 - \pi_s(k_1; \mu, \phi) - \pi_s(k_2; \mu, \phi)}{1 - \pi_s(k_1; \mu, \phi) - \pi_s(k_2; \mu, \phi)} \right) \pi_s(z; \mu, \phi)(1 - I_{\{k_1\}}(z))(1 - I_{\{k_2\}}(z)) \\
&= \left((1 - p_1) + p_0\pi_s(k_1; \mu, \phi) \right) I_{\{k_1\}}(z) + \left((1 - p_2) + p_0\pi_s(k_2; \mu, \phi) \right) I_{\{k_2\}}(z) \\
&+ p_0 \left(1 - \pi_s(k_1; \mu, \phi) - \pi_s(k_2; \mu, \phi) \right) \pi_{k-SS}(z; \mu, \phi),
\end{aligned}$$

em que

$$\pi_{k-SS}(z; \mu, \phi) = \left(\frac{\pi_s(z; \mu, \phi)}{1 - \pi_s(k_1; \mu, \phi) - \pi_s(k_2; \mu, \phi)} \right) (1 - I_{\{k_1\}}(z))(1 - I_{\{k_2\}}(z)),$$

corresponde a FMP da distribuição \mathbf{k} -SS.

Definindo por

$$\omega_1 = p_1 - p_0\pi_s(k_1; \mu, \phi)$$

e

$$\omega_2 = p_2 - p_0\pi_s(k_2; \mu, \phi),$$

temos que a FMP $\pi_{k-IS}(z; \mu, \phi, \mathbf{p})$ escrita como:

$$\begin{aligned}
\pi_{k-IS}(z; \mu, \phi, \boldsymbol{\omega}) &= (1 - \omega_1)I_{\{k_1\}}(z) + (1 - \omega_2)I_{\{k_2\}}(z) + \\
&\quad \left((p_1 + p_2 - 1) - p_0\pi_s(k_1; \mu, \phi) - p_0\pi_s(k_2; \mu, \phi) \right) \pi_{k-SS}(z; \mu, \phi) \\
&= (1 - \omega_1)I_{\{k_1\}}(z) + (1 - \omega_2)I_{\{k_2\}}(z) + (\omega_1 + \omega_2 - 1)\pi_{k-SS}(z; \mu, \phi) \\
&= (1 - \omega_1)I_{\{k_1\}}(z) + (1 - \omega_2)I_{\{k_2\}}(z) + \omega_0\pi_{k-SS}(z; \mu, \phi),
\end{aligned}$$

em que $\boldsymbol{\omega} = (\omega_0, \omega_1, \omega_2)$, sendo $\omega_0 = \omega_1 + \omega_2 - 1$ e sob a condição $\omega_0 + \sum_{i=1}^2 (1 - \omega_i) = 1$.

Uma vez que $0 < p_0 < 1$, com algumas manipulações algébricas, segue que

$$p_1 - \pi_s(k_1; \mu, \phi) < \omega_1 < p_1$$

e

$$p_2 - \pi_s(k_2; \mu, \phi) < \omega_2 < p_2,$$

espaços paramétricos de ω_1 e ω_2 , respectivamente.

Portanto, $\pi_{k-IS}(z; \mu, \phi, \boldsymbol{\omega})$ é a versão *hurdle* da distribuição \mathbf{k} -IS. \square

Uma das vantagens em utilizar a versão *hurdle* é poder interpretar $(1 - \omega_1)$ como a probabilidade do evento $z = k_1$, $(1 - \omega_2)$ como a probabilidade de $z = k_2$ e $(\omega_1 + \omega_2 - 1)\pi_{k-SS}(z; \mu, \phi)$ como a probabilidade de ocorrer uma observação z diferente de k_1 e k_2 , isto é, $\{\forall z \in \mathbb{Z} \mid z \neq k_1 \text{ e } z \neq k_2\}$.

MODELO DE REGRESSÃO SKELLAM k -MODIFICADO

4.1 O Modelo Skellam k -Modificado

Seja Y uma variável aleatória tal que $Y \sim k\text{-MS}(\mu, \phi, p)$. Considere dois conjuntos de variáveis explicativas (covariáveis), x_1, \dots, x_{q_1} e v_1, \dots, v_{q_2} , associados respectivamente aos parâmetros μ e p . O modelo de regressão para conjuntos de dados com valores inteiros, assumindo distribuição k -MS, é construído reescrevendo a distribuição (3.1) como

$$\pi_{k\text{-MS}}(y; \mu(\mathbf{x}), \phi, p(\mathbf{v})) = (1 - p(\mathbf{v}))I_{\{k\}}(y) + p(\mathbf{v})\pi_s(y; \mu(\mathbf{x}), \phi), \quad y \in \mathbb{Z}, \quad (4.1)$$

no qual a condição p dada em (3.2) é reescrita para $p(\mathbf{v})$ como

$$0 \leq p(\mathbf{v}) \leq \frac{1}{1 - \pi_s(k; \mu(\mathbf{x}), \phi)}, \quad (4.2)$$

em que $\mathbf{x} = (1, x_1, \dots, x_{q_1})$ e $\mathbf{v} = (1, v_1, \dots, v_{q_2})$ denotam vetores de covariáveis relacionadas aos parâmetros da regressão.

As covariáveis para os parâmetros μ e p foram incorporadas no modelo por meio de funções de ligação apropriadas, estritamente monótonas e duplamente diferenciáveis: $h_\mu(\mu)$ e $h_p(p)$. Para o parâmetro μ , consideramos a função de ligação identidade, $h_\mu(\mu) = \mu(\mathbf{x}) = \mathbf{x}\boldsymbol{\beta}$, que relaciona μ ao preditor linear com um vetor de parâmetros $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{q_1})^\top$ e o vetor de covariáveis $\mathbf{x} = (1, x_1, \dots, x_{q_1})$. E para o parâmetro p , consideramos uma função de ligação que garanta a restrição definida na Equação (4.2). Neste trabalho consideramos $h_p(p)$ por:

$$h_p(p) = \log \left(\frac{p(\mathbf{v})}{\frac{1}{1 - \pi_s(k; \mu(\mathbf{x}), \phi)} - p(\mathbf{v})} \right) = \mathbf{v}\boldsymbol{\alpha}, \quad (4.3)$$

que relaciona p ao preditor linear com um outro vetor de parâmetros $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_{q_2})^\top$ e o vetor de covariáveis $\mathbf{v} = (1, v_1, \dots, v_{q_2})$.

Desse modo, temos que:

$$\mu(\mathbf{x}) = h_\mu^{-1}(\mathbf{x}\boldsymbol{\beta}) \quad \text{e} \quad p(\mathbf{v}) = h_p^{-1}(\mathbf{v}\boldsymbol{\alpha}).$$

O modelo de regressão k -MS possui $(q_1 + q_2 + 3)$ parâmetros desconhecidos a serem estimados, correspondentes aos vetores $\boldsymbol{\beta}$ e $\boldsymbol{\alpha}$ e o parâmetro ϕ .

4.1.1 Modelo k -MS e sua Versão Hurdle

Com base na versão *hurdle* da distribuição k -MS apresentada em (3.6), o modelo de regressão k -MS é escrito como:

$$\pi_{k-MS}(y; \mu(\mathbf{x}), \phi, \omega(\mathbf{v})) = (1 - \omega(\mathbf{v}))I_{\{k\}}(y) + \omega(\mathbf{v})\pi_{k-SS}(y; \mu(\mathbf{x}), \phi), \quad y \in \mathbb{Z}, \quad (4.4)$$

em que $0 \leq \omega(\mathbf{v}) \leq 1$.

Nesta parametrização, o vetor de parâmetros $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_{q_2})^\top$ e o vetor de covariáveis $\mathbf{v} = (1, v_1, \dots, v_{q_2})$ estão associados ao parâmetro ω .

Uma vez que $\omega(\mathbf{v})$ e $p(\mathbf{v})$ estão relacionados por $\omega(\mathbf{v}) = p(\mathbf{v})(1 - \pi_s(k; \mu(\mathbf{x}), \phi))$, ao explicitar $p(\mathbf{v})$ na Equação (4.3) obtemos:

$$p(\mathbf{v}) = \left(\frac{e^{\mathbf{v}\boldsymbol{\alpha}}}{1 + e^{\mathbf{v}\boldsymbol{\alpha}}} \right) \left(\frac{1}{1 - \pi_s(k; \mu(\mathbf{x}), \phi)} \right).$$

Assim, temos:

$$\begin{aligned} \omega(\mathbf{v}) &= p(\mathbf{v})(1 - \pi_s(k; \mu(\mathbf{x}), \phi)) \\ &= \left(\frac{e^{\mathbf{v}\boldsymbol{\alpha}}}{1 + e^{\mathbf{v}\boldsymbol{\alpha}}} \right) \left(\frac{1}{1 - \pi_s(k; \mu(\mathbf{x}), \phi)} \right) (1 - \pi_s(k; \mu(\mathbf{x}), \phi)) \\ &= \frac{e^{\mathbf{v}\boldsymbol{\alpha}}}{1 + e^{\mathbf{v}\boldsymbol{\alpha}}}, \end{aligned}$$

que equivale a $h_\omega^{-1}(\mathbf{v}\boldsymbol{\alpha})$.

Logo, a função de ligação para ω (associada com a função de ligação proposta para p , $h_p(p)$) é:

$$h_\omega(\omega) = \log \left(\frac{\omega(\mathbf{v})}{1 - \omega(\mathbf{v})} \right) = \mathbf{v}\boldsymbol{\alpha},$$

conhecida como função de ligação *logit*.

Podemos notar que a função de ligação *logit* para o parâmetro $\omega(\mathbf{v})$ exclui dois casos específicos da variável aleatória Y , a saber, quando Y tem uma distribuição degenerada em k e quando Y tem uma distribuição k -SS.

Outras funções de ligação podem ser consideradas para $\omega(\mathbf{v})$ (ver [Huayanay et al. \(2019\)](#)) e, nestes casos, é possível encontrar as funções de ligação correspondentes para $p(\mathbf{v})$.

Ao contrário dos modelos de misturas tradicionais, o modelo (4.4) pode ser ajustado tanto para conjuntos de dados k -inflacionados quanto para conjuntos de dados k -deflacionados. Além disso, essa parametrização leva a uma separação do modelo em duas partes (verossimilhança separável), acarretando a ortogonalidade entre $\omega(\mathbf{v})$ e os parâmetros $\mu(\mathbf{x})$ e ϕ , possibilitando estimá-los independentemente.

4.1.2 Função de Verossimilhança

Seja $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$ um vetor de n variáveis aleatórias independentes, com $Y_i \sim k\text{-MS}(\mu_i, \phi_i, \omega_i)$, para todo $i = 1, 2, \dots, n$, e seja $\mathbf{y} = (y_1, y_2, \dots, y_n)$ o vetor de observações associados a \mathbf{Y} . Considere $\mathbf{x}_i = (1, x_{1i}, \dots, x_{q_1i})$ e $\mathbf{v}_i = (1, v_{1i}, \dots, v_{q_2i})$ os vetores de covariáveis associados a cada y_i e, conseqüentemente, $\mu_i = \mu(\mathbf{x}_i)$ e $\omega_i = \omega(\mathbf{v}_i)$. Denotamos por $\mathcal{D} = \{y_i, \mathbf{x}_i, \mathbf{v}_i : \forall i = 1, \dots, n\}$ as informações relacionadas a todos os indivíduos.

Assim, considerando o modelo (4.4), a função de verossimilhança associada a \mathbf{y} é dada por:

$$\begin{aligned} \mathcal{L}(\mu(\mathbf{x}_i), \phi_i, \omega(\mathbf{v}_i); \mathcal{D}) &= \prod_{i=1}^n \left\{ (1 - \omega(\mathbf{v}_i)) I_{\{k\}}(y_i) + \omega(\mathbf{v}_i) \pi_{k\text{-SS}}(y_i; \mu(\mathbf{x}_i), \phi) \right\} \\ &= \prod_{i=1}^n \left\{ (1 - \omega(\mathbf{v}_i))^{I_{\{k\}}(y_i)} \left(\frac{\omega(\mathbf{v}_i) \pi_S(y_i; \mu(\mathbf{x}_i), \phi)}{1 - \pi_S(k; \mu(\mathbf{x}_i), \phi)} \right)^{1 - I_{\{k\}}(y_i)} \right\}. \end{aligned}$$

O logaritmo natural da função de verossimilhança é

$$\begin{aligned} \ell(\mu(\mathbf{x}_i), \phi_i, \omega(\mathbf{v}_i); \mathcal{D}) &= \sum_{i=1}^n \left\{ I_{\{k\}}(y_i) \log(1 - \omega(\mathbf{v}_i)) + (1 - I_{\{k\}}(y_i)) \log(\omega(\mathbf{v}_i)) \right\} + \\ &\quad \sum_{i=1}^n (1 - I_{\{k\}}(y_i)) \log \left(\frac{\pi_S(y_i; \mu(\mathbf{x}_i), \phi)}{1 - \pi_S(k; \mu(\mathbf{x}_i), \phi)} \right). \end{aligned} \quad (4.5)$$

Uma vez que $\mu(\mathbf{x}_i) = \mathbf{x}_i \boldsymbol{\beta}$ e $\omega(\mathbf{v}_i) = \frac{e^{\mathbf{v}_i \boldsymbol{\alpha}}}{1 + e^{\mathbf{v}_i \boldsymbol{\alpha}}}$, e substituindo em (4.5), temos:

$$\begin{aligned} \ell(\boldsymbol{\beta}, \phi, \boldsymbol{\alpha}; \mathcal{D}) &= \sum_{i=1}^n \left\{ I_{\{k\}}(y_i) \log \left(1 - \frac{e^{\mathbf{v}_i \boldsymbol{\alpha}}}{1 + e^{\mathbf{v}_i \boldsymbol{\alpha}}} \right) + (1 - I_{\{k\}}(y_i)) \log \left(\frac{e^{\mathbf{v}_i \boldsymbol{\alpha}}}{1 + e^{\mathbf{v}_i \boldsymbol{\alpha}}} \right) \right\} + \\ &\quad \sum_{i=1}^n (1 - I_{\{k\}}(y_i)) \log \left(\frac{g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^{y_i} f(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{1 - g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)} \right) \end{aligned}$$

$$\begin{aligned}
\ell(\boldsymbol{\beta}, \phi, \boldsymbol{\alpha}; \mathcal{D}) &= \sum_{i=1}^n \left\{ (1 - I_{\{k\}}(y_i)) \mathbf{v}_i \boldsymbol{\alpha} - \log(1 + e^{\mathbf{v}_i \boldsymbol{\alpha}}) \right\} + \\
&\quad \sum_{i=1}^n (1 - I_{\{k\}}(y_i)) \left\{ y_i \log(g(\mathbf{x}_i \boldsymbol{\beta}, \phi)) + \log \left(\frac{f(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{1 - g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)} \right) \right\} \\
&= \ell_1(\boldsymbol{\alpha}; \mathcal{D}) + \ell_2(\boldsymbol{\beta}, \phi; \mathcal{D}), \tag{4.6}
\end{aligned}$$

em que

$$\begin{aligned}
\ell_1(\boldsymbol{\alpha}; \mathcal{D}) &= \sum_{i=1}^n \left\{ (1 - I_{\{k\}}(y_i)) \mathbf{v}_i \boldsymbol{\alpha} - \log(1 + e^{\mathbf{v}_i \boldsymbol{\alpha}}) \right\}; \text{ e} \\
\ell_2(\boldsymbol{\beta}, \phi; \mathcal{D}) &= \sum_{i=1}^n (1 - I_{\{k\}}(y_i)) \left\{ y_i \log(g(\mathbf{x}_i \boldsymbol{\beta}, \phi)) + \log \left(\frac{f(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{1 - g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)} \right) \right\}.
\end{aligned}$$

Diferentes abordagens podem ser adotadas para a estimação dos parâmetros. Para o modelo k -MS, empregamos abordagem clássica e bayesiana. A estimação com enfoque bayesiano foi realizada considerando um procedimento Monte Carlo em Cadeia de Markov (MCMC, *Monte Carlo Markov Chain*), mais especificamente, o método Hamiltoniano Monte Carlo (HMC), descrito no item 4.3.2. Para este procedimento, é necessário encontrar o gradiente vetorial do logaritmo natural da função de verossimilhança, isto é, o vetor escore $\mathbf{U} = (\mathbf{U}_{\boldsymbol{\beta}}, U_{\phi}, \mathbf{U}_{\boldsymbol{\alpha}})^{\top}$, com $\mathbf{U}_{\boldsymbol{\beta}} = (U_{\beta_0}, U_{\beta_1}, \dots, U_{\beta_{q_1}})$ e $\mathbf{U}_{\boldsymbol{\alpha}} = (U_{\alpha_0}, U_{\alpha_1}, \dots, U_{\alpha_{q_2}})$.

Assim, os elementos do vetor escore relacionados a $\boldsymbol{\beta}$, $\mathbf{U}_{\boldsymbol{\beta}} = (U_{\beta_0}, U_{\beta_1}, \dots, U_{\beta_{q_1}})$, são dados por:

$$\begin{aligned}
U_{\beta_j} = \frac{\partial \ell_2(\boldsymbol{\beta}, \phi; \mathcal{D})}{\partial \beta_j} &= \sum_{i=1}^n (1 - I_{\{k\}}(y_i)) \left\{ y_i \frac{g_{\beta_j}(\mathbf{x}_i \boldsymbol{\beta}, \phi)}{g(\mathbf{x}_i \boldsymbol{\beta}, \phi)} + \frac{f_{\beta_j}(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{f(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)} + \right. \\
&\quad \left. \frac{g_{\beta_j}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi) + g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f_{\beta_j}(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{1 - g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)} \right\}, \quad \forall j = 0, 1, \dots, q_1,
\end{aligned}$$

em que $g_{\beta_j}(\mathbf{x}_i \boldsymbol{\beta}, \phi)$ (bem como $g_{\beta_j}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k$) e $f_{\beta_j}(*; \mathbf{x}_i \boldsymbol{\beta}, \phi)$ ($* = k, y_i$) denotam, respectivamente, as derivadas das funções $g(\mathbf{x}_i \boldsymbol{\beta}, \phi)$ e $f(*; \mathbf{x}_i \boldsymbol{\beta}, \phi)$ em relação a β_j .

O elemento do vetor escore relacionado a ϕ , U_{ϕ} , é dado por:

$$\begin{aligned}
U_{\phi} = \frac{\partial \ell_2(\boldsymbol{\beta}, \phi; \mathcal{D})}{\partial \phi} &= \sum_{i=1}^n (1 - I_{\{k\}}(y_i)) \left\{ y_i \frac{g_{\phi}(\mathbf{x}_i \boldsymbol{\beta}, \phi)}{g(\mathbf{x}_i \boldsymbol{\beta}, \phi)} + \frac{f_{\phi}(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{f(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)} + \right. \\
&\quad \left. \frac{g_{\phi}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi) + g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f_{\phi}(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{1 - g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)} \right\},
\end{aligned}$$

em que $g_{\phi}(\mathbf{x}_i \boldsymbol{\beta}, \phi)$ (bem como $g_{\phi}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k$) e $f_{\phi}(*; \mathbf{x}_i \boldsymbol{\beta}, \phi)$ ($* = k, y_i$) denotam, respectivamente, as derivadas das funções $g(\mathbf{x}_i \boldsymbol{\beta}, \phi)$ e $f(*; \mathbf{x}_i \boldsymbol{\beta}, \phi)$ em relação a ϕ .

Por fim, os elementos do vetor escore relacionados a $\boldsymbol{\alpha}$, $\mathbf{U}_{\boldsymbol{\alpha}} = (U_{\alpha_0}, U_{\alpha_1}, \dots, U_{\alpha_{q_2}})$, são

dados por:

$$U_{\alpha_j} = \frac{\partial \ell_1(\boldsymbol{\alpha}; \mathcal{D})}{\partial \alpha_j} = \sum_{i=1}^n \left\{ (1 - I_{\{k\}}(y_i)) v_{ij} - \frac{e^{v_i \boldsymbol{\alpha}}}{1 + e^{v_i \boldsymbol{\alpha}}} v_{ij} \right\}, \forall j = 0, 1, \dots, q_2.$$

A Tabela 1 apresenta as derivadas das funções $g(\mathbf{x}_i \boldsymbol{\beta}, \phi)$ (bem como $g_\phi(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k$) e $f(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)$ do modelo k -MS com relação a β_j ($j = 0, 1, \dots, q_1$) e ϕ .

Tabela 1 – Derivadas das funções $g(\mathbf{x}_i \boldsymbol{\beta}, \phi)$ (bem como $g_\phi(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k$) e $f(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)$ do modelo k -MS.

$g(\mathbf{x}_i \boldsymbol{\beta}, \phi)$	$g_{\beta_j}(\mathbf{x}_i \boldsymbol{\beta}, \phi)$	$g_\phi(\mathbf{x}_i \boldsymbol{\beta}, \phi)$
$\left(\frac{\phi + \mathbf{x}_i \boldsymbol{\beta}}{\phi - \mathbf{x}_i \boldsymbol{\beta}} \right)^{\frac{1}{2}}$	$\frac{\phi x_{ij}}{\sqrt{(\phi + \mathbf{x}_i \boldsymbol{\beta})(\phi - \mathbf{x}_i \boldsymbol{\beta})^3}}$	$-\frac{\mathbf{x}_i \boldsymbol{\beta}}{\sqrt{(\phi + \mathbf{x}_i \boldsymbol{\beta})(\phi - \mathbf{x}_i \boldsymbol{\beta})^3}}$
$g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k$	$g_{\beta_j}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k$	$g_\phi(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k$
$\left(\frac{\phi + \mathbf{x}_i \boldsymbol{\beta}}{\phi - \mathbf{x}_i \boldsymbol{\beta}} \right)^{\frac{k}{2}}$	$\frac{k \phi x_{ij} \sqrt{(\phi + \mathbf{x}_i \boldsymbol{\beta})^{k-2}}}{\sqrt{(\phi - \mathbf{x}_i \boldsymbol{\beta})^{k+2}}}$	$-\frac{k \mathbf{x}_i \boldsymbol{\beta} \sqrt{(\phi + \mathbf{x}_i \boldsymbol{\beta})^{k-2}}}{\sqrt{(\phi - \mathbf{x}_i \boldsymbol{\beta})^{k+2}}}$
$f(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)$	$f_{\beta_j}(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)$	$f_\phi(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)$
$e^{-\phi \mathcal{J}_y} \left(\sqrt{\phi^2 - (\mathbf{x}_i \boldsymbol{\beta})^2} \right)$	$\frac{-\mathbf{x}_i \boldsymbol{\beta} e^{-\phi \mathcal{J}'_{y_i}(\sqrt{\phi^2 - (\mathbf{x}_i \boldsymbol{\beta})^2)}}{\sqrt{\phi^2 - (\mathbf{x}_i \boldsymbol{\beta})^2}} x_{ij}$	$e^{-\phi} \left(\frac{\phi \mathcal{J}'_{y_i}(\sqrt{\phi^2 - (\mathbf{x}_i \boldsymbol{\beta})^2}}{\sqrt{\phi^2 - (\mathbf{x}_i \boldsymbol{\beta})^2}} - \mathcal{J}_{y_i} \sqrt{\phi^2 - (\mathbf{x}_i \boldsymbol{\beta})^2} \right)$

$$\text{em que } \mathcal{J}'_y = \frac{1}{2} (\mathcal{J}_{y-1}(\xi) + \mathcal{J}_{y+1}(\xi)).$$

É possível obter a matriz hessiana (\mathbf{H}) ao derivar o vetor escore em relação a cada parâmetro. O bloco da matriz \mathbf{H} correspondentes as derivadas dos parâmetros $\boldsymbol{\beta}$ e ϕ , de dimensão $(q_1 + 2) \times (q_1 + 2)$, é formado pelos elementos

$$\begin{aligned} \frac{\partial^2 \ell_2(\boldsymbol{\beta}, \phi; \mathcal{D})}{\partial \beta_j \partial \beta_r} &= \sum_{i=1}^n (1 - I_{\{k\}}(y_i)) \left\{ y_i \frac{g_{\beta_j \beta_r}(\mathbf{x}_i \boldsymbol{\beta}, \phi)}{g(\mathbf{x}_i \boldsymbol{\beta}, \phi)} + \frac{f_{\beta_j \beta_r}(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{f(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)} + \right. \\ &\quad \frac{g_{\beta_j \beta_r}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi) + g_{\beta_j}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f_{\beta_r}(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{1 - g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)} + \\ &\quad \frac{g_{\beta_r}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f_{\beta_j}(k; \mathbf{x}_i \boldsymbol{\beta}, \phi) + g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f_{\beta_j \beta_r}(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{1 - g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)} + \\ &\quad \frac{(g_{\beta_j}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi) + g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f_{\beta_j}(k; \mathbf{x}_i \boldsymbol{\beta}, \phi))}{(1 - g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi))^2} \times \\ &\quad \left. \frac{(g_{\beta_r}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi) + g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f_{\beta_r}(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)) -}{\frac{f_{\beta_j \beta_r}(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi) f_{\beta_j}(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{f(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)^2} - y_i \frac{g_{\beta_j \beta_r}(\mathbf{x}_i \boldsymbol{\beta}, \phi) g_{\beta_j}(\mathbf{x}_i \boldsymbol{\beta}, \phi)}{g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^2}} \right\}, \forall r, j = 0, \dots, q_1; \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \ell_2(\boldsymbol{\beta}, \phi; \mathcal{D})}{\partial \phi \partial \beta_j} &= \sum_{i=1}^n (1 - I_{\{k\}}(y_i)) \left\{ y_i \frac{g_{\beta_j \phi}(\mathbf{x}_i \boldsymbol{\beta}, \phi)}{g(\mathbf{x}_i \boldsymbol{\beta}, \phi)} + \frac{f_{\beta_j \phi}(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{f(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)} + \right. \\ &\quad \frac{g_{\beta_j \phi}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi) + g_{\beta_j}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f_{\phi}(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{1 - g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)} + \\ &\quad \frac{g_{\phi}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f_{\beta_j}(k; \mathbf{x}_i \boldsymbol{\beta}, \phi) + g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f_{\beta_j \phi}(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{1 - g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)} + \\ &\quad \frac{(g_{\beta_j}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi) + g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f_{\beta_j}(k; \mathbf{x}_i \boldsymbol{\beta}, \phi))}{(1 - g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi))^2} \times \\ &\quad (g_{\phi}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi) + g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f_{\phi}(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)) - \\ &\quad \left. \frac{f_{\beta_j \phi}(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi) f_{\beta_j}(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{f(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)^2} - y_i \frac{g_{\beta_j \phi}(\mathbf{x}_i \boldsymbol{\beta}, \phi) g_{\beta_j}(\mathbf{x}_i \boldsymbol{\beta}, \phi)}{g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^2} \right\}, \forall j = 0, \dots, q_1; e \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \ell_2(\boldsymbol{\beta}, \phi; \mathcal{D})}{\partial \phi^2} &= \sum_{i=1}^n (1 - I_{\{k\}}(y_i)) \left\{ y_i \frac{g_{\phi^2}(\mathbf{x}_i \boldsymbol{\beta}, \phi)}{g(\mathbf{x}_i \boldsymbol{\beta}, \phi)} + \frac{f_{\phi^2}(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{f(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)} + \right. \\ &\quad \frac{g_{\phi^2}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi) + 2g_{\phi}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f_{\phi}(k; \mathbf{x}_i \boldsymbol{\beta}, \phi) + g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f_{\phi^2}(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)}{1 - g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi)} + \\ &\quad \frac{(g_{\phi}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi) + g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f_{\phi}(k; \mathbf{x}_i \boldsymbol{\beta}, \phi))^2}{(1 - g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^k f(k; \mathbf{x}_i \boldsymbol{\beta}, \phi))^2} - \\ &\quad \left. \frac{f_{\phi}(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)^2}{f(y_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)^2} - y_i \frac{g_{\phi}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^2}{g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^2} \right\}. \end{aligned}$$

O bloco da matriz \mathbf{H} correspondentes as derivadas dos parâmetros $\boldsymbol{\alpha}$, de dimensão $(q_2 + 1) \times (q_2 + 1)$, tem os elementos

$$\frac{\partial^2 \ell_1(\boldsymbol{\alpha}; \mathcal{D})}{\partial \alpha_j \partial \alpha_r} = \sum_{i=1}^n \left\{ -v_{ji} v_{ri} \frac{e^{v_i \boldsymbol{\alpha}}}{(1 + e^{v_i \boldsymbol{\alpha}})^2} \right\}, \forall j, r = 0, \dots, q_2.$$

4.2 Estimação dos Parâmetros: Abordagem Clássica

Para estimação e inferência dos $q_1 + q_2 + 3$ parâmetros do modelo k -MS, correspondentes aos vetores $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{q_1})^\top$, $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_{q_2})^\top$ e o parâmetro ϕ , apresentamos aqui uma abordagem clássica. A partir das estimativas de $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$ e ϕ , obtém-se as estimativas de $\boldsymbol{\omega}(\mathbf{v})$ e $\boldsymbol{\mu}(\mathbf{x})$. O tipo de k -modificação presente no conjunto de dados é obtido diretamente do cálculo da estimativa do parâmetro $p(\mathbf{v})$ por $\hat{p}(\mathbf{v}) = \frac{\hat{\omega}(\mathbf{v})}{1 - \pi_s(k; \hat{\boldsymbol{\mu}}(\mathbf{x}), \hat{\phi})}$.

