



UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO DE CIÊNCIAS FÍSICAS E MATEMÁTICAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

Igor Dornelles Schoeller Siciliani

**Aplicações de Mecânica Estatística em Econofísica e  
Mineração de Texto**

Florianópolis

2024

Igor Dornelles Schoeller Siciliani

**Aplicações de Mecânica Estatística em Econofísica e  
Mineração de Texto**

Tese submetida ao Programa de Pós-Graduação em Física da Universidade Federal de Santa Catarina, como requisito parcial à obtenção do título de Doutor em Física.

Orientador: Prof. Dr. Marcelo H.R. Tragtenberg

Florianópolis, 15 de março de 2024.

Ficha catalográfica gerada por meio de sistema automatizado gerenciado pela BU/UFSC.  
Dados inseridos pelo próprio autor.

Siciliani, Igor Dornelles Schoeller  
Aplicações de Mecânica Estatística em Econofísica e  
Mineração de Texto / Igor Dornelles Schoeller Siciliani ;  
orientador, Marcelo Henrique Romano Tragtenberg, 2024.  
107 p.

Tese (doutorado) - Universidade Federal de Santa  
Catarina, Centro de Ciências Físicas e Matemáticas,  
Programa de Pós-Graduação em Física, Florianópolis, 2024.

Inclui referências.

1. Física. 2. econofísica. 3. distribuição de riqueza. 4.  
classificação de sentimentos. 5. modelo estocástico. I.  
Tragtenberg, Marcelo Henrique Romano. II. Universidade  
Federal de Santa Catarina. Programa de Pós-Graduação em  
Física. III. Título.

**Igor Dornelles Schoeller Siciliani**

**Aplicações de Mecânica Estatística em Econofísica e Mineração de Texto**

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

**Prof. Dr. Eraldo Sérgio Barbosa da Silva**

Universidade Federal de Santa Catarina - UFSC

**Prof. Dr. Lucas Nicolao**

Universidade Federal de Santa Catarina - UFSC

**Prof. Dr. Osame Kinouchi Filho**

Universidade de São Paulo - USP

**Prof. Dr. Evaldo Mendonça Fleury Curado**

Centro Brasileiro de Pesquisas Físicas - CBPF

**Prof. Dr. Maurício Girardi Schappo**

Universidade Federal de Santa Catarina - UFSC

Certificamos que esta é a versão original e final do trabalho de conclusão que foi julgado adequado para obtenção do título de doutor em física.

---

Coordenação do Programa de Pós-Graduação

---

Prof. Dr. Marcelo H.R. Tragtenberg

Florianópolis, 15 de março de 2024.

# Resumo

Esse trabalho tem como característica suas duas áreas de atuação distintas: Econofísica e Mineração de Textos. Na primeira, uma investigação sobre a distribuição da renda pessoal brasileira, utilizando dados da Pesquisa Nacional por Amostra de Domicílios (PNAD), uma pesquisa anual disponibilizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE), é cuidadosamente articulada. Ainda, a partir da base de dados da PNAD, também, confirmou-se a eficácia de um modelo semi empírico que concilia a lei de Pareto para a distribuição de renda das pessoas mais ricas e a distribuição de Boltzmann-Gibbs para o resto da população. Do mesmo modo, foram realizadas quatro medidas de desigualdade de renda: índice de Pareto, renda média e renda de corte (estas advindas do modelo) e o coeficiente de Gini. Essas são calculadas para a população geral, bem como para dois tipos de dicotomias populacionais: pretos/pardos/indígenas versus brancos/amarelos e, homens versus mulheres. Também é realizada uma análise das séries temporais desses índices para o período 2001-2014. Os resultados sugerem uma diminuição da desigualdade de renda brasileira no período selecionado. Outro resultado importante é que os subgrupos historicamente desfavorecidos (Mulheres e pretos / indígenas / pardos), a maioria da população, têm uma distribuição de renda mais igualitária. No entanto, esses grupos têm uma renda mensal menor que os demais e, essa estrutura social permaneceu praticamente inalterada no período.