Assim, nesta sessão, abordamos o procedimento de estimação para os parâmetros do modelo k -MS via método de máxima verossimilhança. O algoritmo utilizado foi o *Bound*

Optimization by Quadratic Approximation (BOBYQA). A performance dos estimadores foi avaliada por meio de um estudo de simulação. A qualidade do ajuste no procedimento de estimação foi observada por meio de conjuntos de dados artificiais e a aplicabilidade do modelo é representada com a análise de um conjunto de dados reais.

4.2.1 Algoritmo BOBYQA

O algoritmo BOBYQA faz a minimização usando uma aproximação quadrática construída iterativamente para a função objeto, sem a necessidade de fornecer derivadas (POWELL *et al.*, 2009). Optou-se por esse algoritmo devido ao seu desempenho favorável relatado na literatura e à possibilidade da restrição nos limites dos parâmetros, uma necessidade decorrente da condição imposta ao parâmetro ϕ . No *Software R*, o algoritmo está implementado dentro do pacote “nloptr” (YPMA *et al.*, 2018), e para utilizá-lo basta especificar o algoritmo "NLOPT_LN_BOBYQA" nos argumentos da função `nloptr`.

BOBYQA é um pacote de sub-rotinas Fortran que busca o menor valor de uma função $F(\mathbf{x})$, $\mathbf{x} \in R^n$, sujeita aos limites simples

$$a_i \leq x_i \leq b_i, \quad i = 1, 2, \dots, n, \quad (4.7)$$

nas componentes de \mathbf{x} . O usuário define a função por outra sub-rotina que retorna o valor $F(\mathbf{x})$ para qualquer \mathbf{x} em R^n . Nenhuma derivada de F é necessária.

O método de BOBYQA é iterativo, sendo k o número da iteração e n o número de variáveis. Além disso, m é o número de condições de interpolação que são impostas em uma aproximação quadrática $Q_k(\mathbf{x})$, $\mathbf{x} \in R^n$, para $F(\mathbf{x})$, $\mathbf{x} \in R^n$. A aproximação é obtida no início da k -ésima iteração, as equações de interpolação têm a forma

$$Q_k(\mathbf{y}_j) = F(\mathbf{y}_j), \quad j = 1, 2, \dots, m,$$

em que m é um número inteiro do intervalo $[n + 2, \frac{1}{2}(n + 1)(n + 2)]$, escolhido pelo usuário do software.

Seja \mathbf{x}_k o ponto do conjunto $\{\mathbf{y}_j : j = 1, 2, \dots, m\}$ que tem a propriedade

$$F(\mathbf{x}_k) = \min\{F(\mathbf{y}_j) : j = 1, 2, \dots, m\}$$

, quaisquer empates serão quebrados dando prioridade a uma avaliação anterior do menor valor da função $F(\mathbf{x}_k)$. Um número positivo Δ_k , denominado “raio da região de confiança”, também é obtido no início da k -ésima iteração.

Se certas as condições forem alcançadas o término ocorrerá na k -ésima iteração. Caso contrário, um passo \mathbf{d}_k a partir de \mathbf{x}_k é construído tal que $\|\mathbf{d}_k\| \leq \Delta_k$ é válido, tal que $\mathbf{x} = \mathbf{x}_k + \mathbf{d}_k$ está dentro dos limites (4.7), e tal que $\mathbf{x}_k + \mathbf{d}_k$ não é um dos pontos de interpolação \mathbf{y}_j , $j = 1, 2, \dots, m$.

Em seguida, o novo valor da função $F(\mathbf{x}_k + \mathbf{d}_k)$ é calculado e um dos pontos de interpolação, digamos \mathbf{y}_t , é substituído por $\mathbf{x}_k + \mathbf{d}_k$, onde \mathbf{y}_t é diferente de \mathbf{x}_k . Segue-se que \mathbf{x}_{k+1} é definido pela fórmula

$$\mathbf{x}_{k+1} = \begin{cases} \mathbf{x}_k, & \text{se } F(\mathbf{x}_k + \mathbf{d}_k) \geq F(\mathbf{x}_k) \\ \mathbf{x}_k + \mathbf{d}_k, & \text{se } F(\mathbf{x}_k + \mathbf{d}_k) < F(\mathbf{x}_k) \end{cases}$$

Além disso, Δ_{k+1} e Q_{k+1} são gerados para a próxima iteração, Q_{k+1} estando sujeito às restrições

$$Q_{k+1}(\hat{\mathbf{y}}_j) = F(\hat{\mathbf{y}}_j), \quad j = 1, 2, \dots, m$$

nos novos pontos de interpolação por

$$\hat{\mathbf{y}}_j = \begin{cases} \mathbf{y}_j, & \text{para } j \neq t \\ \mathbf{x}_k + \mathbf{d}_k, & \text{para } j = t \end{cases} \quad j = 1, 2, \dots, m$$

.

4.2.2 Critérios de Seleção de Modelos

Os critérios de seleção de modelos são ferramentas usadas para comparar e selecionar modelos ajustados. A seguir é apresentada uma breve descrição de alguns critérios clássicos que serão utilizados para a seleção de modelos.

4.2.2.1 Critério de Informação de Akaike (AIC)

O Critério de Informação de Akaike (AIC), proposto por [Akaike \(1974\)](#), tem seu cálculo dado por:

$$AIC = -2\ell(\hat{\boldsymbol{\theta}}; \mathcal{D}) + 2c,$$

em que $\ell(\boldsymbol{\theta}; \mathcal{D})$ é o valor do logaritmo natural da função de verossimilhança avaliada nas estimativas dos parâmetros $\boldsymbol{\theta}$ e c é o número de parâmetros desconhecidos no modelo.

O AIC penaliza modelos mais complexos, adicionando um termo que cresce com o número de parâmetros. Modelo com menor valor de AIC é considerado o melhor, pois indica um bom ajuste aos dados com um número mínimo de parâmetros.

4.2.2.2 Critério de Informação Bayesiano (BIC)

O Critério de Informação Bayesiano (BIC) foi proposto por [Schwarz \(1978\)](#) e é dado por:

$$BIC = -2\ell(\hat{\boldsymbol{\theta}}; \mathcal{D}) + c \ln(n),$$

sendo n é o tamanho da amostra.

O BIC penaliza modelos mais complexos de forma mais severa do que o AIC, pois o termo de penalizaco ($c \ln(n)$) é multiplicado pelo logaritmo do tamanho da amostra. O BIC favorece modelos mais simples quando o tamanho da amostra é grande.

Em resumo, o AIC e o BIC são critérios que ajudam na escolha do melhor modelo, equilibrando a sua complexidade (número de parâmetros) com a sua capacidade preditiva. Modelos com valores menores desses critérios são preferíveis.

4.2.3 Estudo de Simulaco: Avaliaco da Performance dos Estimadores Clássicos

Para o estudo de simulaco foram considerados conjuntos de dados com diferentes modificaces na frequênciam de k , com o objetivo de caracterizar conjuntos de dados k -inflacionados e k -deflacionados. Para cada cenário, foram gerados $N = 100$ conjuntos de dados do modelo Skellam k -Modificado, cada um com tamanhos amostrais de $n = 100$ e $n = 200$. Além disso, foram consideradas n observaces da variável explicativa X , com $X \sim ZTN(0, 1)$ (ZTN, Normal Zero-Truncada). Os valores de k e dos parâmetros considerados para a geraco nos números pseudos aleatórios com diferentes características (k -inflaco e k -deflaco) são apresentados na Tabela 2.

Tabela 2 – Valores de k e dos parâmetros considerados no estudo de simulaco.

k -Modificaco	k	Valores reais				
		α_0	α_1	β_0	β_1	ϕ
k -Inflaco	-2	2,00	-3,50	3,00	-2,00	4,50
	0	2,00	-3,00	6,00	-2,50	7,50
	4	1,50	-3,00	7,00	-5,00	8,50
k -Deflaco	-2	2,00	1,75	-4,00	1,50	5,50
	0	6,00	4,00	-5,00	2,00	6,50
	4	6,00	2,00	-5,00	7,00	15,50

Para cada conjunto gerado, o modelo k -MS foi ajustado considerando o método de máxima verossimilhança. Com as estimativas obtidas, calculamos os intervalos de confiança via procedimento *bootstrap* não paramétrico considerando 5000 réplicas, obtendo assim os intervalos *bootstrap* com 95% de confiança. De posse destes intervalos, avaliamos as probabilidades de cobertura (PC) para cada parâmetro nos diferentes casos.

O desempenho dos estimadores dos parâmetros do modelo foi avaliado por meio do cálculo das medidas do erro o quadrático médio (EQM), da variância (Var) e do viés absoluto médio (\mathcal{B}). Considerando que θ representa cada um dos parâmetros do modelo, assumindo α_0 , α_1 , β_0 , β_1 e ϕ , as medidas são obtidas da seguinte forma:

$$EQM(\hat{\theta}) = \frac{1}{N} \sum_{i=1}^N (\hat{\theta}_i - \theta)^2, \quad Var(\hat{\theta}) = \frac{1}{N} \sum_{i=1}^N (\hat{\theta}_i - \bar{\theta})^2, \quad \mathcal{B} = \frac{1}{N} \sum_{i=1}^N |\hat{\theta}_i - \theta|, \text{ em que}$$

$\hat{\theta}$ representa o estimador do parâmetro θ .

Já a probabilidade de cobertura é calculada por $PC(\theta) = \frac{1}{N} \sum_{i=1}^N \delta_{\theta}^i$, em que δ_{θ}^i assume 1 se o i -ésimo intervalo de confiança contém o valor verdadeiro do parâmetro θ e 0 caso contrário.

4.2.3.1 Conjuntos de Dados k -Inflacionados

Neste cenário os dados foram gerados de tal forma que os conjuntos de dados fossem caracterizados em k -inflacionados. A Tabela 3 apresenta as medidas do erro quadrático médio, da variância, do viés absoluto médio e da probabilidade de cobertura (em %) dos estimadores para os parâmetros do modelo k -IS, considerando diferentes valores de n e de pontos de modificação k .

Tabela 3 – Avaliação dos estimadores dos parâmetros do modelo k -IS via procedimento clássico, considerando diferentes valores de k e n .

k	Medidas	Estimadores					n	
		$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\phi}$		
-2	EQM	0,2196	0,5119	0,3302	1,2334	0,7362	100	
		0,0999	0,1885	0,1395	0,5331	0,5032	200	
	Var	0,2149	0,4937	0,3324	1,2458	0,7320	100	
		0,1003	0,1892	0,1406	0,5344	0,4965	200	
	B	0,0831	0,1522	0,0347	0,0065	0,1074	100	
		0,0236	0,0337	0,0148	0,0634	0,1081	200	
	PC (%)	95,00	94,00	90,00	92,00	91,00	100	
		97,00	98,00	95,00	93,00	93,00	200	
	0	EQM	0,3095	0,5391	0,5448	1,9669	1,6038	100
			0,1187	0,2415	0,2579	1,0895	1,1036	200
		Var	0,3023	0,5304	0,5499	1,9857	1,6039	100
			0,1115	0,2262	0,2570	1,0673	1,1107	200
B		0,1010	0,1181	0,0208	0,0321	0,1263	100	
		0,0914	0,1328	0,0588	0,1814	0,0631	200	
PC (%)		89,00	91,00	94,00	95,00	89,00	100	
		94,00	92,00	98,00	94,00	93,00	200	
4		EQM	0,2833	0,6750	0,4705	2,4312	2,0817	100
			0,1084	0,2411	0,5141	3,6089	1,8501	200
		Var	0,2566	0,5994	0,4659	2,4501	2,0608	100
			0,1075	0,2413	0,4974	3,3803	1,8283	200
	B	0,1710	0,2856	0,0964	0,0749	0,2037	100	
		0,0452	0,0473	0,1473	0,5122	0,2003	200	
	PC (%)	92,00	91,00	92,00	93,00	91,00	100	
		97,00	95,00	97,00	96,00	96,00	200	

De forma geral, o viés absoluto médio indica a tendência do estimador obter valores próximos do verdadeiro valor do parâmetro, retornando valores próximos de zero. Isto é, em média, as estimativas estão próximas do valor real dos parâmetros. Para o EQM e para a variância,

temos valores pequenos como resultado, indicando para um bom desempenho na estimação dos parâmetros. Quanto menores forem esses valores, mais precisas são as estimativas.

Os resultados obtidos das probabilidades de cobertura para os parâmetros ficaram próximos ao nível nominal de 95% de confiança, com valores levemente superiores para as amostras de tamanho de tamanho 200.

4.2.3.2 Conjuntos de Dados k -Deflacionados

Já para este cenário os dados foram gerados de maneira que os conjuntos de dados fossem caracterizados como k -deflacionados. Na Tabela 4 são apresentadas as medidas do erro quadrático médio, da variância, do viés absoluto médio e da probabilidade de cobertura (em %) dos estimadores para os parâmetros do modelo k -DS, considerando diferentes valores de n e de k .

Tabela 4 – Avaliação dos estimadores dos parâmetros do modelo k -DS via procedimento clássico, considerando diferentes valores de k e n .

k	Medidas	Estimadores					n	
		$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\phi}$		
-2	EQM	0,6445	3,4472	0,1453	0,1535	0,5157	100	
		0,2215	1,0750	0,0680	0,0646	0,2585	200	
	Var	0,6424	3,2057	0,1444	0,1543	0,4977	100	
		0,2236	0,9734	0,0687	0,0651	0,2519	200	
	\mathcal{B}	0,0926	0,5230	0,0479	0,0279	0,1517	100	
		0,0138	0,3336	0,0033	0,0106	0,0955	200	
	PC (%)	95,00	86,00	94,00	95,00	92,00	100	
		99,00	92,00	95,00	96,00	92,00	200	
	0	EQM	86,3790	25,7318	0,1867	0,1905	0,8141	100
			64,7306	28,5561	0,0884	0,07343	0,3768	200
		Var	30,8996	10,8774	0,1871	0,1915	0,7555	100
			30,8435	8,1072	0,0884	0,0736	0,3803	200
\mathcal{B}		7,4692	3,8682	0,0379	0,0304	0,2570	100	
		5,8477	4,5310	0,0297	0,0243	0,0159	200	
PC (%)		88,00	100,00	92,00	95,00	86,00	100	
		92,00	99,00	96,00	96,00	94,00	200	
4		EQM	86,3808	34,4117	0,3874	0,4627	2,1621	100
			53,1701	20,2813	0,1061	0,2126	1,0392	200
		Var	38,4085	12,9768	0,3912	0,4588	2,1789	100
			30,4116	9,2551	0,0621	0,2533	1,0321	200
	\mathcal{B}	6,9539	4,6438	0,0080	0,0921	0,0707	100	
		2,9403	1,5339	0,0000	0,0121	0,0152	200	
	PC (%)	92,00	93,00	96,00	95,00	94,00	100	
		100,00	98,00	98,00	97,00	96,00	200	

Para $k = -2$, temos bons resultados para EQM, Var e \mathcal{B} , indicando para um bom desempenho na estimação dos parâmetros. Nas probabilidades de cobertura, para os parâmetros

α_1 e ϕ ficaram com valores um pouco distantes do nível nominal de 95% de confiança. Já para a modificação nos pontos $k = 0$ e $k = 4$, os valores obtidos para o EQM , Var e \mathcal{B} dos parâmetros α_0 e α_1 são grandes, indicando que não obtivemos um bom desempenho no método de estimação. As probabilidades de cobertura se aproximam do nível nominal quando aumentado o tamanho da amostra, tendo como exceção o parâmetro α_0 para $k = 0$, apresentando PC de 92% apenas.

4.2.4 Aplicação: Conjuntos de Dados Artificiais

Como ilustração, conjuntos de dados artificiais k -modificados (k -inflacionados/ k -deflacionados) foram gerados com a partir do modelo k -MS, considerando tamanhos de amostra $n = 100$ e $n = 200$ e também considerando n observações de uma variável explicativa X , com $X \sim ZTN(0; 1)$. Para obtenção dos intervalos de confiança (IC) dos parâmetros do modelo, foi considerado o método *bootstrap* não paramétrico.

4.2.4.1 Aplicação do Modelo Skellam k -Inflacionado

Considerando o modelo k -MS, os valores dos vetores de parâmetros β e α e ao parâmetro ϕ foram atribuídos de tal forma que os conjuntos de dados gerados apresentem a característica de k -inflação ($0 < p(\mathbf{v}) < 1$) nos pontos $k = -2, 0$ e 4 (valores dos parâmetros conforme Tabela 2). Considerando a abordagem clássica, o modelo k -MS foi ajustado a cada conjunto de dados gerado, obtendo as estimativas dos parâmetros e o seus respectivos intervalos com 95% de confiança. A Tabela 5 apresenta os resultados deste estudo.

Tabela 5 – Estimativas (e intervalos de confiança de 95%) dos parâmetros do modelo k -IS, via procedimento clássico.

k	Estimativas (IC 95%)					n
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\phi}$	
-2	1,58 (1,54; 3,94)	-3,60 (-5,72; -2,59)	2,88 (1,55; 3,72)	-0,58 (-3,84; 0,50)	4,94 (3,14; 5,70)	100
	2,06 (1,21; 2,72)	-3,62 (-4,20; -2,12)	2,64 (2,70; 3,87)	1,96 (-3,07; -1,82)	5,19 (3,17; 5,52)	200
0	2,55 (1,56; 4,24)	-3,37 (-5,50; -2,10)	5,65 (4,33; 6,81)	-2,51 (-4,69; 0,10)	7,00 (6,10; 9,19)	100
	1,80 (1,19; 2,67)	-2,65 (-3,78; -1,83)	6,09 (5,09; 7,07)	-3,37 (-4,94; -1,86)	8,15 (6,08; 10,47)	200
4	1,35 (0,58; 2,29)	-3,02 (-4,89; -1,86)	7,09 (5,80; 8,36)	-5,58 (-7,87; -2,80)	8,25 (7,06; 11,09)	100
	1,59 (0,84; 2,86)	-3,51 (-5,19; -2,32)	7,20 (6,12; 8,19)	-4,92 (-6,42; -2,75)	8,06 (7,13; 10,54)	200

Nota-se que, em todos os casos apresentados, as estimativas estão próximas dos verdadeiros valores dos parâmetros, além destes também estarem contidos nos respectivos intervalos de confiança.

A Figura 1 apresenta graficamente as estimativas do parmetro p (linha superior) e da mdia ajustada (linha inferior) em funo da covarivel X para cada caso de tamanho $n = 200$ apresentado na Tabela 5, comparando com os valores reais. Para a mdia ajustada (e verdadeira), tambm pode ser visto os dados gerados de Y em funo da covarivel X . As proximidades entre as mdias verdadeiras e ajustadas indicam a boa qualidade das estimativas dos parmetros. Tambm  observado que as estimativas para o parmetro p so prximas de seus verdadeiros valores.

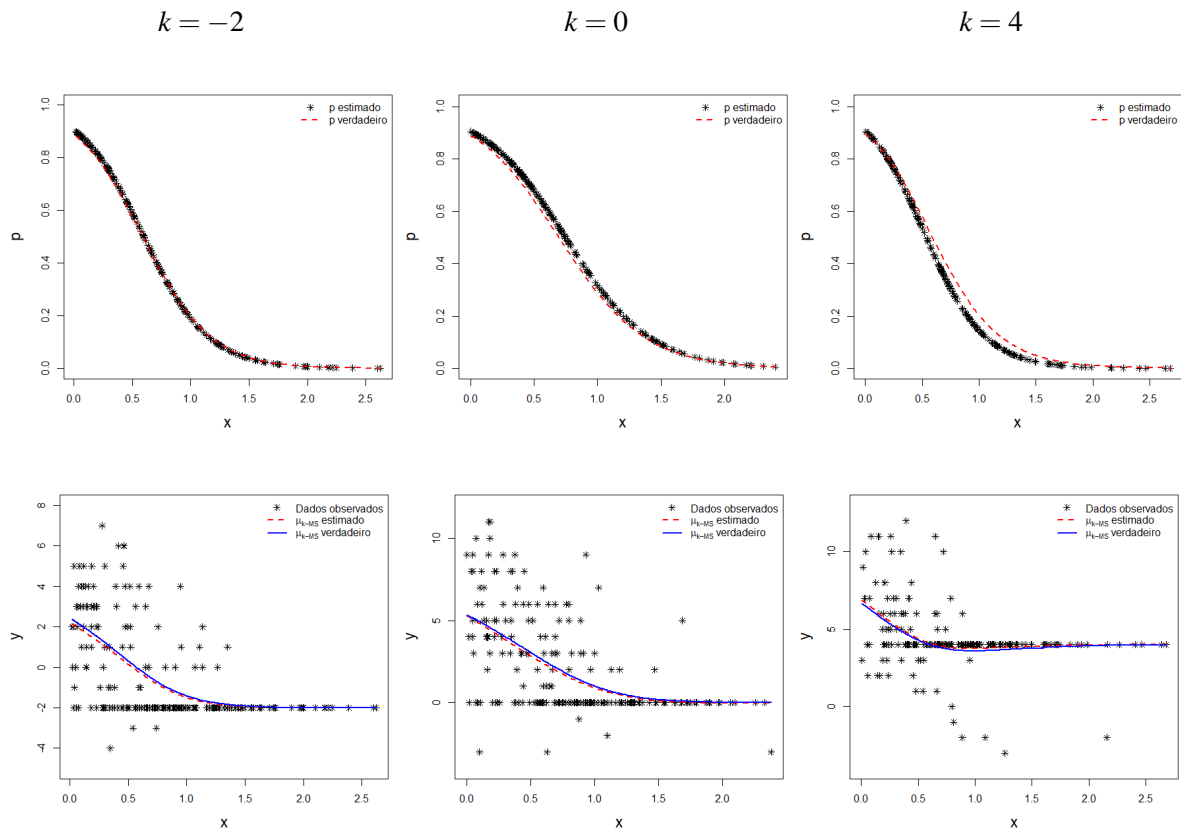


Figura 1 – Estimativas de p e μ_{k-MS} (juntamente com seus valores reais) obtidas a partir do ajuste do modelo k -IS, para $n = 200$.

4.2.4.2 Aplicao do Modelo Skellam k -Deflacionado

Para este caso foram considerados valores dos vetores de parmetros β e α e ao parmetro ϕ de tal forma que os conjuntos de dados obtidos a partir do modelo k -MS apresentem a caracterstica de k -deflao ($1 < p(\mathbf{v}) < 1/(1 - \pi_s(k; \mu(\mathbf{x}), \phi))$), com valores de $k = -2, 0$ e 4 (valores dos parmetros conforme Tabela 2). Similarmente, o modelo k -MS foi ajustado a cada conjunto de dados gerado via procedimento clssico. As estimativas dos parmetros e o seus respectivos intervalos *bootstrap* com 95% de confiana so apresentados na Tabela 6.  possvel notar que em todos os casos considerados as estimativas dos parmetros esto prximas dos verdadeiros valores, bem como estes esto includidos nos respectivos intervalos de confiana.

Tabela 6 – Estimativas (e intervalos de confiança de 95%) dos parâmetros do modelo k -DS, via procedimento clássico.

k	Estimativas (IC 95%)					n
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\phi}$	
-2	1,78 (-0,45;4,75)	2,99 (-0,01;15,84)	-4,37 (-5,24;-3,53)	1,60 (0,76;2,47)	6,29 (4,71;7,70)	100
	1,72 (0,63;2,72)	2,16 (1,32;7,91)	-3,92 (-5,03;-3,91)	1,54 (1,56;2,53)	5,68 (4,52;6,55)	200
0	6,20 (3,48;19,62)	12,61 (3,61;14,39)	-5,36 (-6,32;-4,42)	1,95 (1,01;2,90)	7,77 (5,84;9,44)	100
	6,15 (4,91;5,01)	11,71 (2,38;12,96)	-5,40 (-5,71;-4,55)	2,46 (1,44;2,65)	6,27 (5,20;7,63)	200
4	5,04 (3,76;20,61)	-0,53 (-0,97;12,78)	-5,73 (-7,10;-4,34)	7,23 (5,85;8,63)	15,41 (12,00;19,03)	100
	5,36 (4,57;21,27)	3,14 (0,75;11,49)	-5,04 (-6,49;-4,65)	7,66 (6,99;8,53)	15,46 (14,00;18,98)	200

A partir das estimativas dos parâmetros β , α e ϕ apresentadas na Tabela 6, obtém-se as estimativas de p e da média (μ_{k-MS}), as quais estão apresentadas graficamente na Figura 2. Ao analisar os gráficos desta Figura, nota-se a qualidade das estimativas ao comparar a proximidade das estimativas de p (linha superior da Figura) e de μ_{k-MS} (linha inferior da Figura) com os valores reais em cada caso.

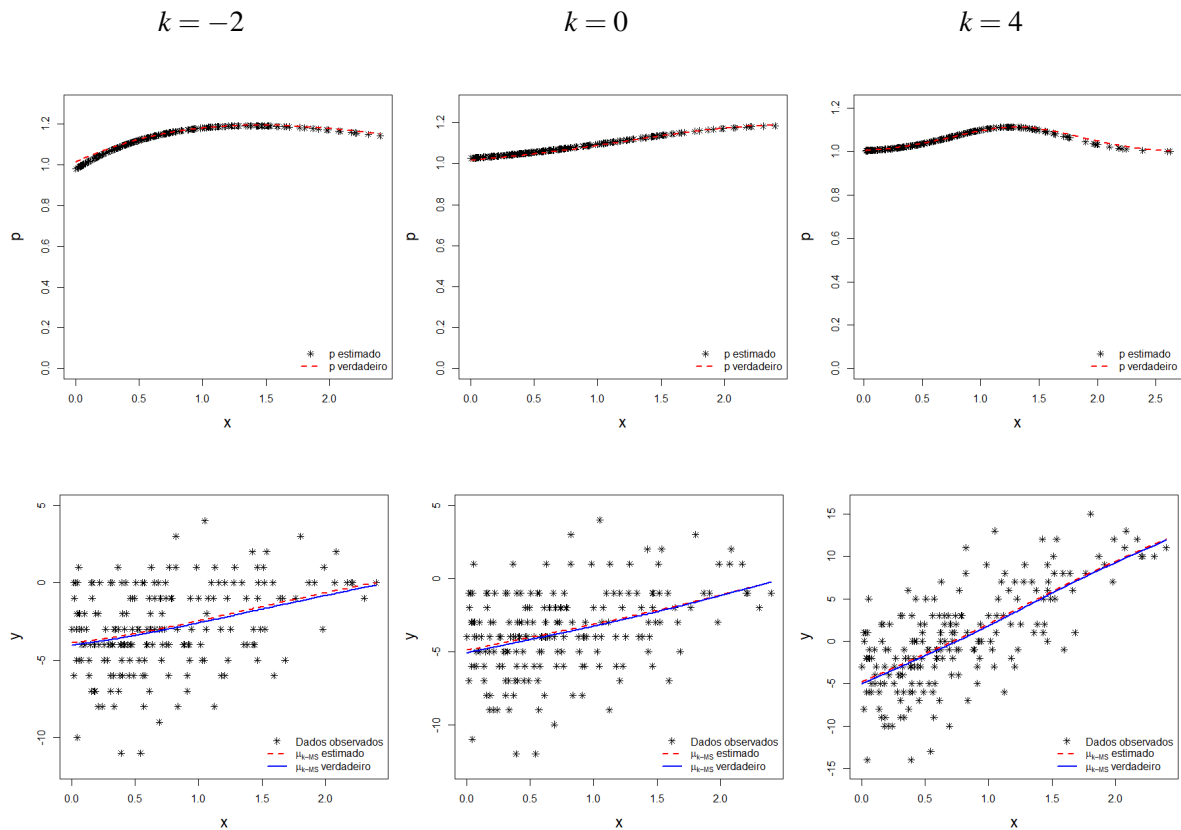


Figura 2 – Estimativas de p e μ_{k-MS} (juntamente com seus valores reais) obtidas a partir do ajuste do modelo k -DS, para $n = 200$.

4.2.4.3 Aplicao do Modelo Skellam

A fim de corroborar com a proposta do modelo k -MS e tambm apresentar outra forma de avaliar a qualidade das estimativas dos parmetros, principalmente no que se refere a estimaco do parmetro p , que  responsvel pela caracterizaco do conjunto de dados sobre presena ou ausncia de k -modificao, nesta Subseo prope-se gerar conjuntos de dados do modelo Skellam tradicional (ou do modelo k -MS fixando $p = 1$, e assim restringindo ao caso particular do modelo Skellam), considerar o ajuste o modelo generalizado (k -MS) para estes dados e avaliar se as estimativas de p so prximas de seu verdadeiro valor ($p = 1$). Dessa forma, consideramos dois tamanhos amostrais, $n = 100$ e $n = 200$, e trs conjuntos de valores de parmetros como apresentados na Tabela 7 (como neste cenrio $p = 1$, os valores do vetor de parmetros α no so definidos).

Tabela 7 – Valores dos parmetros para a gerao de conjuntos de dados do modelo Skellam.

Conjunto de Dados (CD)	Valores reais				
	α_0	α_1	β_0	β_1	ϕ
CD1	–	–	7,00	-5,00	8,50
CD2	–	–	-5,00	2,00	6,50
CD3	–	–	3,00	-2,00	4,50

As Tabelas 8 e 9 apresentam as estimativas dos parâmetros obtidas ao ajustar o modelo k -MS nos conjuntos de dados de tamanho $n = 100$ e $n = 200$, respectivamente, considerando diferentes pontos de k -modificação ($k = -2, 0$ e 4). Ao analisar as estimativas dos parâmetros β_0 , β_1 e ϕ , é possível notar que são bem próximas dos valores reais. Além disso, os intervalos de confiança não contêm o zero e englobam os verdadeiros valores dos parâmetros. Com relação as estimativas dos parâmetros α_0 e α_1 , temos que na maioria dos casos a covariável X não foi significativa, pois o valor zero está contido no intervalo de confiança para α_1 .

A partir dos resultados das estimativas dos parâmetros (especificamente do CD1 da Tabela 8), a Figura 3 apresenta graficamente as estimativas de p (linha superior) e da média μ_{k-MS} (linha inferior), juntamente com seus valores reais, para cada conjunto de dados. Para os gráficos de comparação das médias verdadeira e ajustadas, é possível observar também os dados gerados de Y em função da covariável X . Mais uma vez nota-se as proximidades entre as médias verdadeiras e ajustadas, bem como as proximidades dos valores verdadeiros e estimados de p , indicando o quão robusto é o modelo k -MS, principalmente pelo fato de poder ser ajustado a conjuntos de dados com diferentes características ou até mesmo desconhecendo o ponto de k -modificação e fazendo suposição deste.

Tabela 8 – Estimativas (e intervalos de confiança de 95%) dos parâmetros do modelo k -MS ajustados aos dados gerados do modelo Skellam tradicional de tamanho $n = 100$.

Conjunto de Dados	k	Estimativas				
		$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\phi}$
CD1	-2	6,79 (5,45;20,09)	-1,94 (-3,02;12,22)	6,41 (5,37;7,47)	-4,88 (-5,96;-3,87)	9,23 (7,00;11,68)
	0	3,26 (1,86;7,59)	-0,11 (-2,14;4,47)	6,54 (5,50;7,67)	-4,94 (-6,07;-3,91)	8,63 (7,00;11,01)
	4	1,44 (0,45;2,63)	1,45 (0,31;3,78)	6,56 (5,50;7,69)	-4,87 (-6,00;-3,79)	9,08 (7,00;11,47)
CD2	-2	1,97 (1,08;3,13)	-0,30 (-1,27;0,93)	-5,60 (-6,60;-4,61)	2,15 (1,13;3,18)	7,90 (5,94;9,65)
	0	3,37 (2,41;5,18)	-0,89 (-2,11;0,13)	-5,37 (-6,35;-4,38)	1,90 (0,87;2,84)	7,71 (5,64;9,45)
	4	7,48 (4,66;19,83)	12,37 (3,22;13,92)	-5,44 (-6,35;-4,48)	2,08 (1,11;3,04)	7,85 (5,79;9,63)
CD3	-2	3,02 (1,85;5,11)	-0,52 (-1,95;1,21)	2,77 (2,00;3,58)	-2,02 (-2,83;-1,23)	4,75 (3,47;6,04)
	0	2,00 (1,20;3,09)	-0,25 (-1,16;0,70)	2,56 (1,77;3,41)	-1,82 (-2,65;-1,06)	4,96 (3,62;6,35)
	4	0,96 (0,13;1,93)	1,80 (0,80;3,47)	2,39 (1,46;3,21)	-1,71 (-2,54;-0,84)	4,95 (3,42;6,45)

Tabela 9 – Estimativas (e intervalos de confiança de 95%) dos parâmetros do modelo k -MS ajustados aos dados gerados do modelo Skellam tradicional de tamanho $n = 200$.

Conjunto de Dados	k	Estimativas				
		$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\phi}$
CD1	-2	4,68 (3,71;6,51)	-1,60 (-2,70;-8,83)	6,58 (5,89;7,29)	-4,42 (-5,05;-3,83)	8,54 (7,00;10,32)
	0	3,63 (2,92;4,78)	-0,93 (-1,62;-0,33)	6,46 (5,77;7,18)	-4,34 (-4,99;-3,74)	8,63 (7,00;10,31)
	4	2,38 (0,68;3,03)	-0,07 (-0,41;2,32)	6,99 (6,08;8,14)	-4,57 (-6,06;-4,02)	7,51 (7,00;10,47)
CD2	-2	2,31 (1,67;3,09)	-0,20 (-0,81;0,48)	-5,27 (-5,89;-4,67)	2,39 (1,87;2,90)	6,25 (5,06;7,66)
	0	3,53 (2,72;4,77)	-1,18 (-2,05;-0,47)	-5,39 (-6,01;-4,79)	2,45 (1,91;2,99)	6,41 (5,21;7,80)
	4	9,57 (3,89;19,84)	10,25 (2,72;12,56)	-5,42 (-6,02;-4,83)	2,58 (2,02;3,11)	6,47 (5,24;7,94)
CD3	-2	3,49 (2,68;4,67)	-1,31 (-2,12;-0,63)	2,64 (2,10;3,19)	-1,39 (-1,86;-0,90)	4,40 (3,29;5,48)
	0	2,80 (2,17;3,68)	-0,84 (-1,49;-0,28)	2,43 (1,94;2,95)	-1,42 (-1,87;-0,99)	4,53 (3,57;5,52)
	4	1,02 (0,29;1,76)	1,57 (0,62;3,28)	2,45 (1,89;3,00)	-1,44 (-1,93;-0,97)	4,58 (3,53;5,61)

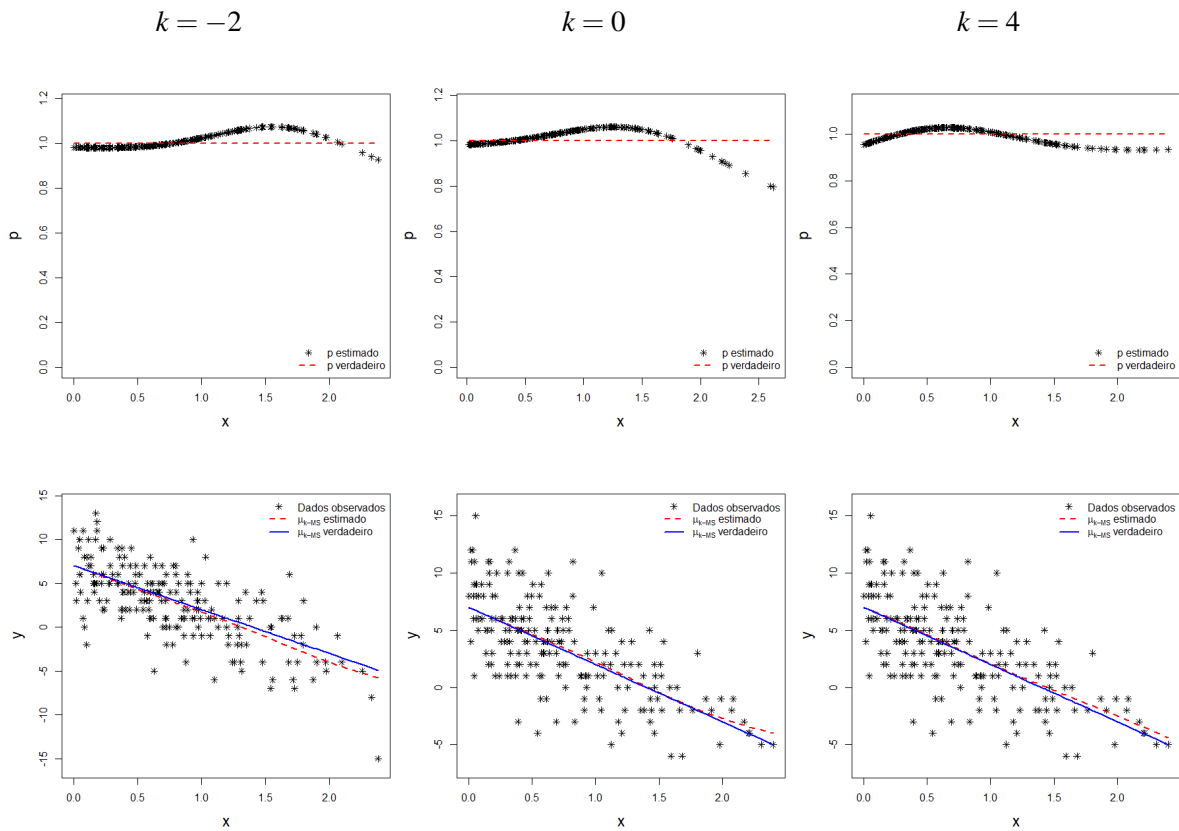


Figura 3 – Estimativas de p e μ_{k-MS} (juntamente com seus valores reais) obtidas a partir do ajuste do modelo k -MS, para valores de parâmetros do CD1 e tamanho de amostra $n = 100$.

4.3 Estimação dos Parâmetros: Abordagem Bayesiana

Para a estimação dos parâmetros, sob o contexto bayesiano, optamos por uma abordagem considerando um procedimento MCMC, mais especificamente, o método do Gradiente Estocástico Hamiltoniano de Monte Carlo.

4.3.1 Distribuição Conjunta a Posteriori

Seja $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$ um vetor de n variáveis aleatórias independentes, com $Y_i \sim k\text{-MS}(\mu_i, \phi, \omega_i)$, para todo $i = 1, 2, \dots, n$, e seja $\mathbf{y} = (y_1, y_2, \dots, y_n)$ o vetor de observações associados a \mathbf{Y} . Considere $\mathbf{x}_i = (1, x_{1i}, \dots, x_{q_1i})$ e $\mathbf{v}_i = (1, v_{1i}, \dots, v_{q_2i})$ os vetores de covariáveis associados a cada y_i e, conseqüentemente, $\mu_i = \mu(\mathbf{x}_i)$ e $\omega_i = \omega(\mathbf{v}_i)$. Denotamos por $\mathcal{D} = \{y_i, \mathbf{x}_i, \mathbf{v}_i : \forall i = 1, \dots, n\}$ as informações relacionadas a todos os indivíduos.

Considerando o modelo k -MS na sua versão *hurdle*, (4.4), a função de verossimilhança associada a \mathbf{y} é dada por:

$$\begin{aligned} \mathcal{L}(\mu(\mathbf{x}_i), \phi, \omega(\mathbf{v}_i); \mathcal{D}) &= \prod_{i=1}^n \left\{ (1 - \omega(\mathbf{v}_i)) I_{\{k\}}(y_i) + \omega(\mathbf{v}_i) \pi_{k\text{-SS}}(y_i; \mu(\mathbf{x}_i), \phi) \right\} \\ &= \prod_{i=1}^n \left\{ (1 - \omega(\mathbf{v}_i))^{I_{\{k\}}(y_i)} \left(\frac{\omega(\mathbf{v}_i) \pi_s(y_i; \mu(\mathbf{x}_i), \phi)}{1 - \pi_s(k; \mu(\mathbf{x}_i), \phi)} \right)^{1 - I_{\{k\}}(y_i)} \right\}. \end{aligned}$$

Como apresentado no Capítulo anterior, em (4.6), podemos escrever a função de verossimilhança em termos de $\ell_1(\boldsymbol{\alpha}; \mathcal{D})$ e $\ell_2(\boldsymbol{\beta}, \phi; \mathcal{D})$ como:

$$\mathcal{L}(\boldsymbol{\beta}, \phi, \boldsymbol{\alpha}; \mathcal{D}) = \exp\{\ell_1(\boldsymbol{\alpha}; \mathcal{D}) + \ell_2(\boldsymbol{\beta}, \phi; \mathcal{D})\}.$$

Para prosseguir com o procedimento de estimação Bayesiano, precisamos especificar uma distribuição a priori para os parâmetros do modelo. A ideia é especificar uma distribuição a priori de forma que, mesmo para tamanhos moderados de amostra, a informação fornecida pelos dados domine a informação a priori. Assim, com a fatoração da função de verossimilhança, sugere distribuições a priori independentes para $\boldsymbol{\alpha}$ e o par $(\boldsymbol{\beta}, \phi)$, denotados por $\pi(\boldsymbol{\alpha})$ e $\pi(\boldsymbol{\beta}, \phi)$, respectivamente. Desta forma, a distribuição conjunta a posteriori é dada proporcionalmente por

$$\pi(\boldsymbol{\beta}, \phi, \boldsymbol{\alpha}; \mathcal{D}) \propto \exp\{\ell_1(\boldsymbol{\alpha}; \mathcal{D}) + \ell_2(\boldsymbol{\beta}, \phi; \mathcal{D})\} \pi(\boldsymbol{\beta}, \phi) \pi(\boldsymbol{\alpha}).$$

Assumimos $\boldsymbol{\beta}$ e ϕ independentes e, portanto, $\pi(\boldsymbol{\beta}, \phi) = \pi(\boldsymbol{\beta})\pi(\phi)$. Consideramos para $\boldsymbol{\beta}$ e $\boldsymbol{\alpha}$ uma distribuição priori normal multivariada independente $N(\boldsymbol{\theta}_0, \sigma_0^2 \mathbf{I})$, em que $\boldsymbol{\theta}_0 = (b_0, b_1, \dots, b_{q_1}, a_0, a_1, \dots, a_{q_2})^\top$ é um vetor de médias; $\sigma_0 > 0$ e \mathbf{I} é uma matriz de identidade $(q_1 + q_2 + 2)$ -dimensional. Para ϕ , devido às restrições impostas a este parâmetro, $\phi \geq |\max\{|\mathbf{x}_i \boldsymbol{\beta}|, i = 1, \dots, n\}|$, consideramos uma distribuição priori normal truncada

$$\pi_0(\phi) = \frac{N(\phi_0, \sigma_1^2)}{1 - \Phi\left(\frac{\delta - \phi_0}{\sigma_1}\right)} = NT(\phi_0, \sigma_1^2, \delta),$$

tal que $\delta > \max\{|\mathbf{x}_i \boldsymbol{\beta}|, i = 1, \dots, n\}$ e $\Phi(\cdot)$ é a distribuição acumulada da normal padrão.

As inferências sobre os parâmetros podem ser baseadas em suas distribuições posteriores marginais, que podem ser obtidas integrando a densidade completa a posteriori $\pi(\boldsymbol{\beta}, \phi, \boldsymbol{\alpha}; \mathcal{D})$ com relação a cada parâmetro (CARLIN; LOUIS, 2008). Porém, não é possível obter soluções analíticas para as integrais. Superamos essa dificuldade usando o método de integração numérica baseado no procedimento de Monte Carlo para obter amostras das distribuições marginais e, assim, fazer inferências sobre os parâmetros considerando estimativas de Monte Carlo. Neste trabalho, consideramos o algoritmo Gradiente Estocástico Hamiltoniano Monte Carlo (SGHMC) para gerar amostras das distribuições marginais (CHEN; FOX; GUESTRIN, 2014). O algoritmo SGHMC é derivado do algoritmo Hamiltoniano Monte Carlo (HMC) proposto por Duane *et al.* (1987), e posteriormente desenvolvido por Neal (1994), que introduziu a aplicação do HMC para problemas estatísticos. Um texto amigável para os leitores interessados no algoritmo HMC pode ser encontrado em Neal (2011).

4.3.2 Algoritmo Hamiltoniano de Monte Carlo

Um grande desafio nos procedimentos de aceitação-rejeição para obter a estimativa de parâmetros é encontrar uma amostra típica da distribuição-alvo. O algoritmo HMC explora de maneira eficiente as trajetórias dentro do espaço paramétrico. Essa exploração resulta não apenas melhor eficiência computacional do que outros algoritmos da classe MCMC, mas também garantias fortes sobre a validade dos estimadores resultantes (Stan Development Team, 2017).

Com base em atualizações simples para as variáveis, a simulação a partir de um sistema dinâmico hamiltoniano permite obter propostas de estados distantes do estado atual. Sendo a distribuição de destino é invariável sob essa dinâmica. O gradiente da distribuição fornece informações sobre o comportamento local da distribuição de destino. Assim, o novo estado é proposto calculando uma trajetória de acordo com a dinâmica hamiltoniana, a obtenção desse estado proposto é obtida com um método de integração denominado *leapfrog*. Um estado proposto desta forma pode estar distante do estado atual, mas tem uma alta probabilidade de aceitação. Assim, permite movimentos maiores através do espaço de estado (espaço paramétrico) e, portanto, convergência mais rápida para a distribuição de destino. Isso supera a lenta exploração do espaço de estados que ocorre quando o algoritmo de Metropolis-Hastings (MH) é utilizado.

O método HMC é uma técnica MCMC para amostragem de sistemas complexos e de alta dimensão. O hamiltoniano \mathcal{H} representa a energia do sistema, que é a soma da energia cinética e potencial tradicionalmente denotadas por $\varepsilon(\boldsymbol{\theta})$ e $K(\mathbf{z})$, onde $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$ e ϕ são as chamadas coordenadas generalizadas e o vetor \mathbf{z} é o chamado momento generalizado. Para simplificar a notação, denotamos o vetor de parâmetros como $\boldsymbol{\theta} = (\boldsymbol{\beta}, \phi, \boldsymbol{\alpha})^\top$ e, para o HMC, a função hamiltoniana pode ser escrita como segue

$$\mathcal{H}(\boldsymbol{\theta}, \mathbf{z}) = \varepsilon(\boldsymbol{\theta}) + K(\mathbf{z}).$$

Aqui, $\varepsilon(\boldsymbol{\theta})$ é chamada de energia potencial e será definida como igual a menos o logaritmo da densidade de probabilidade da distribuição de $\boldsymbol{\theta}$. A função $K(\mathbf{z})$ é chamada de energia cinética e geralmente é definida como

$$K(\mathbf{z}) = \frac{\mathbf{z}'\boldsymbol{\Sigma}^{-1}\mathbf{z}}{2},$$

em que $\boldsymbol{\Sigma}$ é simétrica, definida-positiva, geralmente diagonal, múltipla por um escalar da matriz de identidade. Aqui, \mathbf{z} será representado por uma variável aleatória normal $Z \sim N(0, \boldsymbol{\Sigma})$ da mesma dimensão dos parâmetros ($\boldsymbol{\theta} = (\boldsymbol{\beta}, \phi, \boldsymbol{\alpha})^\top$): $q^* = q_1 + q_2 + 3$. Esta forma para $K(\mathbf{z})$ corresponde ao valor negativo do logaritmo natural da função de densidade de probabilidade (mais uma constante) da distribuição Gaussiana de média-zero com matriz de covariância $\boldsymbol{\Sigma}$. As derivadas parciais do Hamiltoniano determinam como $\boldsymbol{\theta}$ e \mathbf{z} mudam ao longo do tempo, t , de acordo com as equações de Hamilton (aqui entende-se por tempo as iterações):

$$\frac{d\theta_l}{dt} = \frac{\partial \mathcal{H}(\boldsymbol{\theta}, \mathbf{z})}{\partial z_l} \quad \text{e} \quad \frac{dz_l}{dt} = -\frac{\partial \mathcal{H}(\boldsymbol{\theta}, \mathbf{z})}{\partial \theta_l}, \quad l = 1, \dots, q^*.$$

Para qualquer intervalo de tempo de duração s , essas equações definem uma transição do estado em qualquer momento t para o estado no momento $t + s$.

Na estatística bayesiana, o vetor de parâmetros $\boldsymbol{\theta}$ assumirá o papel da posição. Podemos expressar a distribuição posteriori como uma distribuição canônica usando uma função de energia potencial definida por:

$$\varepsilon(\boldsymbol{\theta}) = -\log\{\pi(\boldsymbol{\theta}; \mathcal{D})\} = -\log(\exp\{\ell(\boldsymbol{\theta}; \mathcal{D})\}\pi(\boldsymbol{\theta})) + \log(C),$$

em que $C > 0$ é a constante normalizadora; $\pi(\boldsymbol{\theta})$ e $\ell(\boldsymbol{\theta}; \mathcal{D})$ são, respectivamente,

$$\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta})\pi(\phi)\pi(\boldsymbol{\alpha}); \quad \text{e}$$

$$\ell(\boldsymbol{\theta}; \mathcal{D}) = \ell_1(\boldsymbol{\alpha}; \mathcal{D}) + \ell_2(\boldsymbol{\beta}, \phi; \mathcal{D}).$$

Assim, o hamiltoniano \mathcal{H} é dado por:

$$\mathcal{H}(\boldsymbol{\theta}; \mathbf{z}) = -\log\{\pi(\boldsymbol{\theta}; \mathcal{D})\} + \frac{\mathbf{z}'\boldsymbol{\Sigma}^{-1}\mathbf{z}}{2},$$

e as derivadas parciais da equação hamiltoniana são dadas por:

$$\begin{pmatrix} \frac{d\theta_1}{dt} \\ \vdots \\ \frac{d\theta_{q^*}}{dt} \end{pmatrix} = \boldsymbol{\Sigma}^{-1}\mathbf{z} \quad \text{e} \quad \begin{pmatrix} \frac{dz_1}{dt} \\ \vdots \\ \frac{dz_{q^*}}{dt} \end{pmatrix} = \begin{pmatrix} G_{\theta_1} \\ \vdots \\ G_{\theta_{q^*}} \end{pmatrix},$$

em que $G_{\theta_l} = U_{\theta_l} + \frac{d\log(\pi(\theta_l))}{d\theta_l}$, para todo $l = 1, 2, \dots, q^*$.