Em continuidade a este campo de estudo, desenvolveu-se e validou-se um novo modelo estocástico não-linear para analisar a variável 'riqueza'. Inspirado no mecanismo de integração e disparo, este modelo propõe a divisão dos agentes em duas classes distintas na rede. O modelo é fundamentado em princípios básicos, como a concentração de riqueza, a capacidade de gerar riqueza em um sistema não termodinâmico, interações mutuamente vantajosas e diferentes métodos de acumulação de riqueza. Tanto teoricamente quanto através de simulações, demonstrou-se que a riqueza relativa dos agentes alcança um estado de equilíbrio. Neste ponto de equilíbrio, é possível determinar as distribuições de probabilidade da riqueza. Especificamente, a distribuição de equilíbrio identificada para este modelo é uma combinação das distribuições Laplace e Pareto-1. Nessa mesma linha, utilizando este modelo teórico, foi possível calcular indicadores econômicos, como o índice de Pareto e a curva de Lorenz, para essa nova distribuição de riqueza. Por último, o modelo foi comparado com os dados do DINA (Distribuição Nacional de Renda) dos anos de 2009 e 2019, revelando uma boa concordância entre os resultados teóricos e os dados reais.

Na segunda parte é apresentado a formulação matemática e resultados numéricos de um novo modelo de classificação de sentimentos em textos a partir da

abordagem léxica. Além disso, através da análise de *twitters* são confrontados os resultados obtidos pelo modelo proposto, pelo método de aprendizagem supervisionada de Naive-Bayes e a própria classificação do ser humano. Os resultados desse primeiro tópico sugerem que o novo modelo desenvolvido possui precisão superior ao método de Naive-Bayes, quando esses dois métodos são confrontados com a classificação realizada por humanos.

**Palavras-chave:** 1. classificação de sentimentos 2. econofísica 3. distribuição de riqueza. 4. distribuição de renda. 5. modelo estocástico.

# Abstract

This work is characterized by its two distinct areas of focus: Econophysics and Text Mining. In the first area, an investigation into the distribution of personal income in Brazil, using data from the National Household Sample Survey (PNAD), an annual survey provided by the Brazilian Institute of Geography and Statistics (IBGE), is carefully articulated. Additionally, using the PNAD database, the effectiveness of a semi-empirical model that reconciles Pareto's law for the income distribution of the richest individuals with the Boltzmann-Gibbs distribution for the rest of the population was confirmed. Similarly, four measures of income inequality were carried out: Pareto index, average income, and cutoff income (these derived from the model), and the Gini coefficient. These are calculated for the general population, as well as for two types of population dichotomies: black/brown/indigenous versus white/yellow, and men versus women. An analysis of the time series of these indices for the period 2001-2014 was also conducted. The results suggest a decrease in Brazilian income inequality during the selected period. Another important result is that historically disadvantaged subgroups (women and black/indigenous/brown individuals), the majority of the population, have a more egalitarian income distribution. However, these groups have a lower monthly income than others, and this social structure remained practically unchanged during the period.

Continuing in this field of study, a new non-linear stochastic model was developed and validated to analyze the 'wealth' variable. Inspired by the integrate-and-fire mechanism, this model proposes the division of agents into two distinct classes in the network. The model is based on basic principles such as wealth concentration, the ability to generate wealth in a non-thermodynamic system, mutually beneficial interactions, and different methods of wealth accumulation. Both theoretically and through simulations, it was shown that the relative wealth of agents reaches an equilibrium state. At this equilibrium point, it is possible to determine the probability distributions of wealth. Specifically, the equilibrium distribution identified for this model is a combination of Laplace and Pareto-1 distributions. Along the same lines, using this theoretical model, it was possible to calculate economic indicators, such as the Pareto index and the Lorenz curve, for this new wealth distribution. Finally, the model was compared with the National Income Distribution (DINA) data for the years 2009 and 2019, revealing a good agreement between the theoretical results and the actual data.

In the second part, the mathematical formulation and numerical results of a new sentiment classification model in texts using the lexical approach are presented. Furthermore, through the analysis of tweets, the results obtained by the proposed

model, the Naive Bayes supervised learning method, and human classification were compared. The results of this first topic suggest that the new model developed has superior accuracy to the Naive Bayes method when these two methods are compared with human classification.

**Keywords:**1. sentiment classification 2. econophysics 3. wealth distribution 4. income distribution 5. stochastic model.