O algoritmo HMC completo é apresentado a seguir:

Algoritmo 1 – Hamiltoniano Monte Carlo (HMC)**Requer:** M : número de amostras de Monte Carlo ε : passo de integração L : número de etapas de interações Σ : matriz de massa ($\Sigma \leftarrow \sigma^2 \mathbf{I}_{q^* \times q^*}$) $\boldsymbol{\theta}^{(0)}$ **para** $m \leftarrow 1$ até M **faça** $\boldsymbol{\theta}'^{(0)} \leftarrow \boldsymbol{\theta}^{(m-1)}$ Gere $\mathbf{z}^{(0)}$ de uma $N(0; \Sigma)$ Faça meio passo $\varepsilon/2$, iteração para o momento $\mathbf{z}'^{(0)}$ por:

$$z_i'^{(0)} \leftarrow z_i'^{(0)} + \frac{\varepsilon}{2} \left\{ U_{\theta_i}(\boldsymbol{\theta}'^{(m-1)}) + \frac{d}{d\theta_i} \log(\{\pi(\theta_i'^{(m-1)})\}) \right\}, \forall i = 1, 2, \dots, q^*$$

para $\tau \leftarrow 1$ até L **faça**

$$\theta_i'^{(\tau)} \leftarrow \theta_i'^{(\tau-1)} + \frac{\varepsilon}{\sigma^2} z_i'^{(\tau-1)}, \forall i = 1, 2, \dots, q^*$$

se $\tau \neq L$ **então**

$$z_i'^{(\tau)} \leftarrow z_i'^{(\tau-1)} + \varepsilon \left\{ U_{\theta_i}(\boldsymbol{\theta}'^{(\tau)}) + \frac{d}{d\theta_i} \log(\{\pi(\theta_i'^{(\tau)})\}) \right\}, \forall i = 1, 2, \dots, q^*$$

fim se**fim para**

$$z_i'^{(L)} \leftarrow z_i'^{(L-1)} + \varepsilon \left\{ U_{\theta_i}(\boldsymbol{\theta}'^{(L)}) + \frac{d}{d\theta_i} \log(\{\pi(\theta_i'^{(L)})\}) \right\}, \forall i = 1, 2, \dots, q^*$$

Calcule a probabilidade de aceitação:

$$\alpha \leftarrow \min \left\{ 1; \exp \left\{ -\mathcal{H}(\boldsymbol{\theta}'^{(L)}; \mathbf{z}'^{(L)}) + \mathcal{H}(\boldsymbol{\theta}^{(m-1)}; \mathbf{z}^{(m-1)}) \right\} \right\}$$

Gere $u \sim U(0, 1)$.**se** $\mu \leq \alpha$ **então**

$$\boldsymbol{\theta}^{(m)} \leftarrow \boldsymbol{\theta}'^{(L)} \quad (\text{aceita})$$

senão

$$\boldsymbol{\theta}^{(m)} \leftarrow \boldsymbol{\theta}^{(m-1)} \quad (\text{rejeita})$$

fim se**fim para**

A escolha dos parâmetros no HMC pode afetar o desempenho do método. São três os principais parâmetros ajustáveis: o tamanho da etapa de integração ε , o número de etapas de integração L e a matriz de massa Σ . Geralmente podem ser escolhidos arbitrariamente, sem que a validade do método seja afetada, porém existem casos nos quais a ergodicidade da cadeia é prejudicada pela escolha desses parâmetros (NEAL, 2011).

Por outro lado, o passo de aceitação e rejeição de MH pode levar a uma baixa taxa de aceitação e cada rejeição corresponde a um desperdício de cálculos do gradiente. Dessa forma, uma variante do algoritmo HMC foi proposta por Chen, Fox e Guestrin (2014), a qual consiste na reformulação das equações de Hamilton ao introduzir uma componente estocástica na equação do gradiente, evitando o passo de aceitação e rejeição MH.

4.3.2.1 Algoritmo Gradiente Estocástico Hamiltoniano de Monte Carlo

O HMC simula a dinâmica Hamiltoniana de acordo com as equações de Hamilton para gerar amostras. Chen, Fox e Guestrin (2014) considera a possibilidade de modificar as equações de Hamilton, acrescentando um termo de “fricção” à atualização do momento. As equações de Hamilton com fricção são dadas por:

$$\begin{aligned}\frac{d\boldsymbol{\theta}}{dt} &= \Sigma^{-1}\mathbf{z} \\ \frac{d\mathbf{z}}{dt} &= -\mathbf{G}(\boldsymbol{\theta}) - \mathbf{B}\Sigma^{-1}\mathbf{z} + \mathcal{N}(\mathbf{0}, 2\mathbf{B}),\end{aligned}$$

em que \mathbf{B} é a matriz de difusão e $\mathbf{G}(\boldsymbol{\theta})$ é um vetor cujos os elementos são

$$G_{\theta_i} = U_{\theta_i} + \frac{d \log(\pi(\theta_i))}{d\theta_i}, \text{ para todo } i = 1, 2, \dots, q^*.$$

Cabe destacar que utilizamos um abuso de notação ao adicionar $\mathcal{N}(\mathbf{0}, 2\mathbf{B})$, que denota a introdução de uma variável aleatória que é distribuída de acordo com esta gaussiana multivariada. Na prática, a matriz \mathbf{B} não é conhecida e precisa ser estimada, $\hat{\mathbf{B}}$. Chen, Fox e Guestrin (2014) introduziu um termo de atrito \mathbf{C} , especificado pelo usuário (\mathbf{C} domina \mathbf{B}), e considerou a seguinte dinâmica:

$$\frac{d\boldsymbol{\theta}}{dt} = \Sigma^{-1}\mathbf{z} \quad (4.8)$$

$$\frac{d\mathbf{z}}{dt} = -\mathbf{G}(\boldsymbol{\theta}) - \mathbf{C}\Sigma^{-1}\mathbf{z} + \mathcal{N}(\mathbf{0}, 2(\mathbf{C} - \hat{\mathbf{B}})) + \mathcal{N}(\mathbf{0}, 2\mathbf{B}). \quad (4.9)$$

No cenário mais realista de estimativa imprecisa de \mathbf{B} , a escolha mais simples é $\hat{\mathbf{B}} = \mathbf{0}$. Assim, fazendo $\mathbf{B} = \hat{\mathbf{B}} = \mathbf{0}$, a atualização do momento simplifica-se para

$$\frac{d\mathbf{z}}{dt} = -\mathbf{G}(\boldsymbol{\theta}) - \mathbf{C}\Sigma^{-1}\mathbf{z} + \mathcal{N}(\mathbf{0}, 2\mathbf{C}).$$

Portanto, a dinâmica é governada pelo ruído introduzido (controlável) $\mathcal{N}(\mathbf{0}, 2\mathbf{C})$ e pelo termo de atrito $\mathbf{C}\Sigma^{-1}$. Para aplicações práticas, as equações diferenciais (4.8)-(4.9) não podem ser

resolvidas analiticamente, sendo necessários recorrer a procedimentos numéricos. Reescrevemos (4.8)-(4.9) na forma discretizada onde ε é o tamanho do passo de integração:

$$\Delta\boldsymbol{\theta} = \varepsilon\Sigma^{-1}\mathbf{z} \quad (4.10)$$

$$\Delta\mathbf{z} = -\varepsilon\mathbf{G}(\boldsymbol{\theta}) - \alpha\mathbf{z} + \mathcal{N}(\mathbf{0}, 2\alpha\Sigma). \quad (4.11)$$

em que $\alpha = \varepsilon\mathbf{C}\Sigma^{-1}$. A constante α e o passo de discretização ε são constantes de ajuste; na prática, α tende a ser fixado em um valor pequeno. Na implementação, fixamos $\alpha = 0,10$ e $\varepsilon = 0,01$. Para resolver (4.10)-(4.11), consideramos o método integrador *leapfrog* resumido nos seguintes passos do algoritmo SGHMC:

Algoritmo 2 – Gradiente Estocástico Hamiltoniano Monte Carlo (SGHMC)

Requer:

M : número de amostras de Monte Carlo

ε : passo de integração

L : número de etapas de interações

Σ : matriz de massa ($\Sigma \leftarrow \sigma^2\mathbf{I}_{q^* \times q^*}$)

α : constante

$\boldsymbol{\theta}^{(0)}$

para $m \leftarrow 1$ até M **faça**

$\boldsymbol{\theta}'^{(0)} \leftarrow \boldsymbol{\theta}^{(m-1)}$

Gere $\mathbf{z}^{(0)}$ de uma $\mathcal{N}(\mathbf{0}; \Sigma)$

Faça

$$\mathbf{z}^{(0)} = (1 - \alpha)\mathbf{z}^{(0)} - \frac{\varepsilon}{2}\mathbf{G}(\boldsymbol{\theta}^{(0)}) + \mathcal{N}(\mathbf{0}, 2\alpha\Sigma)$$

para $\tau \leftarrow 1$ até L **faça**

$$\boldsymbol{\theta}'^{(\tau)} = \boldsymbol{\theta}'^{(\tau-1)} + \varepsilon\Sigma^{-1}\mathbf{z}^{(\tau-1)}$$

se $\tau \neq L$ **então**

$$\mathbf{z}^{(\tau)} = (1 - \alpha)\mathbf{z}^{(\tau-1)} - \varepsilon\mathbf{G}(\boldsymbol{\theta}'^{(\tau)}) + \mathcal{N}(\mathbf{0}, 2\alpha\Sigma)$$

fim se

fim para

Calcule:

$$\mathbf{z}^{(L)} = (1 - \alpha)\mathbf{z}^{(L-1)} - \frac{\varepsilon}{2}\mathbf{G}(\boldsymbol{\theta}'^{(L)}) + \mathcal{N}(\mathbf{0}, 2\alpha\Sigma)$$

Faça:

$$\boldsymbol{\theta}^{(m)} \leftarrow \boldsymbol{\theta}'^{(L)} \quad (\text{sem passo de MH})$$

fim para

4.3.3 Critérios de Seleção de Modelos

A seguir, expomos uma breve descrição de alguns critérios que são empregados na seleção de modelos ajustados no contexto bayesiano.

4.3.3.1 Critério de Informação de Akaike Esperado (EAIC)

A adaptação do AIC para o contexto bayesiano é o Critério de Informação de Akaike Esperado (EAIC), uma extensão natural em que se considera a densidade a posteriori dos parâmetros do modelo (BROOKS, 2002). Mais especificamente, o EAIC é um critério baseado na média a posteriori da *deviance* e, para uma cadeia de tamanho M de cada parâmetro do vetor $\boldsymbol{\theta}$, tem seu cálculo aproximado por:

$$\widehat{EAIC} = -\frac{2}{M} \sum_{m=1}^M \ell(\boldsymbol{\theta}^{(m)}; \mathcal{D}) + 2c.$$

O EAIC penaliza modelos mais complexos, adicionando um termo que cresce com o número de parâmetros. Um modelo com um valor menor de EAIC é considerado o mais adequado, sugerindo um bom ajuste aos dados com o mínimo de parâmetros necessários.

4.3.3.2 Critério de Informação Bayesiano Esperado (EBIC)

O Critério de Informação Bayesiano Esperado (EBIC) adapta o BIC para a abordagem bayesiana, utilizando a distribuição posteriori do modelo (CARLIN; LOUIS, 2008). Para uma cadeia de tamanho M de cada parâmetro do vetor $\boldsymbol{\theta}$, esta medida pode ser aproximada por:

$$\widehat{EBIC} = -\frac{2}{M} \sum_{m=1}^M \ell(\boldsymbol{\theta}^{(m)}; \mathcal{D}) + c \ln(n).$$

O EBIC penaliza modelos com maior complexidade de maneira mais rigorosa em comparação com o EAIC, pois o termo de penalização ($c \ln(n)$) é multiplicado por $\ln(n)$. O EBIC tende a favorecer modelos mais parcimoniosos em casos de tamanhos de amostras grandes.

4.3.3.3 Critério de Informação Watanabe-Akaike (WAIC)

O Critério de Informação Watanabe-Akaike (WAIC) foi proposto por Watanabe e Opper (2010) e pode ser interpretado como uma aproximação para a validação cruzada. É um critério usado no contexto bayesiano e seu cálculo é baseado em:

$$elpd_{waic} = lpd - p_{waic},$$

em que $p_{waic} = \sum_{i=1}^n \text{Var}_{\boldsymbol{\theta}|y}(\log(p(y_i|\boldsymbol{\theta})))$ e $lpd = \sum_{i=1}^n \log(p(y_i|y))$ é o logaritmo natural da densidade preditiva pontual.

Ao reescrever

$$\log(p(y_i/y)) = \log \int_{\Theta} p(y_i|\theta)p(\theta|y)d\theta$$

e considerando S simulações para aproximar a variância do logaritmo natural do modelo para cada ponto $y_i|\theta$ pela variância amostral, temos:

$$\hat{p}_{waic} = \sum_{i=1}^n \text{Var}_{s=1}^S (\log(p(y_i|\theta^{(s)}))) \quad \text{e} \quad \widehat{lpd} = \sum_{i=1}^n \log \left(\frac{1}{S} \sum_{s=1}^S p(y_i|\theta^{(s)}) \right).$$

Assim, o WAIC é definido por:

$$\begin{aligned} WAIC &= -2\widehat{elpd}_{waic} \\ &= -2\{\widehat{lpd} - \hat{p}_{waic}\}. \end{aligned}$$

Portanto, o critério WAIC é uma abordagem bayesiana para seleção de modelos que leva em conta a incerteza na estimativa do modelo. Ele usa a distribuição preditiva do modelo e fornece uma estimativa da discrepância esperada entre os valores preditos pelo modelo e os valores reais. Quanto menor o valor de WAIC, melhor o ajuste do modelo aos dados.

Resumidamente, o EAIC, o EBIC e o WAIC são critérios que auxiliam na seleção do modelo ideal, balanceando sua complexidade (número de parâmetros) com sua capacidade preditiva. Modelos com valores mais baixos desses critérios são preferíveis.

4.3.4 Estudo de Simulação: Avaliação da Performance dos Estimadores Bayesianos

Conduzindo de maneira similar ao estudo de simulação descrito na Subseção 4.2.3 para a abordagem clássica, $N = 100$ conjuntos de dados de tamanho n ($n = 100$ e 200) foram gerados do modelo k -MS, considerando uma covariável $X \sim ZTN(0, 1)$. Para fins de comparação dos procedimentos, os valores considerados para o ponto de modificação k , para os vetores de parâmetros β e α e para o parâmetro ϕ foram os mesmos apresentados na Tabela 2.

Para cada um dos N conjuntos de dados gerados, o algoritmo SGHMC é usado para gerar uma cadeia de tamanho 50.000, com 50% dos elementos descartados como *burn-in*, para cada um dos parâmetros do modelo. A convergência das cadeias foi monitorada por meio do critério de Geweke (GEWEKE, 1992), e a média posteriori (associada à função de perda quadrática) foi considerada como estimador bayesiano. Dessa forma, após os ajustes do modelo, obtemos as N estimativas dos parâmetros, e então são calculadas algumas medidas para avaliar o desempenho dos estimadores.

As medidas calculadas foram: o erro quadrático médio (EQM); a variância (Var); o viés absoluto médio (B); e a probabilidade de cobertura dos intervalos de credibilidade bayesianos (PC). Assim, considerando que θ representa cada um dos parâmetros do modelo, isto é, θ assume $\alpha_0, \alpha_1, \beta_0, \beta_1, \phi$, as medidas são obtidas por:

$$EQM(\hat{\theta}) = \frac{1}{N} \sum_{i=1}^N (\hat{\theta}_i - \theta)^2, \quad Var(\hat{\theta}) = \frac{1}{N} \sum_{i=1}^N (\hat{\theta}_i - \bar{\theta})^2,$$

$$\mathcal{B}(\hat{\theta}) = \frac{1}{N} \sum_{i=1}^N |\hat{\theta}_i - \theta|, \quad PC(\theta) = \frac{1}{N} \sum_{i=1}^M \delta_{\theta}^i,$$

em que δ_{θ}^i assume 1 se o i -ésimo intervalo *highest probability density* (HPD) com 95% de credibilidade contém o valor verdadeiro do parâmetro θ e assume 0 caso contrário.

As Tabelas 10 e 11 apresentam os resultados do estudo de simulação para os conjuntos de dados gerados com diferentes valores de k e n do modelo k -IS e do modelo k -DS respectivamente. Tanto para os casos k -inflacionados quanto para os casos k -deflacionados, os resultados apresentaram valores de erro quadrático médio, variância e viés absoluto médio muito próximos de zero, o que indica um ótimo desempenho dos estimadores bayesianos. As probabilidades de cobertura, em sua maioria, foram de 100% (acima do nível nominal de 95%). Ao comparar estes resultados com os obtidos com o procedimento clássico, podemos concluir que o procedimento de estimação bayesiano foi superior na qualidade das estimativas dos parâmetros do modelo k -MS. O procedimento bayesiano também foi mais vantajoso quando comparado o tempo computacional, se mostrando mais significativamente mais rápido.

Tabela 10 – Avaliação dos estimadores dos parâmetros do modelo k -IS via procedimento bayesiano, considerando diferentes valores de k e n .

k	Medidas	Estimadores					n
		$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\phi}$	
-2	EQM	0,0016	0,0006	0,0012	0,0003	0,0002	100
		0,0014	0,0006	0,0006	0,0002	0,0001	200
	Var	0,0016	0,0006	0,0012	0,0003	0,0002	100
		0,0014	0,0006	0,0007	0,0002	0,0001	200
	\mathcal{B}	0,0024	0,0025	0,0045	0,0008	0,0013	100
		0,0019	0,0007	0,0006	0,0000	0,0023	200
PC (%)	100,00	100,00	100,00	100,00	100,00	100	
	100,00	100,00	100,00	100,00	100,00	200	
0	EQM	0,0016	0,0008	0,0012	0,0004	0,0006	100
		0,0015	0,0008	0,0005	0,0002	0,0002	200
	Var	0,0016	0,0008	0,0012	0,0004	0,0001	100
		0,0015	0,0008	0,0004	0,0002	0,0001	200
	\mathcal{B}	0,0031	0,0020	0,0083	0,0064	0,0227	100
		0,0030	0,0023	0,0063	0,0033	0,0127	200
PC (%)	100,00	100,00	100,00	100,00	100,00	100	
	100,00	100,00	100,00	100,00	100,00	200	
4	EQM	0,0019	0,0007	0,0015	0,0003	0,0007	100
		0,0011	0,0006	0,0011	0,0002	0,0004	200
	Var	0,0019	0,0007	0,0013	0,0003	0,0003	100
		0,0011	0,0006	0,0009	0,0002	0,0002	200
	\mathcal{B}	0,0013	0,0013	0,0151	0,0020	0,0201	100
		0,0006	0,0004	0,0119	0,0004	0,0135	200
PC (%)	100,00	100,00	100,00	100,00	100,00	100	
	100,00	100,00	100,00	100,00	100,00	200	

Tabela 11 – Avaliação dos estimadores dos parâmetros do modelo k -DS via procedimento bayesiano, considerando diferentes valores de k e n .

k	Medidas	Estimadores					n	
		$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\phi}$		
-2	EQM	0,0006	0,0002	0,0061	0,0017	0,0225	100	
		0,0004	0,0001	0,0030	0,0015	0,0063	200	
	Var	0,0006	0,0002	0,0022	0,0016	0,0006	100	
		0,0005	0,0002	0,0018	0,0014	0,0003	200	
	B	0,0045	0,0008	0,0628	0,0110	0,1477	100	
		0,0004	0,0003	0,0350	0,0082	0,0776	200	
	PC (%)	100,00	100,00	99,00	100,00	98,00	100	
		100,00	100,00	100,00	100,00	100,00	200	
	0	EQM	0,0000	0,0000	0,0023	0,0016	0,0059	100
			0,0000	0,0000	0,0014	0,0011	0,0016	200
		Var	0,0000	0,0000	0,0015	0,0014	0,0002	100
			0,0000	0,0000	0,0013	0,0011	0,0001	200
B		0,0002	0,0001	0,0287	0,0145	0,0755	100	
		0,0006	0,0001	0,0133	0,0051	0,0388	200	
PC (%)		100,00	100,00	100,00	100,00	100,00	100	
		100,00	100,00	100,00	100,00	100,00	200	
4		EQM	0,0000	0,0000	0,0007	0,0007	0,3121	100
			0,0000	0,0000	0,0009	0,0010	0,0068	200
		Var	0,0000	0,0000	0,0007	0,0006	0,2850	100
			0,0000	0,0000	0,0009	0,0009	0,0068	200
	B	0,0005	0,0004	0,0045	0,0085	0,1729	100	
		0,0004	0,0001	0,0022	0,0097	0,0014	200	
	PC (%)	100,00	100,00	100,00	100,00	100,00	100	
		100,00	100,00	100,00	100,00	100,00	200	

4.3.5 Aplicação: Conjuntos de Dados Artificiais

De maneira similar à descrita na Subseção 4.2.4, ilustrações com um único conjunto de dados de cada caso apresentado no estudo acima será aqui apresentado. Isto é: conjuntos de dados artificiais de tamanho amostral $n = 100$ e 200 gerados do modelo k -MS considerando n observações de uma variável explicativa $X \sim ZTN(0; 1)$ e diferentes pontos de k -modificação. Os valores de k , dos vetores de parâmetros β e α e do parâmetro ϕ foram os mesmos apresentados na Tabela 2.

Considerando a abordagem bayesiana, o modelo k -MS foi ajustado a cada conjunto de dados gerado e as estimativas bayesianas dos parâmetros (média a posteriori) e o intervalo HPD com 95% de credibilidade foram obtidos. A Tabela 12 apresenta estes resultados. Em todos os casos considerados, é possível notar que as estimativas encontram-se próximas dos verdadeiros valores dos parâmetros, além destes estarem contidos nos respectivos intervalos de credibilidade.

Tabela 12 – Estimativas (e intervalos HPD com 95% de credibilidade) dos parâmetros do modelo k -MS (k -IS e k -DS), via procedimento bayesiano.

k	Estimativas (IC 95%)					n
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\phi}$	
Modelo Skellam k -Inflacionado						
-2	1,94 (1,75; 2,12)	-3,53 (-3,72; -3,34)	3,03 (2,84; 3,21)	-1,98 (-2,18; -1,79)	4,50 (4,32; 4,71)	100
	2,00 (1,83;2,17)	-3,51 (-3,69;-3,32)	2,95 (2,77;3,13)	-2,02 (-2,20;-1,83)	4,52 (4,33;4,71)	200
0	1,93 (1,76; 2,12)	-3,04 (-3,23; -2,86)	6,02 (5,84; 6,21)	-2,49 (-2,68; -2,29)	7,49 (7,31; 7,70)	100
	2,02 (1,85; 2,20)	-3,01 (-3,19; -2,82)	5,96 (5,77; 6,14)	-2,50 (-2,69; -2,31)	7,50 (7,31; 7,69)	200
4	1,56 (1,38; 1,74)	-2,97 (-3,16; -2,78)	6,97 (6,79; 7,16)	-5,02 (-5,21; -4,83)	8,51 (8,32; 8,71)	100
	1,48 (1,31; 1,65)	-3,03 (-3,21; -2,84)	7,00 (6,82; 7,18)	-5,01 (-5,19; -4,81)	8,49 (8,29; 8,68)	200
Modelo Skellam k -Deflacionado						
-2	2,01 (1,82; 2,20)	1,76 (1,57; 1,95)	-4,08 (-4,26; -3,91)	1,46 (1,28; 1,64)	5,42 (5,23; 5,62)	100
	2,00 (1,81;2,19)	1,76 (1,56;1,95)	-4,08 (-4,26;-3,92)	1,54 (1,38;1,71)	5,35 (5,15;5,54)	200
0	6,00 (5,80; 6,20)	4,00 (3,81; 4,19)	-5,06 (-5,24; -4,88)	1,95 (1,77; 2,14)	6,48 (6,28; 6,67)	100
	6,00 (5,81;6,20)	4,00 (3,80;4,19)	-5,04 (-5,21;-4,86)	2,03 (1,86;2,20)	6,42 (6,23;6,61)	200
4	5,99 (5,79; 6,19)	1,99 (1,80; 2,18)	-5,03 (-5,22; -4,85)	6,98 (6,80; 7,17)	15,50 (15,31; 15,71)	100
	6,00 (5,81;6,20)	2,00 (1,80;2,19)	-4,99 (-5,18;-4,81)	7,05 (6,86;7,23)	15,49 (15,30;15,68)	200

A Figura 4 apresenta os dados gerados de Y versus X , juntamente com μ_{k-MS} e $\hat{\mu}_{k-MS}$, para cada caso apresentado na Tabela 12. Na Figura 5 são apresentadas as estimativas para o parâmetro p . Nota-se as proximidades entre os valores verdadeiros e ajustados em todos os casos, indicando a boa qualidade do procedimento bayesiano via algoritmo SGHMC para a estimação dos parâmetros.

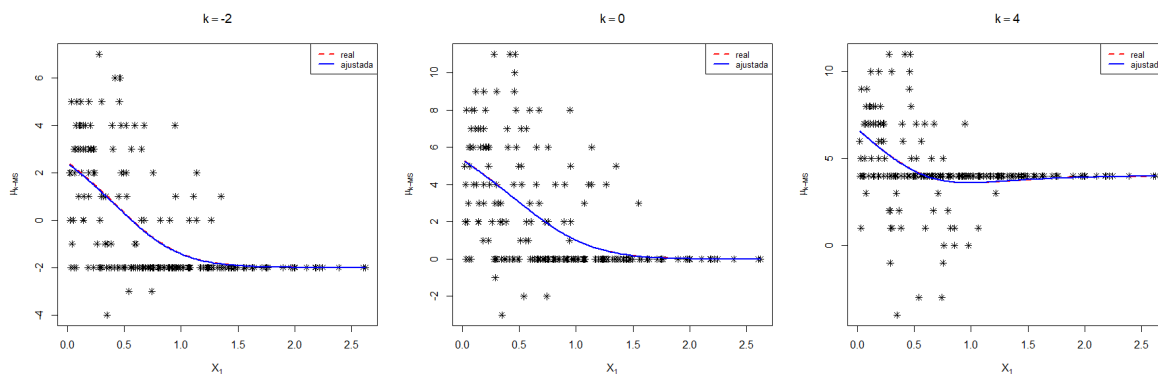
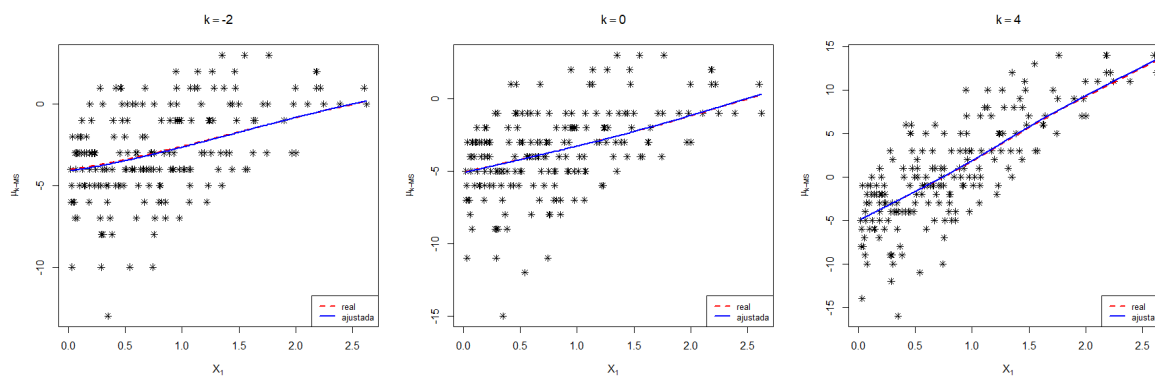
Modelo Skellam k -InflacionadoModelo Skellam k -Deflacionado

Figura 4 – Estimativas bayesianas de μ_{k-MS} (juntamente com seus valores reais) obtidas a partir do ajuste do modelo k -MS, para $n = 200$.

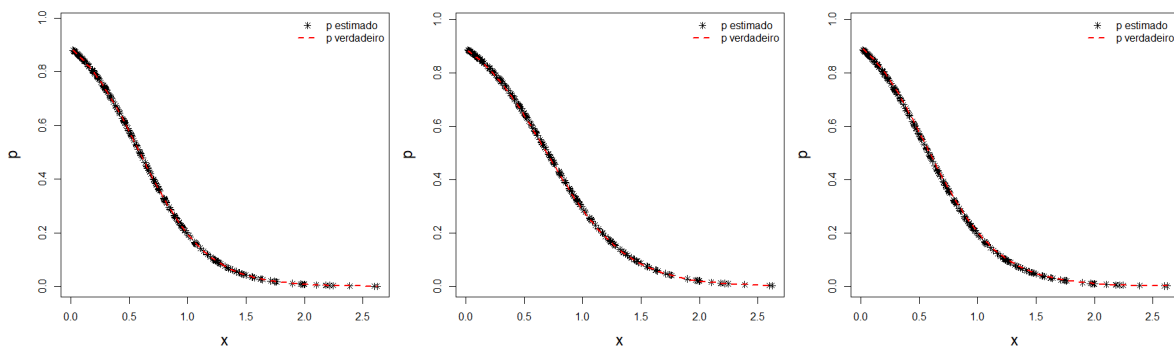
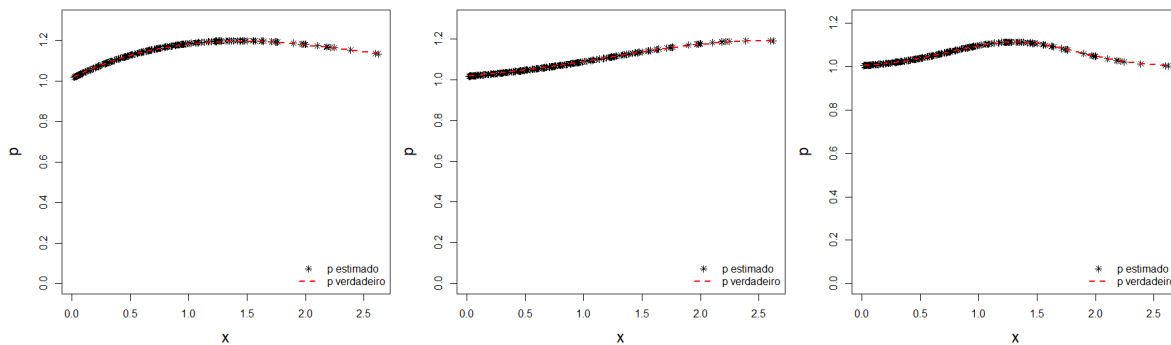
Modelo Skellam k -InflacionadoModelo Skellam k -Deflacionado

Figura 5 – Estimativas bayesianas do parâmetro p obtidas a partir do ajuste do modelo k -MS, para $n = 200$.

4.4 Aplicação: Conjunto de Dados Reais

4.4.1 Dados do Índice da Bolsa de Valores de São Paulo

O Índice da Bolsa de Valores de São Paulo (Ibovespa) desempenha um papel fundamental no mercado financeiro brasileiro como indicador do comportamento das transações de compra e venda de ações. É composto pelas principais ações negociadas na Bolsa de Valores brasileira, visto que contém os ativos que movimentam em torno de 80% do volume total de negociações da empresa B3 (uma das principais empresas de infraestrutura de mercado financeiro no mundo, com atuação em ambiente de bolsa e de balcão).

A pontuação do Ibovespa é expressa em reais e o cálculo é dado pela multiplicação do preço de cada ação pelo percentual de participação da empresa na carteira. Dessa forma, o índice total reflete o valor total das ações que compõem a B3. Se o índice sobe, isso indica que, em média, as ações que o compõem se valorizaram, e se ele cair, significa que boa parte das ações perderam valor. Essas flutuações são reflexos das expectativas dos investidores em relação aos ativos e ao cenário econômico, tanto interno quanto externo.

Dessa forma, o Ibovespa é um importante indicador da dinâmica do mercado de ações no Brasil e é influenciado por uma série de fatores, tais como a saúde financeira das empresas, a política econômica do país, a situação internacional e as condições do mercado financeiro em geral. Atualmente, o índice é composto por cerca de 60 empresas dos mais distintos ramos, como varejo, *commodities* e fabricantes de bens de consumo.

Um termo frequentemente empregado para comparar a mudança em valores do mercado de ações é chamado de *tick*. Essa medida é usada para avaliar as mudanças de preços de forma discreta. Tendo isso em vista, nossos dados consistem em mudanças discretas de preços, *tick-by-tick*, do Ibovespa. Aqui definimos o tamanho do *tick* (*tick size*) no valor de 1 centavo da moeda monetária. Assim, usaremos como variável resposta a diferença de preço, medida em *ticks* (centavos), da pontuação do Ibovespa, do dia t e do dia $t - 1$. A partir desses dados, construímos uma nova série de dados históricos usando o logaritmo da razão entre o dia atual e o dia anterior, isto é, o log-retorno dos valores do índice dado por

$$\log \left(\frac{y_t}{y_{t-1}} \right).$$

Neste contexto, o conjunto de dados é composto por 1194 observações da variação semanal da pontuação do Ibovespa, correspondente ao período de janeiro de 2000 até dezembro de 2022. As negociações das ações acontecem diariamente, de segunda a sexta-feira, no período compreendido entre as 10h20 e as 16h. A distribuição de frequências das quantidades de *ticks* é apresentada na Tabela 13 e melhor ilustrada na Figura 6.

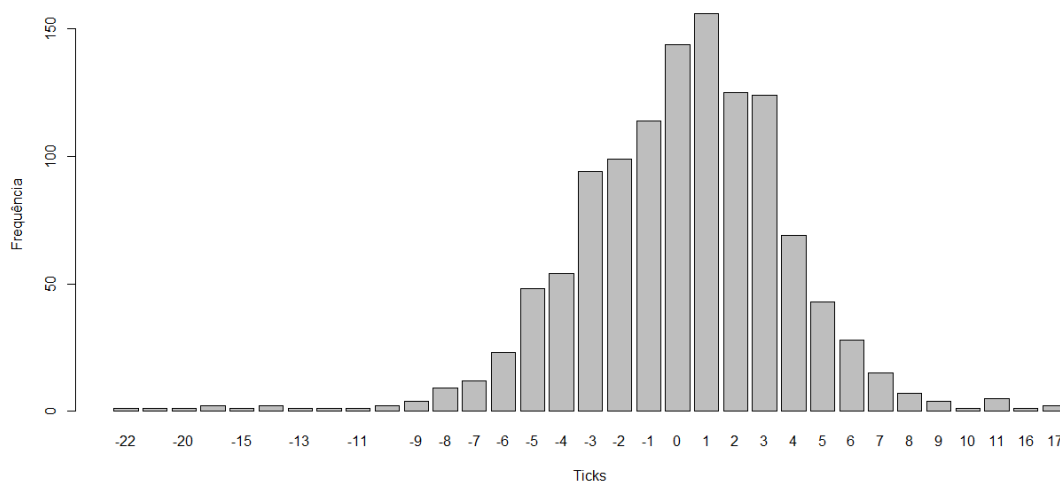


Figura 6 – Distribuição de frequências das quantidades de *ticks* registradas na variação semanal do Ibovespa no período de 2000-2022.

Tabela 13 – Distribuição de frequências e algumas estatísticas descritivas das quantidades de *ticks* registradas na variação semanal do Ibovespa no período de 2000-2022.

y_i	Conjunto de Dados																												\bar{y}	SD					
	-22	-21	-20	-17	-15	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8			9	10	11	16	17
f_i	1	1	1	2	1	2	1	1	1	2	4	9	12	23	48	54	94	99	114	144	156	125	124	69	43	28	15	7	4	1	5	1	2	0,14	3,76

Para a escolha da covariável levamos em conta a opinião dos economistas, que consideram o valor do dólar (em reais) um dos principais fatores responsáveis pela flutuação do índice brasileiro. Compreender a relação entre o preço do dólar e o comportamento do mercado brasileiro é fundamental tanto para investidores quanto para as medidas governamentais. Com esse objetivo, utilizamos o log-retorno do preço do dólar dos dias t e $t - 1$ como uma covariável no modelo.

Para o ajuste do modelo k -MS faz-se necessário o conhecimento do ponto de modificação k , que em aplicações reais, muitas vezes, esse número é desconhecido. Assim, para identificar o valor de k mais adequado, ajustamos o modelo k -MS considerando $k = -2, -1, 0, 1$ e 2 via método de máxima verossimilhança. Esses valores de k foram definidos com base na distribuição de frequências dos *ticks* apresentada na Figura 6. Após os ajustes, calculamos os critérios de seleção de modelos AIC e BIC . Menores valores do AIC e do BIC apontam para uma melhor adequação do modelo. A Tabela 14 apresenta os resultados destes ajustes, de onde é possível notar que o modelo k -MS com $k = -2$ foi apontado como o melhor ajustado, por apresentar os menores valores de AIC e BIC . Pelos valores estimados para o parâmetro p , temos o indicativo de inflação de valores no ponto -2 , caracterizando o conjunto de dados como inflacionado de -2 .

Tabela 14 – Estimativas dos parâmetros e critérios de seleção dos modelos k -MS ajustados aos dados do Ibovespa, considerando diferentes valores de k .

k	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\phi}$	AIC	BIC	$\min \hat{p}$	$\max \hat{p}$
-2	2,43	-7,10	0,25	-80,31	10,19	6036,85	6062,27	0,82	1,03
-1	2,24	-2,81	0,22	-80,19	10,21	6039,67	6065,09	0,87	1,03
0	2,02	-1,22	0,22	-80,68	10,26	6041,51	6066,94	0,86	1,00
1	1,89	0,38	0,24	-82,29	10,46	6045,11	6070,54	0,87	0,99
2	2,16	0,46	0,26	-81,14	10,29	6047,43	6072,85	0,90	1,02

Uma vez que os critérios de seleção de modelos apontaram para o modelo k -MS com $k = -2$, devido ao esforço computacional, os intervalos de confiança foram obtidos apenas para os parâmetros do melhor modelo ajustado. Para a obtenção destes intervalos, os dados foram reamostrados por meio do método *bootstrap* não paramétrico. Na Tabela 15 são apresentados os intervalos com 95% de confiança obtidos. Os intervalos de confiança indicam que os parâmetros são significativos, uma vez que nenhum deles contém o valor zero.

Tabela 15 – Estimativas e IC *bootstrap* (95%) dos parâmetros do modelo k -MS, com $k = -2$, ajustado aos dados do Ibovespa.

k	Estimativas (IC 95%)				
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\phi}$
-2	2,43 (2,20 ; 2,65)	-7,10 (-7,59 ; -0,93)	0,25 (0,09 ; 0,45)	-80,31 (-93,93 ; -74,09)	10,19 (8,67 ; 10,92)

A Figura 7 apresenta algumas ilustrações gráficas dos resultados do ajuste. O gráfico (A) desta Figura apresenta os valores estimados para o parâmetro p e o intervalo de confiança calculado via *bootstrap*. Já gráfico (B) apresenta as observações y (*ticks*) em função da covariável X (log-retorno do câmbio do dólar), juntamente com a média μ_{k-MS} estimada.

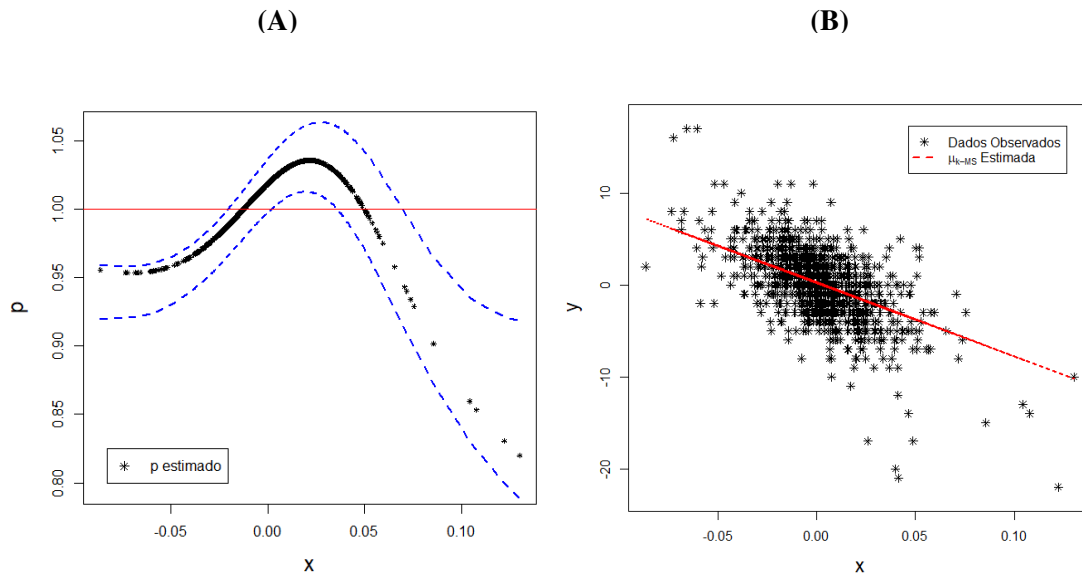


Figura 7 – Estimativas de p e de μ_{k-MS} obtidas a partir do modelo k -MS ajustado, com $k = -2$: (A) \hat{p} versus a covariável X ; (B) y versus a covariável X , juntamente com a estimativa de μ_{k-MS} .

Dos ajuste do modelo k -MS aos dados, podemos notar a relação entre variação no valor do Ibovespa, em *ticks*, com a variação no preço do dólar. A medida que a variação no preço do dólar se torna negativa, maior tende a ser a variação da pontuação do Ibovespa. Assim, ocorrências de baixas no câmbio do dólar influenciam positivamente o valor das ações que compõem a carteira do Ibovespa. No caso de aumento no preço do dólar, a pontuação Ibovespa apresenta tendência de queda, isto é, o valor das ações sofrem variação negativa.

4.4.2 Dados da Temporada Regular 2022-2023 da NBA

A *National Basketball Association* (NBA) é a principal liga de basquete dos Estados Unidos e de grande destaque no mercado esportivo internacional. Atualmente conta com 30 times (franquias), sendo que 29 desses estão localizadas em território estadunidense e um no Canadá. Os times são divididos em 2 conferências, Leste e Oeste, com 15 equipes em cada.

A NBA é composta por quatro fases: temporada regular, *play-in*, *playoffs* e final da NBA. Durante a temporada regular, cada time disputa 82 jogos, cuja pontuação de cada time na sua conferência é dada pela diferença entre o número total de vitórias e o número total de derrotas, podendo resultar em valores inteiros positivos ou negativos. No final da temporada regular, a classificação dos times dentro de cada conferência funciona pelo sistema de “pontos corridos”, sendo o primeiro (último) colocado o time de maior (menor) pontuação dentro de

cada conferência. Ao todo, os 10 melhores colocados de cada conferência avançam para as próximas fases (*play-in* e/ou *playoffs*). Os seis primeiros classificados de cada conferência seguem diretamente para a fase *playoffs*, e as equipes entre a sétima até a décima colocação disputam no *play-in* mais duas vagas restantes, enfrentando jogos únicos de mata-mata, para definir os dois últimos classificados a disputar o *playoffs*.

Neste sentido, dois conjuntos de dados foram considerados neste estudo, que correspondem aos valores das diferenças entre partidas ganhas e partidas perdidas na temporada regular de 2022-2023 dos times em cada conferência (dados obtidos de [Basketball Reference \(2024\)](#)). A Tabela 16 apresenta a distribuição de frequências da pontuação das equipes no final da temporada na sua respectiva conferência (Leste e Oeste), além das medidas descritivas como a média amostral (\bar{y}) e o desvio-padrão (SD). Analisando esta Tabela, podemos ter uma visão das performances das equipes de cada conferência, com relação às suas vitórias e derrotas durante a temporada regular. Interpretando essas pontuações, como exemplo um time da conferência Leste, o time com pontuação -48 significa que este perdeu 48 jogos a mais do que ganhou (dos 82 jogos, ganhou 17 e perdeu 65). Outro exemplo, de pontuação positiva, o time dessa mesma conferência com pontuação 34 significa que este ganhou 34 jogos a mais do que perdeu (dos 82 jogos, ganhou 58 e perdeu 24). A conferência Leste apresenta uma pontuação média de 1,48 jogos. Isso significa que, em média, as equipes desta conferência ganharam mais jogos do que perderam. Porém o SD é de 22,43 jogos, sugerindo que as diferenças entre as equipes variam consideravelmente em torno da média. Na conferência Oeste, a pontuação média é de -1,47 jogos, isto é, em média, as equipes perderam cerca de um jogo a mais do que ganharam. Mesmo apresentando um valor alto para o desvio padrão, de 18,02 jogos, a variação de pontuações foi um pouco menor do que em relação a conferência Leste. A diferença nas amplitudes das pontuações das Conferência Leste e a Oeste indica uma distinção de comportamento das equipes, de modo que a Conferência Leste (Oeste) apresenta uma variabilidade maior (menor).

Tabela 16 – Distribuição de frequências e estatísticas descritivas da diferença entre jogos ganhos e perdidos (a pontuação) das equipes de cada conferência da NBA na temporada regular de 2022-2023.

y_i	Conjunto de dados																			\bar{y}	SD		
	-48	-38	-28	-16	-14	-12	-8	-6	-2	0	2	4	6	8	12	14	20	24	26			32	34
Geral	1	2	1	1	1	2	1	1	2	2	2	1	3	2	1	1	2	1	1	1	1	0	20,05
Leste	f_i	1		1		1	2		1	2			1	1	1		1		1	1	1	1,48	22,43
Oeste		2			1			1	1	1	2	1	2	1		1	1	1				-1,47	18,02

Como variável explicativa X , foi considerada a diferença entre a quantidade de jogos ganhos com no mínimo 10 pontos de vantagem (J_G) e jogos perdidos com no mínimo 10 pontos de desvantagem (J_P) pela equipe, ou seja, $X = J_G - J_P$. A Figura 8 ilustra graficamente os valores da variável Y , as pontuações (diferenças entre jogos ganhos e perdidos das equipes), e também da covariável $X = J_G - J_P$, por conferência, na temporada regular 2022-2023 da NBA. Analisando os gráficos, podemos notar que, em sua maioria, as equipes com pontuações positivas sobressaem por vencer mais jogos com uma vantagem mínima de 10 pontos do que por perder com uma

desvantagem equivalente. Por outro lado, equipes com pontuações negativas destacam-se por perder mais jogos com uma desvantagem mínima de 10 pontos do que por vencer com uma desvantagem equivalente.

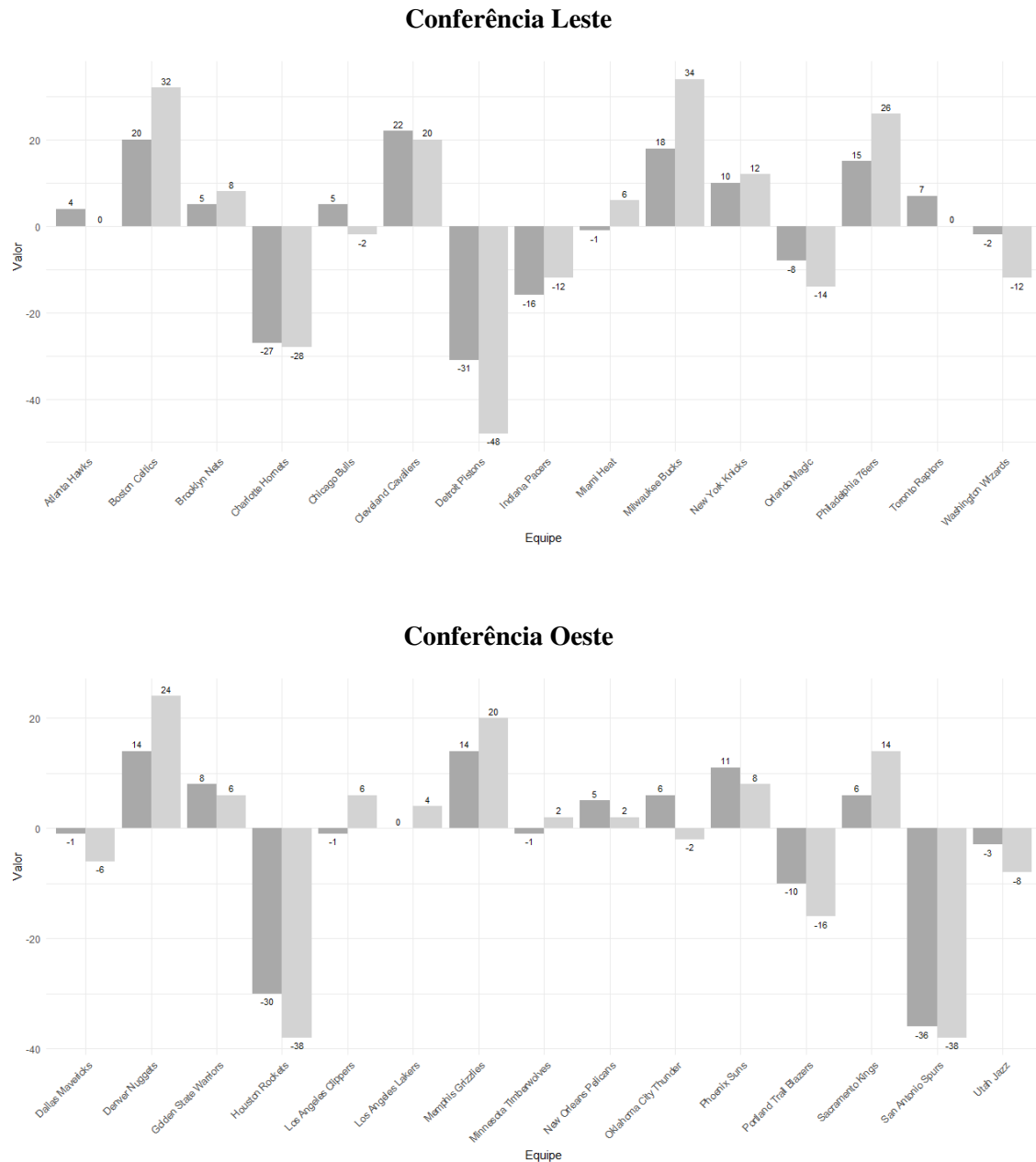


Figura 8 – Ilustrações em gráficos de barras das pontuações (cinza claro) e também da covariável $X = J_G - J_P$ (cinza escuro), por conferência, na temporada regular 2022-2023 da NBA.

Para determinar o valor mais apropriado de k -modificação nos dados observados para o ajuste do modelo k -MS, fizemos a suposição de alguns valores: na conferência Leste, consideramos ajustes com $k = -48, -12, e 0$; e na conferência Oeste, consideramos $k = -38, 2, e 6$. Tais valores foram definidos com base na distribuição de frequências das pontuações (Tabela 16). Por meio do procedimento bayesiano com o algoritmo SGHMC, os resultados dos modelos ajustados, especificamente a estimativa e intervalo HPD com 95% de credibilidade de cada parâmetro,

para cada conferência, podem ser encontrados na Tabela 17. É importante ressaltar que, na conferência Oeste, a covariável X não demonstrou significância para o parâmetro correspondente à k -modificação (p ou ω) nos modelos k -MS ajustados, considerando tanto $k = 2$ quanto $k = 6$.

Tabela 17 – Estimativas (e HPD 95%) dos parâmetros dos modelos k -MS ajustados aos dados de pontuação da temporada regular 2022-20023 da NBA, considerando diferentes valores de k .

k	Estimativas (HPD 95%)				
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\phi}$
Conferência Leste					
-48	2,73 (1,99; 3,22)	0,90 (0,12; 1,29)	2,53 (1,96; 3,26)	0,67 (0,56; 0,78)	69,19 (66,17;72,96)
-12	1,97 (1,29;2,66)	0,40 (0,05;1,13)	0,25 (0,05;0,96)	1,36 (1,17;1,58)	58,01 (56,93;59,13)
0	2,00 (1,32;2,70)	-0,13 (-0,71;-0,03)	0,56 (0,16;1,21)	1,35 (1,16;1,56)	57,84 (56,79;58,91)
Conferência Oeste					
-38	2,67 (1,96;3,17)	0,23 (0,07;0,41)	0,91 (0,23;1,53)	1,09 (0,83;139)	61,06 (59,47;63,12)
2	1,97 (1,29;2,67)	-0,04 (-0,17;0,07)	0,39 (0,11;1,06)	1,22 (1,01;1,45)	58,10 (57,07; 59,15)
6	2,00 (1,34;2,72)	-0,06 (-0,19;0,06)	0,32 (0,09;1,01)	1,23 (1,02;1,45)	58,40 (57,35;59,48)

Após ajustarmos o modelo k -MS para diferentes valores de k , procedemos ao cálculo dos critérios de seleção de modelos $WAIC$, $EAIC$ e $EBIC$. A Tabela 18 exibe os resultados desses critérios, evidenciando que o modelo k -MS com $k = -12$ foi identificado como o mais bem ajustado para a conferência Leste, enquanto para a conferência Oeste, o valor indicado é $k = -38$ como o mais adequado.