# Lista de ilustrações

2.2.1	Exemplo de distribuição de renda . . . . .	25
2.2.2	Exemplo de distribuição de dinheiro: Modelo de Duas Classes . . . . .	27
2.2.3	Exemplo de distribuição de dinheiro: Modelo de Duas Classes com dinheiro negativo . . . . .	28
2.2.4	Modelo LVG . . . . .	30
3.2.1	Curva de Lorenz . . . . .	34
3.3.1	PNAD 2014 . . . . .	36
3.3.2	Curva de Lorenz e Coeficiente de Gini: PNAD 2014 . . . . .	37
3.4.1	Série temporal da renda média e Valor de Corte . . . . .	40
3.4.2	Série temporal Coeficiente de Gini . . . . .	41
3.4.3	Diferença Percentual Relativa a População Total . . . . .	42
3.4.4	Série Temporal Índice de Pareto . . . . .	44
4.1.1	MECR densidade de probabilidade . . . . .	57
4.1.2	MECR densidade de probabilidade acumulada . . . . .	58
4.2.1	Diagrama de um passo de simulação . . . . .	59
4.2.2	Funções de disparo . . . . .	60
4.2.3	Densidade dos ricos e tempo de decorrelação . . . . .	61
4.2.4	Evolução temporal da riqueza relativa média . . . . .	62
4.2.5	Densidade de probabilidade MECR+sigmóide vs. simulação . . . . .	63
4.2.6	Densidade de probabilidade MECR+degrau vs. simulação . . . . .	64
4.2.7	Variação da proporção de ricos e Renda de Corte $V_C$ . . . . .	65
4.2.8	Método para determinação de $V_C$ . . . . .	66
4.2.9	Curva de Lorenz . . . . .	67
4.2.10	DINA 2009 . . . . .	69
4.2.11	DINA 2019 . . . . .	70
5.1.1	Processo de Análise de Sentimento . . . . .	76
6.1.1	Diagrama MONEMC . . . . .	82
6.1.2	Exemplo de classificação . . . . .	83

# Lista de tabelas

3.3.1	PNAD de 2014 . . . . .	38
6.1.1	Exemplo de dicionário base $D(T)$ . . . . .	80
6.2.1	Twitters e classificação . . . . .	84
6.2.3	Comparação NB vs. MONEMC . . . . .	88

# Lista de abreviaturas e siglas

AS	Análise de Sentimento
BA	Branco-Amarelo
FDA	Função Distribuição Acumulada
FDAC	Função Distribuição Acumulada Complementar
DBG	Distribuição de Boltzmann-Gibbs
FD	Função de Disparo
GGL	Gerstner-Galves-Löcherbach
IA	Inteligência Artificial
KS	Kolmogorov-Smirnov
LVG	Lotka-Volterra Generalizado
ME	Máxima Entropia
MVS	Máquina de Vetores de Suporte
MO	Mineração de Opinião
MECR	Modelo Econômico com Concentração de Riqueza
MONEMC	Modelo Neural de Múltiplos Canais
NB	Naive Bayes
PNAD	Pesquisa Nacional por Amostra de Domicílios
PPI	Preto-Pardo-Indígena

# Sumário

	<b>Resumo</b> . . . . .	<b>5</b>
	<b>Abstract</b> . . . . .	<b>7</b>
<b>1</b>	<b>INTRODUÇÃO</b> . . . . .	<b>15</b>
<b>I</b>	<b>ECONOFÍSICA</b>	<b>19</b>
<b>2</b>	<b>INTRODUÇÃO A ECONOFÍSICA E AOS MODELOS DE AGENTES PARA DISTRIBUIÇÃO DE RIQUEZA</b> . . . . .	<b>23</b>
2.1	Econofísica e a distribuição de renda e riqueza . . . . .	23
2.2	Modelos de distribuição de renda e riqueza . . . . .	24
2.2.1	<b>Lei de Pareto</b> . . . . .	<b>24</b>
2.2.2	<b>Modelo de Duas Classes</b> . . . . .	<b>25</b>
2.2.3	<b>O Modelo Lotka-Volterra Generalizado</b> . . . . .	<b>28</b>
2.2.4	<b>O Modelo Duplo Pareto</b> . . . . .	<b>30</b>
<b>3</b>	<b>TEORIA CINÉTICA E DISTRIBUIÇÃO DE RENDA NO BRASIL</b> .	<b>33</b>
3.1	Modificações no Modelo de Duas Classes . . . . .	33
3.2	Indicadores de desigualdade social . . . . .	34
3.3	Distribuição de renda no Brasil . . . . .	35
3.4	Séries temporais dos coeficientes de desigualdade . . . . .	39
<b>4</b>	<b>MODELO ECONÔMICO COM CONCENTRAÇÃO DE RIQUEZA (MECR)</b> . . . . .	<b>47</b>
4.1	Fundamentação do modelo . . . . .	47
4.1.1	<b>Classe baixa</b> . . . . .	<b>54</b>
4.1.2	<b>Classe alta</b> . . . . .	<b>54</b>
4.1.3	<b>Condições de contorno do MECR</b> . . . . .	<b>56</b>
4.2	MECR simulações e resultados . . . . .	59
<b>II</b>	<b>MINERAÇÃO DE TEXTOS</b>	<b>71</b>
<b>5</b>	<b>ANÁLISE DE SENTIMENTO</b> . . . . .	<b>75</b>
5.1	Análise de sentimento: uma visão geral . . . . .	75
5.1.1	<b>Métodos estatísticos de classificação</b> . . . . .	<b>77</b>

---

5.1.1.1	Naive Bayes . . . . .	77
5.1.1.2	Máxima Entropia . . . . .	78
6	<b>MODELO NEURAL DE MÚLTIPLOS CANAIS PARA ANÁLISE DE SENTIMENTO . . . . .</b>	<b>79</b>
6.1	Fundamentos teóricos do MONEMC . . . . .	79
6.2	Resultados e trabalhos futuros . . . . .	84
7	<b>CONCLUSÃO . . . . .</b>	<b>89</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>91</b>
	<b>APÊNDICES . . . . .</b>	<b>97</b>
	<b>APPENDIX A – O MODELO DE GERSTNER-GALVES-LÖCHERBACH</b>	<b>99</b>
	<b>APPENDIX B – TESTE DE KOLMOGOROV-SMIRNOV PARA DUAS AMOSTRAS . . . . .</b>	<b>101</b>
	<b>APPENDIX C – ALGORITMO PARA DETERMINAÇÃO DE <math>V_C</math> E <math>\omega</math> . . . . .</b>	<b>103</b>
	<b>APPENDIX D – L-BFGS-B . . . . .</b>	<b>105</b>



# 1 Introdução

A Mecânica Estatística é uma das áreas fundamentais da Física, sendo utilizada em quase todos os seus campos, como a Astrofísica e a Mecânica Quântica. Do mesmo modo, devido ao seu grande êxito no tratamento de problemas de extrema complexidade envolvendo teorias probabilísticas e dinâmicas de muitos corpos, sua aplicação através de novos métodos computacionais tem estendido sua gama de atuação para outras áreas do conhecimento, das quais podemos citar: a Química, a Biologia, a Economia e, mais recentemente, a Aprendizagem de Máquina (*Machine Learning*).

Nesse sentido, este trabalho reflete esse intenso e produtivo movimento interdisciplinar que a Mecânica Estatística, inerentemente, é capaz de prover aos pesquisadores dessa área. Assim, esse texto será dividido em duas grandes partes: I Econofísica e II Mineração de Textos.

O termo “Econofísica” é oriundo dos anos 90 e foi difundido por um livro de 1999 intitulado “*Introduction to Econophysics*”[1]. Esse sugere que uma abordagem alternativa para a Economia poderia ser fornecida pela Física ou até mesmo que a Economia e seus princípios poderiam ser baseados e explicados por princípios físicos, sobretudo, da Mecânica Estatística. A “invasão” dos físicos na economia e finanças na virada do milênio é um fascinante fenômeno. Numa definição mais formal, poderíamos dizer que a Econofísica procura utilizar métodos originalmente desenvolvidos para aplicações em problemas físicos, na descrição, modelagem e solução de questões econômicas, envolvendo, em grande parte das vezes, dinâmica não linear, processos estocásticos e incertezas.

Na economia moderna, desproporções na distribuição de renda e riqueza ao longo das sociedades representam fontes de instabilidade social. Desse fato, pode-se entender o motivo do tópico em questão receber tamanho destaque em se tratando de pesquisas relativas à Economia, à Sociologia e à Política, em especial, a partir do século XIX com as ideias de Vilfredo Pareto[2]. Desse ponto, apresenta-se nesse estudo uma investigação sobre a distribuição de renda da população brasileira através dos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) ao longo do período de 2001 a 2014. A abordagem em questão, chamada de Modelo de Duas Classes, propõe uma distribuição de probabilidade dupla para a renda pessoal dos agentes econômicos. Ou seja, os mais ricos seguiriam uma distribuição de Pareto e, o resto da população uma distribuição de Boltzmann-Gibbs. Como consequência, algumas medidas que refletem a desigualdade social de uma população podem ser calculadas, como exemplo, temos o índice de Pareto e a a renda média. Ainda, para lidar com as muitas dimensões da desigualdade de renda, alguns desses indicadores são calculados para a população