Tabela 18 – Critérios de seleção dos modelos k -MS ajustados aos dados de pontuação da temporada regular 2022-20023 da NBA, considerando diferentes valores de k .

k	$WAIC$	$EAIC$	$EBIC$
Conferência Leste			
-48	115,65	120,01	124,91
-12	101,75	108,15	113,73
0	102,97	109,53	115,03
Conferência Oeste			
-38	88,45	96,37	100,87
2	97,82	104,52	109,94
6	96,91	103,70	109,08

A Figura 9 apresenta os gráficos obtidos a partir do ajuste do modelo k -MS para os valores de k identificados como mais adequados, para cada conferência. Os gráficos (A1) e

(A2) representam as estimativas de p e μ em relação à covariável X para a conferência Leste (modelo k -MS com $k = -12$), enquanto os gráficos (B1) e (B2) ilustram, respectivamente, as estimativas de p e μ em relação à mesma covariável para a conferência Oeste (modelo k -MS com $k = -38$). Para ambas as conferências, os valores estimados de p indicam inflação no ponto k (inflação de $k = -12$ na conferência Leste e inflação de $k = -38$ na conferência Oeste). O bom ajuste do modelo k -MS aos dados em ambos os casos, evidenciado pela concordância entre o comportamento da média ajustada e os dados observados, sugere uma relação significativa entre a covariável e a variável resposta, indicando que quanto maior o valor de X , maiores as pontuações das equipes.

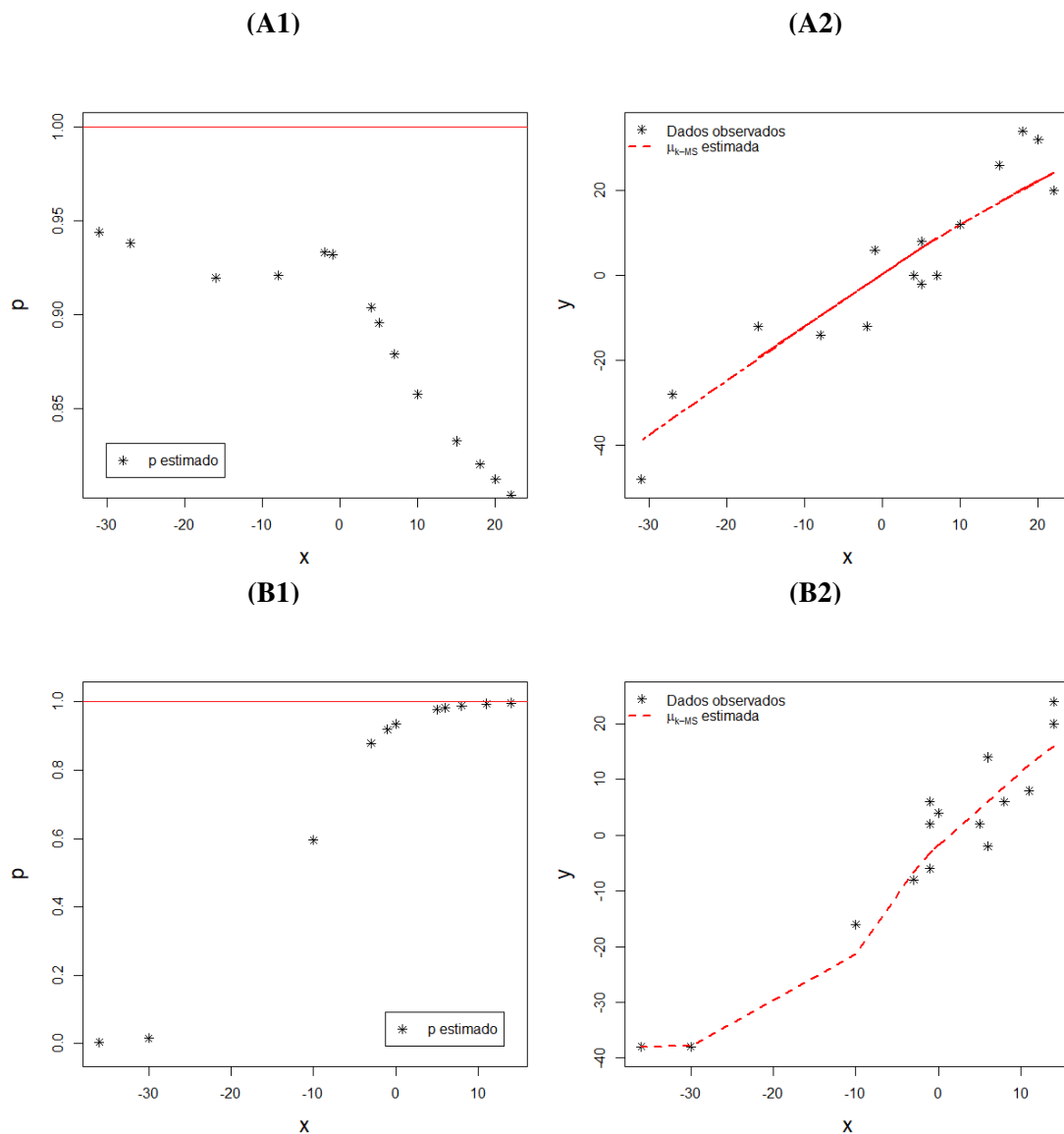


Figura 9 – Estimativas de p e de μ_{k-MS} , versus a covariável, X obtidas a partir do modelo k -MS ajustado, com: (A1) e (A2) para os dados da conferência Leste, modelo k -MS com $k = -12$; e (B1) e (B2) para os dados da conferência Oeste, modelo k -MS com $k = -38$.

A fim de validar os resultados obtidos a partir da análise destes dados, obtivemos a

pontuação média predita para os dados da temporada regular atual, de 2023-2024, e comparamos com os seus respectivos valores reais. A Figura 10 apresenta a pontuação da temporada atual e o predito em função da covariável X (também da temporada atual) para cada conferência. É importante destacar que o erro médio de predição para a conferência Leste foi de 7 pontos, enquanto para a conferência Oeste foi de 8 pontos.

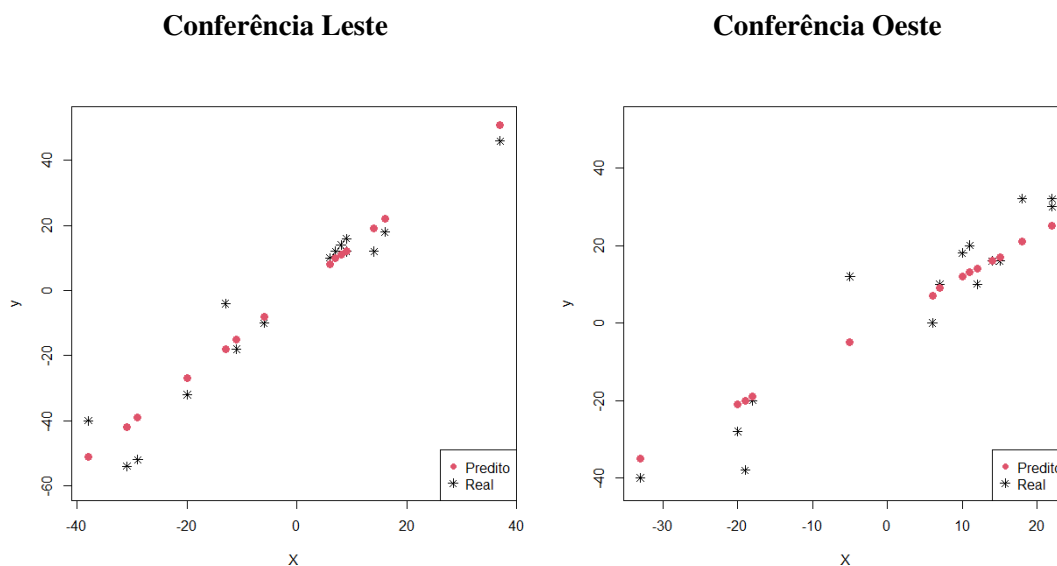


Figura 10 – Pontuação final da temporada regular da NBA 2023-2024 *versus* a diferença entre a quantidade de jogos ganhos com no mínimo 10 pontos de vantagem e jogos perdidos com no mínimo 10 pontos de desvantagem, para as conferências Leste e Oeste.

A partir desta análise, considerando que a temporada regular é determinante para a seleção dos times que avançam e competem pelo título da NBA, o uso do desempenho das equipes, conforme descrito pela covariável, oferece uma vantagem adicional na previsão das chances de chegada aos *playoffs*. Isso significa identificar com maior precisão as equipes com maior probabilidade de se classificarem entre as 10 primeiras posições na temporada regular da NBA.

MODELO DE REGRESSÃO SKELLAM k_1 E k_2 INFLACIONADO

5.1 Modelo Skellam k_1 e k_2 Inflacionado

Seja Z uma variável aleatória tal que $Z \sim \mathbf{k}\text{-IS}(\mu, \phi, \mathbf{p})$ como apresentado em (3.8). Considere os conjuntos de covariáveis, x_1, \dots, x_{q_0} e v_{11}, \dots, v_{1q_1} , $\iota = 1, 2$, associados respectivamente aos parâmetros μ e p_ι . Reescrevendo a FMP dada em (3.8), obtemos o seguinte modelo de regressão Skellam \mathbf{k} -Inflacionado:

$$\begin{aligned} \pi_{\mathbf{k}\text{-IS}}(z; \mu(\mathbf{x}), \phi, p_\iota(\mathbf{v}_\iota)) &= (1 - p_1(\mathbf{v}_1))I_{\{k_1\}}(z) + (1 - p_2(\mathbf{v}_2))I_{\{k_2\}}(z) + \\ & p_0(\mathbf{v}_1, \mathbf{v}_2)\pi_S(z; \mu(\mathbf{x}), \phi), \quad z \in \mathbb{Z}, \end{aligned} \quad (5.1)$$

em que $\mathbf{x} = (1, x_1, \dots, x_{q_0})$ e $\mathbf{v}_\iota = (1, v_{\iota 1}, \dots, v_{\iota q_\iota})$ são os vetores das covariáveis relacionadas aos parâmetros da regressão; e $p_0(\mathbf{v}_1, \mathbf{v}_2) = p_1(\mathbf{v}_1) + p_2(\mathbf{v}_2) - 1$ sob as condições:

C1. $p_\iota(\mathbf{v}_\iota) \in \mathcal{P} = (0, 1)$, $\forall \iota = 0, 1, 2$;

C2. $(p_1(\mathbf{v}_1) + p_2(\mathbf{v}_2) - 1) + \sum_{\iota=1}^2 (1 - p(\mathbf{v}_\iota)) = 1$.

Para incorporar covariáveis no parâmetro μ , consideramos a função de ligação identidade, $h_\mu(\mu) = \mu(\mathbf{x})$, que relaciona μ ao preditor linear $\mathbf{x}\boldsymbol{\beta}$, $h_\mu(\mu) = \mu(\mathbf{x}) = \mathbf{x}\boldsymbol{\beta}$, sendo $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{q_0})^\top$ o vetor de parâmetros. Para cada parâmetro p_ι , assegurando as restrições paramétricas, consideramos a função de ligação *logit*. Dessa forma, temos

$$h_{p_\iota}(p_\iota) = \log \left(\frac{p_\iota(\mathbf{v}_\iota)}{1 - p_\iota(\mathbf{v}_\iota)} \right) = \mathbf{v}_\iota \boldsymbol{\alpha}_\iota, \quad \iota = 1, 2, \quad (5.2)$$

que relaciona cada p_ι ao preditor linear com vetor de parâmetros $\boldsymbol{\alpha}_\iota = (\alpha_{\iota 0}, \alpha_{\iota 1}, \dots, \alpha_{\iota q_\iota})^\top$ e o vetor de covariáveis $\mathbf{v}_\iota = (1, v_{\iota 1}, \dots, v_{\iota q_\iota})$.

Ao explicitar $p_l(\mathbf{v}_l)$ a partir da Equação (5.2), obtemos:

$$p_l(\mathbf{v}_l) = \left(\frac{e^{\mathbf{v}_l \boldsymbol{\alpha}_l}}{1 + e^{\mathbf{v}_l \boldsymbol{\alpha}_l}} \right).$$

Portanto, o modelo de regressão \mathbf{k} -IS possui $(q_0 + q_1 + q_2 + 4)$ parâmetros desconhecidos a serem estimados, correspondentes aos vetores $\boldsymbol{\beta}$, $\boldsymbol{\alpha}_1$ e $\boldsymbol{\alpha}_2$ e ao parâmetro ϕ .

5.1.1 Modelo \mathbf{k} -IS e sua Versão Hurdle

A versão *hurdle* da distribuição Skellam k_1 e k_2 Inflacionada apresentada em (3.11), tem o modelo de regressão escrito como

$$\begin{aligned} \pi_{\mathbf{k}\text{-IS}}(z; \boldsymbol{\mu}(\mathbf{x}), \phi, \boldsymbol{\omega}_l(\mathbf{v}_l)) &= (1 - \omega_1(\mathbf{v}_1))I_{\{k_1\}}(z) + (1 - \omega_2(\mathbf{v}_2))I_{\{k_2\}}(z) + \\ &(\omega_1(\mathbf{v}_1) + \omega_2(\mathbf{v}_2) - 1)\pi_{\mathbf{k}\text{-SS}}(z; \boldsymbol{\mu}(\mathbf{x}), \phi), \quad z \in \mathbb{Z}. \end{aligned} \quad (5.3)$$

Os espaços paramétricos de $\omega_1(\mathbf{v}_1)$ e $\omega_2(\mathbf{v}_2)$ são dados por:

$$p_1(\mathbf{v}_1) - \pi_s(k_1; \boldsymbol{\mu}(\mathbf{x}), \phi) < \omega_1(\mathbf{v}_1) < p_1(\mathbf{v}_1)$$

e

$$p_2(\mathbf{v}_2) - \pi_s(k_2; \boldsymbol{\mu}(\mathbf{x}), \phi) < \omega_2(\mathbf{v}_2) < p_2(\mathbf{v}_2).$$

Nesta parametrização, o vetor de parâmetros $\boldsymbol{\alpha}_l = (\alpha_{l0}, \alpha_{l1}, \dots, \alpha_{lq_l})^\top$ e o vetor de covariáveis $\mathbf{v}_l = (1, v_{l1}, \dots, v_{lq_l})$ estão associados ao parâmetro ω_l , $l = 1, 2$.

5.1.2 Função de Verossimilhança

Seja $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)$ um vetor de n variáveis aleatórias independentes, com $Z_i \sim \mathbf{k}$ -IS($\mu_i, \phi, \mathbf{p}_i$), $i = 1, 2, \dots, n$, e seja $\mathbf{z} = (z_1, z_2, \dots, z_n)$ o vetor de observações associado a \mathbf{Z} . Considere $\mathbf{x}_i = (1, x_{1i}, \dots, x_{q_0i})$, $\mathbf{v}_{1i} = (1, v_{11i}, \dots, v_{1q_1i})$ e $\mathbf{v}_{2i} = (1, v_{21i}, \dots, v_{2q_2i})$ os vetores de covariáveis associados a cada z_i e, conseqüentemente, $\mu_i = \boldsymbol{\mu}(\mathbf{x}_i)$, $p_{1i} = p_1(\mathbf{v}_{1i})$ e $p_{2i} = p_2(\mathbf{v}_{2i})$. Denotamos por $\mathcal{D} = \{z_i, \mathbf{x}_i, \mathbf{v}_{1i}, \mathbf{v}_{2i} : \forall i = 1, \dots, n\}$ as informações relacionadas às n observações.

Assim, ao reescrever o modelo $\pi_{\mathbf{k}\text{-IS}}(z; \boldsymbol{\mu}(\mathbf{x}), \phi, p_l(\mathbf{v}_l))$, $l = 1, 2$, como

$$\begin{aligned} \pi_{\mathbf{k}\text{-IS}}(z; \boldsymbol{\mu}(\mathbf{x}), \phi, p_l(\mathbf{v}_l)) &= \left\{ \pi_{\mathbf{k}\text{-IS}}(k_1; \boldsymbol{\mu}(\mathbf{x}), \phi, p_l(\mathbf{v}_l)) \right\}^{I_{\{k_1\}}(z)} \times \left\{ \pi_{\mathbf{k}\text{-IS}}(k_2; \boldsymbol{\mu}(\mathbf{x}), \phi, p_l(\mathbf{v}_l)) \right\}^{I_{\{k_2\}}(z)} \times \\ &\left\{ \pi_{\mathbf{k}\text{-IS}}(z; \boldsymbol{\mu}(\mathbf{x}), \phi, p_l(\mathbf{v}_l)) \right\}^{(1 - I_{\{k_1\}}(z) - I_{\{k_2\}}(z))}, \end{aligned}$$

temos que a função de verossimilhança associada a \mathbf{z} é dada por:

$$\begin{aligned} \mathcal{L}(\boldsymbol{\mu}(\mathbf{x}_i), \phi, p_1(\mathbf{v}_{1i}), p_2(\mathbf{v}_{2i}); \mathcal{D}) &= \prod_{i=1}^n \left\{ \left[(1 - p_1(\mathbf{v}_{1i})) + (p_1(\mathbf{v}_{1i}) + p_2(\mathbf{v}_{2i}) - 1)\pi_s(k_1; \boldsymbol{\mu}(\mathbf{x}_i), \phi) \right]^{I_{\{k_1\}}(z_i)} \times \right. \\ &\left[(1 - p_2(\mathbf{v}_{2i})) + (p_1(\mathbf{v}_{1i}) + p_2(\mathbf{v}_{2i}) - 1)\pi_s(k_2; \boldsymbol{\mu}(\mathbf{x}_i), \phi) \right]^{I_{\{k_2\}}(z_i)} \times \\ &\left. \left[(p_1(\mathbf{v}_{1i}) + p_2(\mathbf{v}_{2i}) - 1)\pi_s(z_i; \boldsymbol{\mu}(\mathbf{x}_i), \phi) \right]^{(1 - I_{\{k_1\}}(z_i) - I_{\{k_2\}}(z_i))} \right\}. \end{aligned}$$

Assim, o logaritmo natural da função de verossimilhança é

$$\begin{aligned} \ell(\boldsymbol{\mu}(\mathbf{x}_i), \phi, p_1(\mathbf{v}_{1i}), p_2(\mathbf{v}_{2i}); \mathcal{D}) &= \sum_{i=1}^n \left\{ I_{\{k_1\}}(z_i) \log \left((1 - p_1(\mathbf{v}_{1i})) + (p_1(\mathbf{v}_{1i}) + p_2(\mathbf{v}_{2i}) - 1) \pi_s(k_1; \boldsymbol{\mu}(\mathbf{x}_i), \phi) \right) + \right. \\ & I_{\{k_2\}}(z_i) \log \left((1 - p_2(\mathbf{v}_{2i})) + (p_1(\mathbf{v}_{1i}) + p_2(\mathbf{v}_{2i}) - 1) \pi_s(k_2; \boldsymbol{\mu}(\mathbf{x}_i), \phi) \right) + \\ & \left. (1 - I_{\{k_1\}}(z_i) - I_{\{k_2\}}(z_i)) \log \left((p_1(\mathbf{v}_{1i}) + p_2(\mathbf{v}_{2i}) - 1) \pi_s(z_i; \boldsymbol{\mu}(\mathbf{x}_i), \phi) \right) \right\}, \end{aligned}$$

que, ao substituir $\boldsymbol{\mu}(\mathbf{x}_i) = \mathbf{x}_i \boldsymbol{\beta}$ e $p_t(\mathbf{v}_{ti}) = \frac{e^{\mathbf{v}_{ti} \boldsymbol{\alpha}_t}}{1 + e^{\mathbf{v}_{ti} \boldsymbol{\alpha}_t}}$ ($t = 1, 2$) em $\ell(\boldsymbol{\mu}(\mathbf{x}_i), \phi, p_1(\mathbf{v}_{1i}), p_2(\mathbf{v}_{2i}); \mathcal{D})$, obtemos:

$$\begin{aligned} \ell(\boldsymbol{\beta}, \phi, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2; \mathcal{D}) &= \sum_{i=1}^n \left\{ I_{\{k_1\}}(z_i) \log \left(\left(1 - \frac{e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1}}{1 + e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1}} \right) + \left(\frac{e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1}}{1 + e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1}} + \frac{e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2}}{1 + e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2}} - 1 \right) \pi_s(k_1; \mathbf{x}_i \boldsymbol{\beta}, \phi) \right) + \right. \\ & I_{\{k_2\}}(z_i) \log \left(\left(1 - \frac{e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2}}{1 + e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2}} \right) + \left(\frac{e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1}}{1 + e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1}} + \frac{e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2}}{1 + e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2}} - 1 \right) \pi_s(k_2; \mathbf{x}_i \boldsymbol{\beta}, \phi) \right) + \\ & \left. \left(1 - I_{\{k_1\}}(z_i) - I_{\{k_2\}}(z_i) \right) \log \left(\left(\frac{e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1}}{1 + e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1}} + \frac{e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2}}{1 + e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2}} - 1 \right) \pi_s(z_i; \mathbf{x}_i \boldsymbol{\beta}, \phi) \right) \right\}. \end{aligned}$$

Com algumas manipulações algébricas alcançamos a seguinte equação para a função log-verossimilhança $\ell(\boldsymbol{\beta}, \phi, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2; \mathcal{D})$:

$$\begin{aligned} \ell(\boldsymbol{\beta}, \phi, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2; \mathcal{D}) &= \sum_{i=1}^n \left\{ I_{\{k_1\}}(z_i) \log (1 + e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2} + (e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1} e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2} - 1) \pi_s(k_1; \mathbf{x}_i \boldsymbol{\beta}, \phi)) + \right. \\ & I_{\{k_2\}}(z_i) \log (1 + e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1} + (e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1} e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2} - 1) \pi_s(k_2; \mathbf{x}_i \boldsymbol{\beta}, \phi)) + \\ & \left(1 - I_{\{k_1\}}(z_i) - I_{\{k_2\}}(z_i) \right) \log ((e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1} e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2} - 1) \pi_s(z_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)) - \\ & \left. \log(1 + e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1}) - \log(1 + e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2}) \right\}. \end{aligned} \quad (5.4)$$

5.1.3 Função de Verossimilhança para a Versão Hurdle

Considere agora que as n variáveis aleatórias independentes $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)$ tem, cada uma, a distribuição $\mathbf{k} - IS$ escrita na versão hurdle, isto é, $Z_i \sim \mathbf{k} - IS(\mu_i, \phi, \boldsymbol{\omega}_i)$, para todo $i = 1, 2, \dots, n$. Seja $\mathbf{z} = (z_1, z_2, \dots, z_n)$ o vetor de observações associado a \mathbf{Z} e $\mathbf{x}_i = (1, x_{1i}, \dots, x_{q_0i})$, $\mathbf{v}_{1i} = (1, v_{11i}, \dots, v_{1q_1i})$ e $\mathbf{v}_{2i} = (1, v_{21i}, \dots, v_{2q_2i})$ os vetores de covariáveis associados a cada z_i e, conseqüentemente, $\mu_i = \boldsymbol{\mu}(\mathbf{x}_i)$, $\boldsymbol{\omega}_{1i} = \boldsymbol{\omega}_1(\mathbf{v}_{1i})$ e $\boldsymbol{\omega}_{2i} = \boldsymbol{\omega}_2(\mathbf{v}_{2i})$. Denotamos por $\mathcal{D} = \{z_i, \mathbf{x}_i, \mathbf{v}_{1i}, \mathbf{v}_{2i} : \forall i = 1, \dots, n\}$ as informações relacionadas a todos os indivíduos. Assim, considerando o modelo (5.3), a função de verossimilhança associada a \mathbf{z} é dada por:

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\mu}(\mathbf{x}_i), \phi, \boldsymbol{\omega}_1(\mathbf{v}_{1i}), \boldsymbol{\omega}_2(\mathbf{v}_{2i}); \mathcal{D}) &= \prod_{i=1}^n \left\{ (1 - \boldsymbol{\omega}_1(\mathbf{v}_{1i})) I_{\{k_1\}}(z_i) + (1 - \boldsymbol{\omega}_2(\mathbf{v}_{2i})) I_{\{k_2\}}(z_i) + \right. \\
&\quad \left. (\boldsymbol{\omega}_1(\mathbf{v}_{1i}) + \boldsymbol{\omega}_2(\mathbf{v}_{2i}) - 1) \pi_{k-SS}(z_i; \boldsymbol{\mu}(\mathbf{x}_i), \phi) \right\} \\
&= \prod_{i=1}^n \left\{ (1 - \boldsymbol{\omega}_1(\mathbf{v}_{1i}))^{I_{\{k_1\}}(z_i)} \times (1 - \boldsymbol{\omega}_2(\mathbf{v}_{2i}))^{I_{\{k_2\}}(z_i)} \times \left((\boldsymbol{\omega}_1(\mathbf{v}_{1i}) + \boldsymbol{\omega}_2(\mathbf{v}_{2i}) - 1) \times \right. \right. \\
&\quad \left. \left. \frac{\pi_S(z_i; \boldsymbol{\mu}(\mathbf{x}_i), \phi)}{1 - \pi_S(k_1; \boldsymbol{\mu}(\mathbf{x}_i), \phi) - \pi_S(k_2; \boldsymbol{\mu}(\mathbf{x}_i), \phi)} \right)^{(1 - I_{\{k_1\}}(z_i) - I_{\{k_2\}}(z_i))} \right\},
\end{aligned}$$

e o logaritmo natural dessa função é

$$\begin{aligned}
\ell(\boldsymbol{\mu}(\mathbf{x}_i), \phi, \boldsymbol{\omega}_1(\mathbf{v}_{1i}), \boldsymbol{\omega}_2(\mathbf{v}_{2i}); \mathcal{D}) &= \sum_{i=1}^n \left\{ I_{\{k_1\}}(z_i) \log(1 - \boldsymbol{\omega}_1(\mathbf{v}_{1i})) + I_{\{k_2\}}(z_i) \log(1 - \boldsymbol{\omega}_2(\mathbf{v}_{2i})) + \right. \\
&\quad (1 - I_{\{k_1\}}(z_i) - I_{\{k_2\}}(z_i)) \left[\log(\boldsymbol{\omega}_1(\mathbf{v}_{1i}) + \boldsymbol{\omega}_2(\mathbf{v}_{2i}) - 1) + \right. \\
&\quad \left. \left. \log\left(\frac{\pi_S(z_i; \boldsymbol{\mu}(\mathbf{x}_i), \phi)}{1 - \pi_S(k_1; \boldsymbol{\mu}(\mathbf{x}_i), \phi) - \pi_S(k_2; \boldsymbol{\mu}(\mathbf{x}_i), \phi)}\right) \right] \right\}.
\end{aligned}$$

Ao substituir

$$\boldsymbol{\omega}_l(\mathbf{v}_{li}) = \left(\frac{e^{\mathbf{v}_{li} \boldsymbol{\alpha}_l}}{1 + e^{\mathbf{v}_{li} \boldsymbol{\alpha}_l}} \right) - \left(\frac{e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1} e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2} - 1}{(1 + e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1})(1 + e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2})} \right) \pi_S(k_l; \boldsymbol{\mu}(\mathbf{x}_i), \phi), \quad l = 1, 2,$$

em $\ell(\boldsymbol{\mu}(\mathbf{x}_i), \phi, \boldsymbol{\omega}_1(\mathbf{v}_{1i}), \boldsymbol{\omega}_2(\mathbf{v}_{2i}); \mathcal{D})$, temos:

$$\begin{aligned}
\ell(\boldsymbol{\mu}(\mathbf{x}_i), \phi, \boldsymbol{\omega}_1(\mathbf{v}_{1i}), \boldsymbol{\omega}_2(\mathbf{v}_{2i}); \mathcal{D}) &= \sum_{i=1}^n \left\{ I_{\{k_1\}}(z_i) \log\left(1 + e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2} + (e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1} e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2} - 1) \pi_S(k_1; \boldsymbol{\mu}(\mathbf{x}_i), \phi)\right) + \right. \\
&\quad I_{\{k_2\}}(z_i) \log\left(1 + e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1} + (e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1} e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2} - 1) \pi_S(k_2; \boldsymbol{\mu}(\mathbf{x}_i), \phi)\right) + \\
&\quad (1 - I_{\{k_1\}}(z_i) - I_{\{k_2\}}(z_i)) \log\left((e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1} e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2} - 1) \pi_S(z_i; \boldsymbol{\mu}(\mathbf{x}_i), \phi)\right) - \\
&\quad \left. \log(1 + e^{\mathbf{v}_{1i} \boldsymbol{\alpha}_1}) - \log(1 + e^{\mathbf{v}_{2i} \boldsymbol{\alpha}_2}) \right\}.
\end{aligned}$$

Diferentemente do modelo k -MS descrito no Capítulo 4, em que a versão *hurdle* proporciona como vantagem a separação das probabilidades de ocorrência de (algumas) observações e facilidades no processo de estimação, aqui a complexidade adicional introduzida para esse contexto devido às restrições paramétricas, não viabiliza o seu uso. A versão tradicional produz uma log-verossimilhança mais tratável, e isso, então, nos guiou para a utilização do modelo com esta parametrização para o procedimento de estimação dos parâmetros.

5.1.4 Vetor Escore

Derivando a função $\ell(\boldsymbol{\beta}, \phi, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2; \mathcal{D})$ dada em (5.4) em relação a cada parâmetro, obtemos o gradiente vetorial do logaritmo natural da função de verossimilhança, isto é, o vetor escore $\mathbf{U} = (\mathbf{U}_\beta, U_\phi, \mathbf{U}_{\boldsymbol{\alpha}_1}, \mathbf{U}_{\boldsymbol{\alpha}_2})^\top$, sendo $\mathbf{U}_\beta = (U_{\beta_0}, U_{\beta_1}, \dots, U_{\beta_{q_0}})$, $\mathbf{U}_{\boldsymbol{\alpha}_1} = (U_{\alpha_{10}}, U_{\alpha_{11}}, \dots, U_{\alpha_{1q_1}})$ e $\mathbf{U}_{\boldsymbol{\alpha}_2} = (U_{\alpha_{20}}, U_{\alpha_{21}}, \dots, U_{\alpha_{2q_2}})$.

Os elementos do vetor escore relacionados a $\boldsymbol{\beta}$, são obtidos calculando a derivada parcial do logaritmo natural da função de verossimilhança, $U_{\beta_j} = \frac{\partial \ell(\boldsymbol{\beta}, \phi, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2; \mathcal{D})}{\partial \beta_j}$. Assim, temos:

$$U_{\beta_j} = \sum_{i=1}^n \left\{ I_{\{k_1\}}(y_i) \left(\frac{(e^{v_{1i}\boldsymbol{\alpha}_1} e^{v_{2i}\boldsymbol{\alpha}_2} - 1)(g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_1} f_{\beta_j}(k_1; \mathbf{x}_i\boldsymbol{\beta}, \phi) + g_{\beta_j}(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_1} f(k_1; \mathbf{x}_i\boldsymbol{\beta}, \phi))}{1 + e^{v_{2i}\boldsymbol{\alpha}_2} + (e^{v_{1i}\boldsymbol{\alpha}_1} e^{v_{2i}\boldsymbol{\alpha}_2} - 1)g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_1} f(k_1; \mathbf{x}_i\boldsymbol{\beta}, \phi)} \right) + \right. \\ \left. I_{\{k_2\}}(y_i) \left(\frac{(e^{v_{1i}\boldsymbol{\alpha}_1} e^{v_{2i}\boldsymbol{\alpha}_2} - 1)(g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_2} f_{\beta_j}(k_2; \mathbf{x}_i\boldsymbol{\beta}, \phi) + g_{\beta_j}(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_2} f(k_2; \mathbf{x}_i\boldsymbol{\beta}, \phi))}{1 + e^{v_{1i}\boldsymbol{\alpha}_1} + (e^{v_{1i}\boldsymbol{\alpha}_1} e^{v_{2i}\boldsymbol{\alpha}_2} - 1)g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_2} f(k_2; \mathbf{x}_i\boldsymbol{\beta}, \phi)} \right) + \right. \\ \left. (1 - I_{\{k_1\}}(y_i) - I_{\{k_2\}}(y_i)) \left(\frac{f_{\beta_j}(k_2; \mathbf{x}_i\boldsymbol{\beta}, \phi)}{f(k_2; \mathbf{x}_i\boldsymbol{\beta}, \phi)} + \frac{g_{\beta_j}(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_2}}{g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_2}} \right) \right\}, \forall j = 0, 1, \dots, q_0,$$

em que $g_{\beta_j}(\mathbf{x}_i\boldsymbol{\beta}, \phi)^*$ e $f_{\beta_j}(*; \mathbf{x}_i\boldsymbol{\beta}, \phi)$ ($*$ = k_1, k_2, z_i) denotam, respectivamente, as derivadas das funções $g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^*$ e $f(*; \mathbf{x}_i\boldsymbol{\beta}, \phi)$ em relação a β_j .

O elemento do vetor escore relacionado a ϕ , $U_\phi = \frac{\partial \ell(\boldsymbol{\beta}, \phi, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2; \mathcal{D})}{\partial \phi}$, é dado por:

$$U_\phi = \sum_{i=1}^n \left\{ I_{\{k_1\}}(y_i) \left(\frac{(e^{v_{1i}\boldsymbol{\alpha}_1} e^{v_{2i}\boldsymbol{\alpha}_2} - 1)(g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_1} f_\phi(k_1; \mathbf{x}_i\boldsymbol{\beta}, \phi) + g_\phi(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_1} f(k_1; \mathbf{x}_i\boldsymbol{\beta}, \phi))}{1 + e^{v_{2i}\boldsymbol{\alpha}_2} + (e^{v_{1i}\boldsymbol{\alpha}_1} e^{v_{2i}\boldsymbol{\alpha}_2} - 1)g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_1} f(k_1; \mathbf{x}_i\boldsymbol{\beta}, \phi)} \right) + \right. \\ \left. I_{\{k_2\}}(y_i) \left(\frac{(e^{v_{1i}\boldsymbol{\alpha}_1} e^{v_{2i}\boldsymbol{\alpha}_2} - 1)(g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_2} f_\phi(k_2; \mathbf{x}_i\boldsymbol{\beta}, \phi) + g_\phi(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_2} f(k_2; \mathbf{x}_i\boldsymbol{\beta}, \phi))}{1 + e^{v_{1i}\boldsymbol{\alpha}_1} + (e^{v_{1i}\boldsymbol{\alpha}_1} e^{v_{2i}\boldsymbol{\alpha}_2} - 1)g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_2} f(k_2; \mathbf{x}_i\boldsymbol{\beta}, \phi)} \right) + \right. \\ \left. (1 - I_{\{k_1\}}(y_i) - I_{\{k_2\}}(y_i)) \left(\frac{f_\phi(k_2; \mathbf{x}_i\boldsymbol{\beta}, \phi)}{f(k_2; \mathbf{x}_i\boldsymbol{\beta}, \phi)} + \frac{g_\phi(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_2}}{g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_2}} \right) \right\},$$

em que $g_\phi(\mathbf{x}_i\boldsymbol{\beta}, \phi)^*$ e $f_\phi(*; \mathbf{x}_i\boldsymbol{\beta}, \phi)$ ($*$ = k_1, k_2, z_i) denotam, respectivamente, as derivadas das funções $g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^*$ e $f(*; \mathbf{x}_i\boldsymbol{\beta}, \phi)$ em relação a ϕ .

Já os elementos do vetor escore relacionados a $\boldsymbol{\alpha}_1$, $U_{\alpha_{1j}} = \frac{\partial \ell(\boldsymbol{\beta}, \phi, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2; \mathcal{D})}{\partial \alpha_{1j}}$, são dados por:

$$U_{\alpha_{1j}} = \sum_{i=1}^n \left\{ I_{\{k_1\}}(y_i) \left(\frac{(e^{v_{1i}\boldsymbol{\alpha}_1} e^{v_{2i}\boldsymbol{\alpha}_2})g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_1} f(k_1; \mathbf{x}_i\boldsymbol{\beta}, \phi) \cdot v_{1ij}}{1 + e^{v_{2i}\boldsymbol{\alpha}_2} + (e^{v_{1i}\boldsymbol{\alpha}_1} e^{v_{2i}\boldsymbol{\alpha}_2} - 1)g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_1} f(k_1; \mathbf{x}_i\boldsymbol{\beta}, \phi)} \right) + \right. \\ \left. I_{\{k_2\}}(y_i) \left(\frac{v_{1ij} e^{v_{1i}\boldsymbol{\alpha}_1} + (e^{v_{1i}\boldsymbol{\alpha}_1} e^{v_{2i}\boldsymbol{\alpha}_2})g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_2} f(k_2; \mathbf{x}_i\boldsymbol{\beta}, \phi) \cdot v_{1ij}}{1 + e^{v_{1i}\boldsymbol{\alpha}_1} + (e^{v_{1i}\boldsymbol{\alpha}_1} e^{v_{2i}\boldsymbol{\alpha}_2} - 1)g(\mathbf{x}_i\boldsymbol{\beta}, \phi)^{k_2} f(k_2; \mathbf{x}_i\boldsymbol{\beta}, \phi)} \right) + \right. \\ \left. (1 - I_{\{k_1\}}(y_i) - I_{\{k_2\}}(y_i)) \left(\frac{(e^{v_{1i}\boldsymbol{\alpha}_1} e^{v_{2i}\boldsymbol{\alpha}_2}) \cdot v_{1ij}}{e^{v_{1i}\boldsymbol{\alpha}_1} e^{v_{2i}\boldsymbol{\alpha}_2} - 1} - \frac{e^{v_{1i}\boldsymbol{\alpha}_1}}{1 + e^{v_{1i}\boldsymbol{\alpha}_1} \cdot v_{1ij}} \right) \right\}, \forall j = 0, 1, \dots, q_1.$$

Por fim, os elementos do vetor escore relacionados a α_2 , $U_{\alpha_{2j}} = \frac{\partial \ell(\boldsymbol{\beta}, \phi, \alpha_1, \alpha_2; \mathcal{D})}{\partial \alpha_{1j}}$, são dados por:

$$U_{\alpha_{2j}} = \sum_{i=1}^n \left\{ I_{\{k_1\}}(y_i) \left(\frac{v_{2ij} e^{v_{2i} \alpha_2} + (e^{v_{1i} \alpha_1} e^{v_{2i} \alpha_2}) g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^{k_1} f(k_1; \mathbf{x}_i \boldsymbol{\beta}, \phi) \cdot v_{2ij}}{1 + e^{v_{2i} \alpha_2} + (e^{v_{1i} \alpha_1} e^{v_{2i} \alpha_2} - 1) g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^{k_1} f(k_1; \mathbf{x}_i \boldsymbol{\beta}, \phi)} \right) + \right. \\ \left. I_{\{k_2\}}(y_i) \left(\frac{(e^{v_{1i} \alpha_1} e^{v_{2i} \alpha_2}) g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^{k_2} f(k_2; \mathbf{x}_i \boldsymbol{\beta}, \phi) \cdot v_{2ij}}{1 + e^{v_{1i} \alpha_1} + (e^{v_{1i} \alpha_1} e^{v_{2i} \alpha_2} - 1) g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^{k_2} f(k_2; \mathbf{x}_i \boldsymbol{\beta}, \phi)} \right) + \right. \\ \left. (1 - I_{\{k_1\}}(y_i) - I_{\{k_2\}}(y_i)) \left(\frac{(e^{v_{1i} \alpha_1} e^{v_{2i} \alpha_2}) \cdot v_{2ij}}{e^{v_{1i} \alpha_1} e^{v_{2i} \alpha_2} - 1} \right) - \frac{e^{v_{2i} \alpha_2}}{1 + e^{v_{2i} \alpha_2} \cdot v_{2ij}} \right\}, \forall j = 0, 1, \dots, q_2.$$

A Tabela 19 apresenta as derivadas das funções $g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^*$ e $f(*; \mathbf{x}_i \boldsymbol{\beta}, \phi)$ ($* = k_1, k_2, z_i$) do modelo k -IS com relação a β_j ($j = 0, 1, \dots, q_0$) e ϕ .

Tabela 19 – Derivadas das funções $g(\mathbf{x}_i \boldsymbol{\beta}, \phi)$ e $f(z_i; \mathbf{x}_i \boldsymbol{\beta}, \phi)$ do modelo k -IS.

$g(\mathbf{x}_i \boldsymbol{\beta}, \phi)^*$	$g_{\beta_j}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^*$	$g_{\phi}(\mathbf{x}_i \boldsymbol{\beta}, \phi)^*$
$\left(\frac{\phi + \mathbf{x}_i \boldsymbol{\beta}}{\phi - \mathbf{x}_i \boldsymbol{\beta}} \right)^{\frac{*}{2}}$	$\frac{* \phi x_{ij} \sqrt{(\phi + \mathbf{x}_i \boldsymbol{\beta})^{*-2}}}{\sqrt{(\phi - \mathbf{x}_i \boldsymbol{\beta})^{*+2}}}$	$-\frac{* \mathbf{x}_i \boldsymbol{\beta} \sqrt{(\phi + \mathbf{x}_i \boldsymbol{\beta})^{*-2}}}{\sqrt{(\phi - \mathbf{x}_i \boldsymbol{\beta})^{*+2}}}$
$f(*; \mathbf{x}_i \boldsymbol{\beta}, \phi)$	$f_{\beta_j}(*; \mathbf{x}_i \boldsymbol{\beta}, \phi)$	$f_{\phi}(*; \mathbf{x}_i \boldsymbol{\beta}, \phi)$
$e^{-\phi} \mathcal{J}_* \left(\sqrt{\phi^2 - (\mathbf{x}_i \boldsymbol{\beta})^2} \right)$	$\frac{-\mathbf{x}_i \boldsymbol{\beta} e^{-\phi} \mathcal{J}'_* \left(\sqrt{\phi^2 - (\mathbf{x}_i \boldsymbol{\beta})^2} \right)}{\sqrt{\phi^2 - (\mathbf{x}_i \boldsymbol{\beta})^2}} x_{ij}$	$e^{-\phi} \left(\frac{\phi \mathcal{J}'_* \left(\sqrt{\phi^2 - (\mathbf{x}_i \boldsymbol{\beta})^2} \right)}{\sqrt{\phi^2 - (\mathbf{x}_i \boldsymbol{\beta})^2}} - \mathcal{J}_* \sqrt{\phi^2 - (\mathbf{x}_i \boldsymbol{\beta})^2} \right)$

$$\text{em que } \mathcal{J}'_z = \frac{1}{2} (\mathcal{J}_{z-1}(\xi) + \mathcal{J}_{z+1}(\xi)).$$

5.2 Estimação dos Parâmetros: Abordagem Bayesiana

Devido ao bom desempenho obtido com o enfoque bayesiano via método SGHMC para o modelo k -MS, seguimos com a mesma abordagem para a estimação dos parâmetros do modelo Skellam k_1 e k_2 Inflacionado.

5.2.1 Distribuição Conjunta a Posteriori

Prosseguindo com a descrição do procedimento de estimação bayesiano, considerando a função função de log-verossimilhança apresentada em (5.4) e definindo o vetor de parâmetros $\boldsymbol{\theta}$, agora $\boldsymbol{\theta} = (\boldsymbol{\beta}, \alpha_1, \alpha_2, \phi)$, temos que a distribuição conjunta a posteriori é dada proporcionalmente por

$$\pi(\boldsymbol{\theta}; \mathcal{D}) \propto \exp \{ \ell(\boldsymbol{\beta}, \phi, \alpha_1, \alpha_2; \mathcal{D}) \} \pi(\boldsymbol{\beta}, \alpha_1, \alpha_2) \pi(\phi). \quad (5.5)$$

A distribuição a priori considerada para os parâmetros $\boldsymbol{\beta}$, α_1 e α_2 foi uma distribuição normal multivariada $N(\boldsymbol{\theta}_0, \sigma_0^2 \mathbf{I})$, onde $\boldsymbol{\theta}_0$ é um vetor de médias; $\sigma_0 > 0$ é uma constante; e \mathbf{I} é uma

matriz de identidade $(q_0 + q_1 + q_2 + 3)$ -dimensional. Para ϕ , como temos a condição de que $\phi \geq |\max\{|\mathbf{x}_i\boldsymbol{\beta}|, i = 1, \dots, n\}|$, consideramos uma distribuição priori normal truncada

$$\pi_0(\phi) = \frac{N(\phi_0, \sigma_1^2)}{1 - \Phi\left(\frac{\delta - \phi_0}{\sigma_1}\right)} = NT(\phi_0, \sigma_1^2, \delta),$$

tal que $\delta > \max\{|\mathbf{x}_i\boldsymbol{\beta}|, i = 1, \dots, n\}$ e $\Phi(\cdot)$ é a função de distribuição acumulada da normal padrão.

De maneira similar ao descrito na Subseção 4.3.2, o algoritmo SGHMC foi usado para gerar amostras das distribuições posteriores marginais, e assim realizar inferências sobre os parâmetros sobre os parâmetros do modelo.

5.3 Estudo de Simulação: Avaliação da Performance dos Estimadores Bayesiano

Neste estudo de simulação, $N = 100$ conjuntos de dados do modelo $\mathbf{k} - IS$ ($\mathbf{k} = (k_1, k_2)$) foram gerados, considerando tamanhos amostrais de $n = 100$ e 200 . Além disso, foram consideradas n observações da variável explicativa X , com $X \sim ZTN(0, 1)$. Para este estudo consideramos um único cenário, e os valores considerados de \mathbf{k} e de cada parâmetro de $\boldsymbol{\theta}$ para a geração do dados foram: $\mathbf{k} = (k_1; k_2) = (-4; 1)$ e $\boldsymbol{\theta} = (\beta_0; \beta_1; \alpha_{10}; \alpha_{11}; \alpha_{20}; \alpha_{21}; \phi) = (-7; 6; 3; 2; 4; -3; 7, 5)$

O modelo $\mathbf{k} - IS$ é ajustado a cada conjunto de dados gerados e a partir do algoritmo SGHMC obtém-se uma cadeia de tamanho 50.000 para cada parâmetro do modelo, com 50% dos elementos descartados como *burn-in*. O critério de Geweke foi empregado para avaliar a convergência das cadeias, e a media posteriori foi considerada como estimador bayesiano. Após obtidas as N estimativas dos parâmetros, calculamos o erro quadrático médio (EQM), a variância (Var), o viés absoluto médio (\mathcal{B}) e a probabilidade de cobertura dos intervalos de credibilidade bayesianos (PC).

A Tabela 20 apresenta os resultados das medidas obtidas para os estimadores dos parâmetros do modelo \mathbf{k} -IS. Analisando estes resultados, observamos bons desempenho dos estimadores, com valores de EQM , Var e \mathcal{B} bem próximos de zero. Corroborando a isso, temos também ainda probabilidades de cobertura com percentual de 100% (acima do nível nominal de 95%) para todos os parâmetros.

Tabela 20 – Avaliação dos estimadores dos parâmetros do modelo k -IS via procedimento bayesiano, considerando diferentes valores de n .

k		Medidas	Estimadores						n		
k_1	k_2		$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_{20}$	$\hat{\alpha}_{21}$	$\hat{\beta}_0$	$\hat{\beta}_1$		$\hat{\phi}$	
-4	1	EQM	0,0273	0,0050	0,0524	0,0354	0,0352	0,0255	0,0351	100	
			0,0279	0,0042	0,0299	0,0280	0,0307	0,0407	0,0380	200	
		Var	0,0276	0,0050	0,0453	0,0349	0,0355	0,0255	0,0250	100	
			0,0275	0,0041	0,0292	0,0260	0,0308	0,0402	0,0347	200	
		\mathcal{B}	0,0002	0,0000	0,0869	0,0294	0,0021	0,0144	0,1018	100	
			0,0260	0,0122	0,0313	0,0468	0,0134	0,0288	0,0599	200	
		PC (%)	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100
			100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	200

5.4 Aplicação: Conjuntos de Dados Artificiais

Aqui, um conjunto de dados de tamanho n considerado no estudo de simulação anterior é apresentado como ilustração, a fim de analisar principalmente as estimativas bayesianas pontuais e intervalar. Isto é: um conjunto de dado de tamanho n ($n = 100$ e 200) é gerado do modelo k -IS considerando a mesma variável explicativa X , o mesmo valor de k e dos parâmetros de θ do estudo de simulação.

Como curiosidade do comportamento de cada um desses conjuntos de dados, as Tabelas 21 e 22 apresentam as distribuições de frequências e algumas estatísticas descritivas dos conjuntos de dados artificiais gerado do modelo k -IS, com $k_1 = -4$, $k_2 = 1$, considerando $n = 100$ e $n = 200$ respectivamente.

Tabela 21 – Distribuição de frequências e estatísticas descritivas do conjunto de dados artificiais gerado do modelo k -IS, com $k_1 = -4$, $k_2 = 1$, considerando $n = 100$.

y_i	Conjunto de dados															\bar{y}	SD	
	-12	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3			5
f_i	1	4	1	2	6	8	12	17	6	11	7	5	14	4	1	1	-3,12	3,34

Tabela 22 – Distribuição de frequências e estatísticas descritivas do conjunto de dados artificiais gerado do modelo k -IS, com $k_1 = -4$, $k_2 = 1$, considerando $n = 200$.

y_i	Conjunto de dados																	\bar{y}	SD	
	-14	-13	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3			5
f_i	2	2	2	5	3	10	12	14	17	22	23	18	22	10	32	2	3	1	-3,25	3,54

Considerando a abordagem bayesiana via algoritmo SGHMC, vários modelos k -IS foram ajustados a cada conjunto de dados gerado, considerando diferentes valores para k . As estimativas bayesianas dos parâmetros (média a posteriori) e o intervalo HPD com 95% de credibilidade estão disponíveis na Tabela 23. Apesar do ajuste de diferentes modelos (consequência de considerar

diferentes \mathbf{k}), nota-se que em todos os casos considerados, as estimativas encontram-se próximas dos verdadeiros valores dos parâmetros, além destes estarem contidos nos respectivos intervalos de credibilidade.

Tabela 23 – Estimativas (e intervalos HPD com 95% de credibilidade) dos parâmetros dos modelos \mathbf{k} -IS ajustados considerando diferentes valores de \mathbf{k} , via procedimento bayesiano.

\mathbf{k}		Estimativas (HPD 95%)							n
k_1	k_2	$\hat{\alpha}_{10}$	$\hat{\alpha}_{11}$	$\hat{\alpha}_{20}$	$\hat{\alpha}_{21}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\phi}$	
-4	1	2,76	1,97	4,12	-2,99	-6,90	6,19	7,45	100
		(1,79;3,76)	(1,00;2,94)	(3,40;4,85)	(-3,82;-2,14)	(-7,52;-6,31)	(5,35;6,99)	(6,71;8,23)	
		3,11	2,08	3,99	-3,17	-6,99	6,17	7,57	200
		(2,21;4,00)	(1,11;3,01)	(3,36;4,61)	(-3,96;-2,42)	(-7,51;-6,48)	(5,40;6,93)	(6,88;8,31)	
0	1	1,86	1,85	4,13	-2,97	-6,77	6,34	7,62	100
		(1,10;2,66)	(0,93;2,77)	(3,40;4,87)	(-3,83;-2,14)	(-7,42;-6,11)	(5,50;7,15)	(6,82;8,48)	
		2,21	1,90	4,00	-3,14	-6,85	6,28	7,84	200
		(1,48;2,94)	(1,01;2,82)	(3,37;4,64)	(-3,93;-2,39)	(-7,38;-6,29)	(5,51;7,06)	(7,02;8,62)	
-4	-1	2,77	1,98	4,31	-2,89	-6,92	6,12	7,39	100
		(1,81;3,77)	(1,01;2,95)	(3,53;5,07)	(-3,77;-2,03)	(-7,52;-6,32)	(5,29;6,91)	(6,70;8,17)	
		3,13	2,08	4,26	-3,05	-6,97	6,08	7,42	200
		(2,23;4,01)	(1,12;3,01)	(3,60;4,97)	(-3,86;-2,22)	(-7,47;-6,45)	(5,31;6,81)	(6,77;8,12)	
-5	-2	2,78	1,95	4,33	-2,89	-7,06	6,07	7,43	100
		(1,80;3,76)	(0,98;2,93)	(3,56;5,10)	(-3,76;-2,02)	(-7,65;-6,45)	(5,25;6,87)	(6,79;8,16)	
		3,14	2,07	4,32	-3,058	-7,09	6,01	7,44	200
		(2,27;4,05)	(1,11;3,01)	(3,66;5,03)	(-3,87;-2,21)	(-7,62;-6,61)	(5,25;6,75)	(6,86;8,11)	
-5	-1	2,77	1,95	4,30	-2,90	-7,02	6,13	7,44	100
		(1,79;3,75)	(0,97;2,93)	(3,52;5,06)	(-3,77;-2,04)	(-7,64;-6,44)	(5,30;6,92)	(6,78;8,19)	
		3,13	2,06	4,25	-3,06	-7,07	6,10	7,48	200
		(2,26;4,04)	(1,12;3,02)	(3,59;4,95)	(-3,86;-2,23)	(-7,58;-6,57)	(5,33;6,83)	(6,85;8,16)	
-3	1	2,66	1,98	4,13	-2,99	-6,79	6,20	7,46	100
		(1,69;3,66)	(1,03;2,96)	(3,41;4,87)	(-3,86;-2,17)	(-7,41;-6,19)	(5,37;7,01)	(6,67;8,26)	
		3,03	2,08	3,99	-3,16	-6,90	6,17	7,57	200
		(2,14;3,94)	(1,12;3,00)	(3,38;4,64)	(-3,96;-2,41)	(-7,43;-6,39)	(5,42;6,96)	(6,85;8,34)	
-5	1	2,75	1,93	4,11	-3,01	-7,01	6,19	7,51	100
		(1,76;3,73)	(0,96;2,91)	(3,40;4,84)	(-3,82;-2,14)	(-7,65;-6,43)	(5,35;6,99)	(6,82;8,29)	
		3,11	2,05	3,98	-3,17	-7,09	6,18	7,62	200
		(2,23;4,01)	(1,12;3,01)	(3,36;4,60)	(-3,97;-2,43)	(-7,61;-6,59)	(5,42;6,96)	(6,96;8,35)	

Uma vez que vários modelos \mathbf{k} -IS foram ajustados por considerar diferentes valores para \mathbf{k} (apesar de se conhecer que os conjuntos de dados foram gerados considerando $k_1 = -4$ e $k_2 = 1$), obtivemos os valores dos critérios de seleção de modelo como uma forma de identificar o melhor modelo ajustado. Isso foi feito pensando no ponto de vista prático que, muitas vezes, os valores de k_1 e k_2 podem ser também desconhecidos e precisam ser determinados previamente para o ajuste do modelo. Idealmente esperamos que os critérios apontem para o verdadeiro modelo. A Tabela 24 apresenta os valores dos critérios de seleção de modelos WAIC, AIC e EBIC para todos os modelos ajustes, considerando $n = 100$ e 200 . Destaca-se que todos os critérios apontaram para os verdadeiro valores de k_1 e k_2 considerados para a geração dos dados, isto é, $k_1 = -4$ e $k_2 = 1$.

Tabela 24 – Valores dos critérios de seleção dos modelos \mathbf{k} -IS ajustados ao considerar diferentes valores de \mathbf{k} .

\mathbf{k}		Critérios de Seleção			n
k_1	k_2	WAIC	EAIC	EBIC	
-4	1	468,63	478,94	498,93	100
		957,07	966,29	991,95	200
0	1	488,10	490,69	511,68	100
		982,33	979,02	1005,53	200
-4	-1	504,64	501,21	521,59	100
		1028,12	1014,21	1040,21	200
-5	-2	503,19	500,41	520,77	100
		1073,69	1032,69	1059,05	200
-5	-1	511,08	504,34	524,95	100
		1036,58	1017,41	1043,49	200
-3	1	476,83	484,97	505,59	100
		958,53	967,17	992,86	200
-5	1	471,36	481,09	501,32	100
		959,70	967,88	993,62	200

A Figura 11 apresenta os dados gerados de Y versus X , juntamente com as médias verdadeiras $\mu_{\mathbf{k}-IS}$ e as médias ajustadas $\hat{\mu}_{\mathbf{k}-IS}$ obtidas com o modelo ajustado $\mathbf{k}-IS$, com $\mathbf{k} = (-4, 1)$, e tamanho de amostra $n = 200$. Observamos as proximidades entre as médias verdadeiras e ajustadas, indicando a boa qualidade do procedimento bayesiano via algoritmo SGHMC na estimação dos parâmetros.

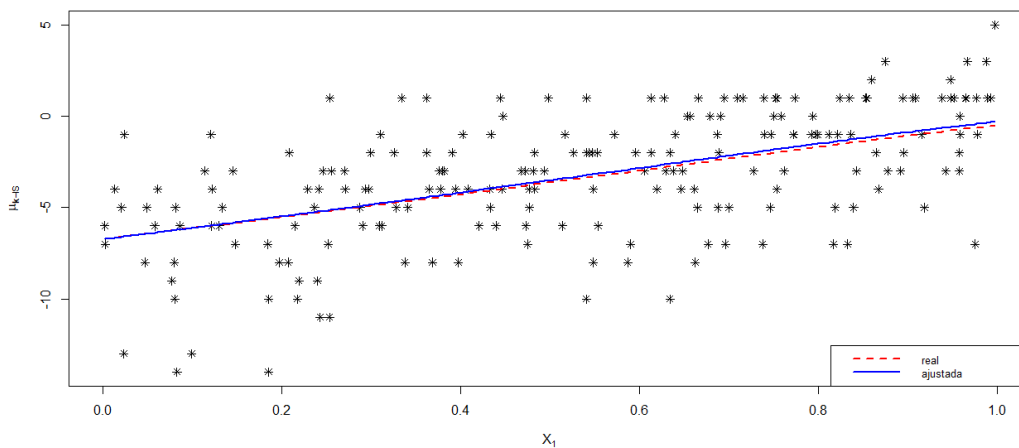


Figura 11 – Estimativas bayesianas de $\mu_{\mathbf{k}-IS}$ (juntamente com seus valores reais) obtidas a partir do modelo \mathbf{k} -IS ajustado, com $k_1 = -4$ e $k_2 = 1$, e tamanho amostral $n = 200$: \mathbf{y} versus a covariável X , juntamente com a estimativa de $\mu_{\mathbf{k}-IS}$.

5.5 Aplicação: Conjunto de Dados Reais

5.5.1 Dados do Índice da Bolsa de Valores de São Paulo

Iremos considerar o mesmo conjunto de dados do Ibovespa apresentado na Subseção 4.4.1 (ver Tabela 13 e Figura 6). O objetivo aqui é identificar possíveis pontos k_1 e k_2 que ocorrem inflação nos dados observados e assim obter o melhor modelo ajustado. Para isso, o ajuste do modelo k -IS aos dados foi feito considerando diferentes pares para k_1 e k_2 . Os valores de k foram definidos com base na distribuição de frequências dos *ticks* apresentada na Tabela 13 (e/ou Figura 6).

Os resultados dos modelos ajustados aos dados são apresentados na Tabela 25. Nesta Tabela encontra-se descrito os valores de k_1 e k_2 , bem como as estimativas e os intervalos HPD com 95% de credibilidade dos parâmetros de cada modelo ajustado.

Tabela 25 – Estimativas (e intervalo HPD com 95% de credibilidade) dos parâmetros dos modelos k -IS, considerando diferentes valores de k , ajustados aos dados do Ibovespa.

k		Estimativas (Intervalo HPD 95%)						
k_1	k_2	$\hat{\alpha}_{10}$	$\hat{\alpha}_{11}$	$\hat{\alpha}_{20}$	$\hat{\alpha}_{21}$	$\hat{\beta}_0$	$\hat{\beta}_1$	ϕ
-3	3	4,53 (3,65;5,51)	1,93 (0,01;3,87)	4,50 (3,53;5,60)	-5,13 (-9,09;-1,36)	0,27 (0,05;0,48)	-61,86 (-68,32;-55,32)	9,02 (8,24;9,79)
-2	-3	4,71 (3,82;5,68)	2,00 (0,03;3,90)	5,21 (3,72;7,03)	-5,35 (-11,13;-0,64)	0,25 (0,05;0,47)	-61,80 (-68,11;-55,14)	9,03 (8,29;9,83)
-2	-1	4,73 (3,81;5,71)	1,99 (0,10;3,89)	5,81 (4,08;7,72)	-4,91 (-10,77;-0,74)	0,24 (0,02;0,45)	-62,17 (-68,40;-56,10)	9,05 (8,32;9,86)
-2	0	4,71 (3,82;5,74)	2,01 (0,11;3,93)	4,42 (3,39;5,58)	-5,04 (-8,78;-1,42)	0,23 (0,02;0,45)	-61,95 (-68,18;-55,91)	9,14 (8,38;9,96)
-2	1	4,72 (3,80;5,69)	2,00 (0,00;3,90)	4,66 (3,25;6,40)	-4,68 (-10,38;-1,03)	0,22 (0,01;0,44)	-61,83 (-68,24;-55,38)	9,13 (8,35;9,95)
-2	2	2,72 (3,81;5,70)	1,99 (0,05;3,93)	5,33 (3,71;7,27)	-4,68 (-10,49;-0,98)	0,22 (0,03;0,42)	-61,74 (-67,97;-55,44)	9,07 (8,32;9,86)
-2	3	4,72 (3,84;5,71)	2,01 (0,12;3,99)	3,90 (3,12;4,78)	-4,38 (-8,30;-0,46)	0,17 (0,04;0,39)	-62,06 (-68,19;-55,66)	9,04 (8,28;9,81)
0	1	4,30 (3,39;5,27)	1,97 (0,10;4,00)	4,02 (3,11;5,04)	-4,62 (-8,58;-0,80)	0,20 (0,01;0,42)	-61,79 (-68,28;-55,81)	9,31 (8,50;10,16)
2	3	4,31 (3,45;5,30)	2,04 (0,10;3,98)	3,84 (3,08;4,68)	-4,30 (-8,15;-0,51)	0,14 (0,02;0,36)	-62,01 (-68,20;-55,89)	9,10 (8,36;9,92)

Para cada modelo ajustado, calculamos os critérios de seleção de modelos *WAIC*, *EAIC* e *EBIC*. A Tabela 26 apresenta os resultados destas medidas, de onde é possível notar a concordância unânime dos critérios, apontando o modelo k -IS com $k_1 = 2$ e $k_2 = 3$ como o melhor (menores valores para os três critérios). Temos assim, o conjunto de dados caracterizado como inflacionado de observações 2 e 3.

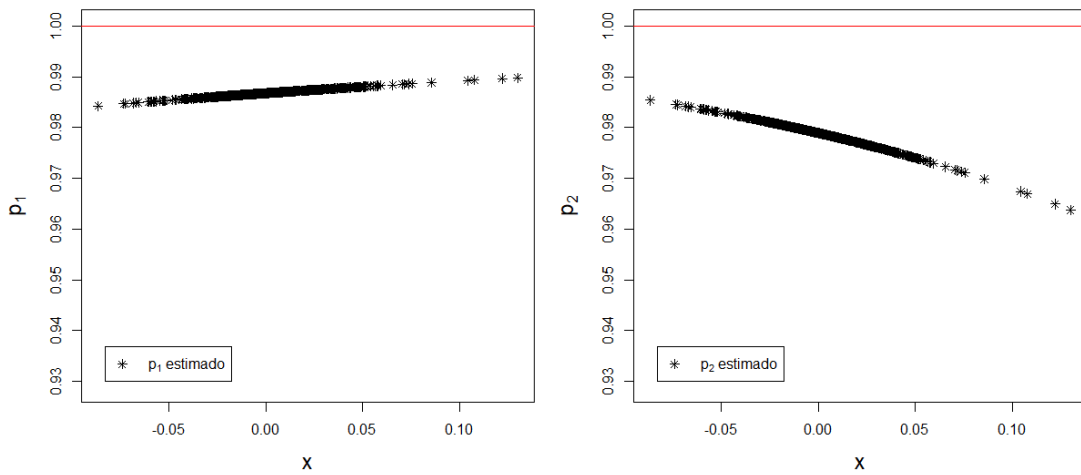
Algumas ilustrações gráficas dos resultados do ajuste são apresentadas na Figura 12. Os gráficos (A) e (B) desta Figura apresentam, respectivamente, os valores estimados para do parâmetro p_1 e do parâmetro p_2 . Já o gráfico (C) apresenta as observações y (*ticks*) em função da covariável X (log-retorno do câmbio do dólar), juntamente com a média μ_{k-IS} estimada. Podemos notar que a média estimada acompanha o comportamento dos dados.

Tabela 26 – Valores dos critérios de seleção dos modelos k -IS ajustados aos dados do Ibovespa, ao considerar diferentes valores de k .

k		Critérios de Seleção		
k_1	k_2	WAIC	EAIC	EBIC
-3	3	6207,33	6090,34	6130,48
-2	-3	6226,48	6094,04	6134,06
-2	-1	6217,81	6093,62	6133,68
-2	0	6218,78	6094,98	6135,13
-2	1	6220,66	6093,15	6133,17
-2	2	6218,56	6094,60	6134,63
-2	3	6220,99	6088,47	6128,55
0	1	6215,39	6093,14	6133,28
2	3	6207,96	6088,36	6128,51

(A)

(B)



(C)

$k_1 = 2$ e $k_2 = 3$

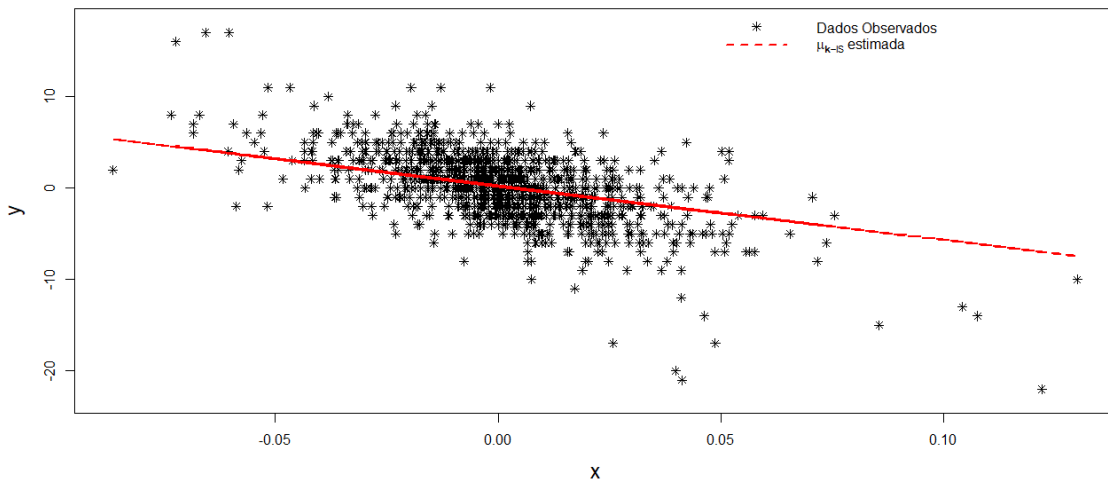


Figura 12 – Estimativas de p_1 , p_2 e de μ_{k-IS} obtidas a partir do modelo k -IS ajustado, com $k_1 = 2$ e $k_2 = 3$: (A) \hat{p}_1 versus a covariável X ; (B) \hat{p}_2 versus a covariável X ; (C) y versus a covariável X , juntamente com a estimativa de μ_{k-MS} .

Os valores de k_1 e k_2 apontados como inflacionados no conjunto de dados possuem frequências semelhantes. Durante o período analisado, houve uma variação de 2 *ticks* na pontuação do Ibovespa em 125 semanas, enquanto 3 *ticks* foram registrados em outras 125 semanas. Assim, nessas semanas as ações que compõem a carteira do Ibovespa apresentaram aumento de valor em comparação com a semana anterior, uma vez que as variações foram positivas.

COMENTÁRIOS FINAIS

Neste estudo, introduzimos generalizações da distribuição Skellam, que incluem um único ponto de k -modificação, abrangendo tanto inflação quanto deflação, e também considerando a inclusão de dois pontos de modificação, k_1 e k_2 , focando exclusivamente em casos de inflação. Para isso, foi necessário abordar algumas definições, características, propriedades e reparametrização da distribuição Skellam tradicional.

Para as distribuições propostas, a distribuição Skellam k -Modificada (k -MS) e a distribuição Skellam k_1 e k_2 Inflacionada (\mathbf{k} -IS), fornecemos uma análise abrangente de suas propriedades e particularidades. Extensões dessas distribuições foram feitas para o contexto de modelos de regressão. Apresentamos a formulação dos modelos k -MS e \mathbf{k} -IS, introduzindo variáveis explicativas para o parâmetro μ , comum nos dois modelos, e também para o(s) parâmetro(s) responsável(eis) pela modificação das probabilidades.

No procedimento de estimação e inferência dos parâmetros do modelo k -MS, adotamos uma abordagem clássica utilizando o método de máxima verossimilhança, além de uma abordagem bayesiana utilizando o método da classe MCMC, com o propósito de comparação. Os resultados obtidos através do procedimento bayesiano mostraram-se altamente promissores, impulsionados pelo uso eficaz do método do Gradiente Estocástico Hamiltoniano de Monte Carlo (SGHMC). Foi possível observar a adequação do ajuste do modelo k -MS a conjuntos de dados com diferentes pontos de modificação k , sem a necessidade de conhecimento prévio sobre o valor de ponto k e/ou mesmo sobre o tipo de modificação (inflação/deflação).

Aplicações do modelo k -MS foram apresentadas para dois problemas reais: O primeiro problema envolveu o conjunto de dados correspondente às observações da variação semanal da pontuação do Ibovespa. Isso se refere à diferença de preço, medida em *ticks* (centavos), do dia atual em relação ao dia anterior, durante o período entre janeiro de 2000 e dezembro de 2022; Já o segundo problema envolveu dois conjuntos de dados correspondentes às diferenças entre partidas ganhas e partidas perdidas pelos times durante a temporada regular de 2022-2023 da

National Basketball Association, em cada conferência (Leste e Oeste).

Na aplicação que utiliza os dados da variação semanal da pontuação do Ibovespa, onde o preço do dólar (log-retorno) é empregado como covariável, o modelo k -MS com $k = -2$ foi identificado como o mais bem ajustado segundo os critérios de seleção. Os valores estimados para o parâmetro p , caracterizou o conjunto de dados como inflacionado de observações -2 (*ticks*). Já para a aplicação dos dados da temporada regular da NBA de 2022-2023, os critérios de seleção evidenciaram que o modelo k -MS com $k = -12$ foi identificado como o mais bem ajustado para a conferência Leste, enquanto para a conferência Oeste, o valor indicado foi $k = -38$.

Mediante a constatação do bom desempenho do procedimento bayesiano na estimação e inferência dos parâmetros do modelo k -MS, optamos por seguir essa abordagem na estimação dos parâmetros do modelo Skellam k_1 e k_2 Inflacionado. Os resultados das estimativas pontuais e dos intervalos de credibilidade obtidos no estudo de simulação também foram altamente satisfatórios. A aplicação, realizada para os dados do Ibovespa, teve o indicativo de inflação nos pontos 2 e 3, que representam variação positivas na pontuação do índice.

Com base nos resultados obtidos, tanto os modelos k -MS quanto os k -IS demonstraram ser boas alternativas para explicar o comportamento de dados com valores inteiros: o modelo k -MS por abranger tanto inflação quanto deflação de qualquer observação k e possibilitar a identificação do tipo de modificação por meio da estimativa do parâmetro p ; o modelo k -IS por permitir a análise de conjuntos de dados com a presença de dois pontos de inflação. Em resumo, são potenciais ferramentas no âmbito de distribuição e de modelagem para conjuntos de dados inteiros.

Por fim, há várias pesquisas que podem dar continuidade ao trabalho desenvolvido nesta tese. Entre elas, propomos os seguintes tópicos:

- Considerar uma abordagem bayesiana para identificar o ponto de k -Modificação, introduzindo uma priori para k ;
- Considerar o modelo k -MS no contexto multivariado;
- Estender a ideia da inflação na probabilidade de ocorrência de k_1 e de k_2 feita no modelo k -IS, para o caso k_1 e k_2 Modificado. Assim, teríamos o Modelo Skellam k_1 e k_2 Modificado, que acomodaria também a deflação.

REFERÊNCIAS

AKAIKE, H. A new look at the statistical model identification. **IEEE transactions on automatic control**, Ieee, v. 19, n. 6, p. 716–723, 1974. Citado na página 52.

ALZAID, A. A.; OMAIR, M. A. On the poisson difference distribution inference and applications. **Bulletin of the Malaysian Mathematical Sciences Society. Second Series**, Malaysian Mathematical Sciences Society, Bangi, v. 33, n. 1, p. 17–45, 2010. Citado nas páginas 23, 25, 27, 28 e 29.

_____. Poisson difference integer valued autoregressive model of order one. **Bulletin of the Malaysian Mathematical Sciences Society**, Springer Science & Business Media, v. 37, n. 2, p. 465–485, 2014. Citado na página 23.

ARFKEN, G. B.; WEBER, H. J.; HARRIS, F. E. Chapter 14 - Bessel functions. In: ARFKEN, G. B.; WEBER, H. J.; HARRIS, F. E. (Ed.). **Mathematical Methods for Physicists**. 7^a. ed. Boston: Academic Press, 2013. p. 643–713. ISBN 978-0-12-384654-9. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780123846549000141>>. Citado na página 25.

Basketball Reference. **Basketball Reference**. 2024. Disponível em: <<https://www.basketball-reference.com/>>. Citado na página 81.

BROOKS, S. Discussion of the paper Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society**, v. 64, p. 616–618, 2002. Citado na página 69.

CARLIN, B. P.; LOUIS, T. A. **Bayesian methods for data analysis**. [S.l.]: CRC Press, 2008. Citado nas páginas 64 e 69.

CARVALHO, S. O. **Distribuições k -Modificadas da Família Série de Potência Uniparamétrica**. Tese (Dissertação de Mestrado) — PIGGES – Programa Interinstitucional de Pós Graduação em Estatística, Universidade de São Paulo e Universidade Federal de São Carlos, São Carlos – SP, 2017. Citado na página 22.

CHEN, T.; FOX, E.; GUESTRIN, C. Stochastic gradient hamiltonian monte carlo. In: XING, E. P.; JEBARA, T. (Ed.). **Proceedings of the 31st International Conference on Machine Learning**. Beijing, China: [s.n.], 2014. p. 1683–1691. Citado nas páginas 64 e 67.

CONCEIÇÃO, K. S. **Modelos série de potência zero-modificados**. Tese (Doutorado) — Universidade Federal de São Carlos, 2013. Citado na página 21.

CONCEIÇÃO, K. S.; SUZUKI, A. K.; ANDRADE, M. G. A bayesian approach for zero-modified skellam model with hamiltonian mcmc. **Statistical Methods & Applications**, Springer, p. 1–19, 2020. Citado nas páginas 23 e 32.

DIETZ, E.; BÖHNING, D. On Estimation of the Poisson Parameter in Zero-Modified Poisson Models. **Computational Statistics & Data Analysis**, v. 34, p. 441–459, 2000. Citado na página 21.

DUANE, S.; KENNEDY, A. D.; PENDLETON, B. J.; ROWETH, D. Hybrid monte carlo. **Physics letters B**, Elsevier, v. 195, n. 2, p. 216–222, 1987. Citado na página 64.

GERSTENKORN, T. Moment recurrence relations for the inflated negative binomial, poisson and geometric distributions. **Demonstratio Mathematica**, De Gruyter Open, v. 12, n. 2, p. 389–410, 1979. Citado na página 21.

GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments. **Bayesian statistics**, Clarendon Press, v. 4, p. 641–649, 1992. Citado na página 70.

HIRAKAWA, K.; BAQAI, F.; WOLFE, P. J. Wavelet-based poisson rate estimation using the skellam distribution. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Computational Imaging VII**. [S.l.], 2009. v. 7246, p. 72460R. Citado na página 23.

HUAYANAY, A.; N, J. B.; CANCHO, V.; DEY, D. Performance of asymmetric links and correction methods for imbalanced data in binary regression. **Journal of Statistical Computation and Simulation**, v. 89, n. 9, p. 1694–1714, 2019. Citado na página 47.

IRWIN, J. O. The frequency distribution of the difference between two independent variates following the same poisson distribution. **Journal of the Royal Statistical Society**, Wiley Online Library, v. 100, n. 3, p. 415–416, 1937. Citado na página 22.

KARLIS, D.; NTZOUFRAS, I. Bayesian analysis of the differences of count data. **Statistics in medicine**, Wiley Online Library, v. 25, n. 11, p. 1885–1905, 2006. Citado nas páginas 23 e 28.

_____. Bayesian modelling of football outcomes: using the skellam's distribution for the goal difference. **IMA Journal of Management Mathematics**, Oxford University Press, v. 20, n. 2, p. 133–145, 2008. Citado nas páginas 23 e 27.

KING, G. Event Count Models for International Relations: Generalizations and Applications. **International Studies Quarterly**, v. 33, p. 123–147, 1989. Citado na página 37.

KOOPMAN, S. J.; LIT, R.; LUCAS, A. Intraday stochastic volatility in discrete price changes: the dynamic skellam model. **Journal of the American Statistical Association**, Taylor & Francis, v. 112, n. 520, p. 1490–1503, 2017. Citado na página 23.

MULLAHY, J. Specification and Testing of Some Modified Count Data Models. **Journal of Econometrics**, v. 33, p. 341–365, 1986. Citado na página 37.

MURAT, M.; SZYNAL, D. Non-zero inflated modified power series distributions. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 27, n. 12, p. 3047–3064, 1998. Citado na página 22.

NEAL, R. M. **Bayesian learning for neural networks Ph. D.** Tese (Doutorado) — thesis. Dept. of Computer Science, University of Toronto, Toronto, 1994. Citado na página 64.

NEAL, R. M. **Handbook of Markov Chain Monte**. [S.l.]: Chapman & Hall, New York, 2011. 113–162 p. Citado nas páginas 64 e 67.

NTZOUFRAS, I.; PALASKAS, V.; DRIKOS, S. Bayesian models for prediction of volleyball set differences. **arXiv preprint arXiv:1911.04541**, 2019. Citado na página 22.

PANDEY, K. On generalized inflated poisson distribution. **J. Sci. Res. Banaras Hindu Univ**, v. 15, n. 2, p. 157–162, 1964. Citado nas páginas 21 e 22.

POWELL, M. J. *et al.* The bobyqa algorithm for bound constrained optimization without derivatives. **Cambridge NA Report NA2009/06, University of Cambridge, Cambridge**, v. 26, p. 26–46, 2009. Citado na página 51.

SCHWARZ, G. Estimating the dimension of a model. **The annals of statistics**, JSTOR, p. 461–464, 1978. Citado na página 52.

SINGH, M. Inflated binomial distribution. **Banaras Hindu Univ. J. Sci. Res**, v. 16, n. 1, p. 87–90, 1965. Citado na página 21.

SINGH, S. Note on inflated poisson-distribution. In: INST MATHEMATICAL STATISTICS IMS BUSINESS OFFICE-SUITE 7, 3401 INVESTMENT **ANNALS OF MATHEMATICAL STATISTICS**. [S.l.], 1963. v. 33, n. 3, p. 1210. Citado na página 21.

SKELLAM, J. G. The frequency distribution of the difference between two poisson variates belonging to different populations. **Journal of the Royal Statistical Society: Series A**, v. 109, p. 296, 1946. Citado nas páginas 22 e 23.

UMBACH, D. On inference for a mixture of a poisson and a degenerate distribution. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 10, n. 3, p. 299–306, 1981. Citado na página 21.

WATANABE, S.; OPPER, M. Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. **Journal of machine learning research**, v. 11, n. 12, 2010. Citado na página 69.

YPMA, J.; BORCHERS, H. W.; EDELBUETTEL, D.; YPMA, M. J. Package ‘nloptr’. **R package version**, Citeseer, v. 1, n. 1, 2018. Citado na página 51.