brasileira em geral, bem como para dois tipos de dicotomias populacionais: pretos / indígenas / pardos versus brancos / amarelos; e homens versus mulheres ao longo do período proposto. Assim, o capítulo 2 intitulado Teoria Cinética e Distribuição de Renda no Brasil parte dessa dissertação refere-se ao artigo "*Kinetic Theory and Brazilian Income Distribution, 2019*", estudo realizado pelo mesmo autor dessa tese e seu orientador e publicado na revista "*Physica A: Statistical Mechanics and its Applications - Journal - Elsevier*" (<https://doi.org/10.1016/j.physa.2018.08.164>). Ainda, nessa mesma linha, uma extensão ao estudo em questão, que não consta no escopo dessa tese, foi desenvolvida e publicada na revista "*Physical Review E: covering statistical, nonlinear, biological, and soft matter physics -APS*" com o título "*Optimal income crossover for a two-class model using particle swarm optimization, 2022*" (<https://doi.org/10.1103/PhysRevE.106.034313>).

Para além do Modelo de Duas Classes, a literatura sugere que a distribuição da riqueza pode ser adequadamente descrita pela combinação de várias distribuições de probabilidade, como a distribuição de Pareto para as faixas de alta renda e outras distribuições (exponencial, Gompertz, Normal, ou Galton) para o restante da população. Essa abordagem tem mostrado grande sucesso ao ser aplicada em diferentes conjuntos de dados, incluindo pesquisas domiciliares e informações fiscais, indicando a existência de uma distribuição de equilíbrio segmentada de acordo com as classes econômicas. A modelagem desse fenômeno por sua vez, tem se focado em dinâmicas estocásticas onde o comportamento de cauda pesada está associado a região assintótica da distribuição de equilíbrio do sistema, refletindo a concentração de riqueza nas mãos de uma pequena parcela da população. Nesse contexto, também, apresentaremos no corrente trabalho um novo modelo estocástico não-linear que visa aprofundar a compreensão da distribuição de riqueza. Este modelo, chamado pelo autor de MECR, propõe que a população seja dividida em classes alta e baixa definidas a partir de uma função de disparo e, por consequência, com processos diferenciados para a geração de riqueza. O modelo MECR enfatiza a importância de considerar a diversidade de processos de geração de riqueza e as interações econômicas entre os indivíduos. Ele aborda a concentração de riqueza e o impacto das redes de interação econômica, sugerindo que as variações na riqueza média da população são fortemente influenciadas pela classe mais rica (Efeito de Transbordamento). Ainda, uma das premissas fundamentais do proposto modelo é a acumulação de riqueza por uma parcela pequena da população. A partir disso, é desenvolvido um construto teórico que culmina em um estado de equilíbrio com uma distribuição da riqueza relativa do tipo: Laplace-Pareto-1. Desse fato, a possibilidade de se estudar riquezas negativas também é alcançada. O "efeito de riqueza negativa", nessa situação, significa que alguns dos agentes mais pobres gastam mais do que recebem, ou ainda, que possuem uma dívida que excede seu total de ativos. Por fim, os resultados teóricos serão confrontados com simulações computacionais para análises e validações.



---

Na parte II, um assunto bem diferente do anterior será abordado: o processamento de textos para Análise de Sentimento. A Mineração de Textos é um novo e excitante ramo de pesquisa multidisciplinar que tenta resolver a crise de sobrecarga de informação combinando técnicas de mineração de dados (*Big Data*), Aprendizagem de Máquina (*Machine Learning*), Processamento de Linguagem Natural, Recuperação de Informações e Gestão do Conhecimento. Dentre suas várias aplicações, uma técnica chamada de Análise de Sentimento (AS) ou Mineração de Opinião (MO) se destaca, devido a sua ampla aplicabilidade: classificação de opiniões de um modo geral, análise de satisfação, métricas de rede social, previsão do mercado de ações, etc. Nesse escopo, grande parte dos classificadores de emoções tem sua formulação matemática baseada em conceitos e teorias físicas, como a classificação por entropia máxima, por métodos associados ao modelo de Ising e as já consagradas redes neurais[3, 4, 5, 6].

Nessa linha, esse trabalho tem como objetivo apresentar a formulação de um novo modelo de classificação de sentimentos a partir da abordagem léxica, e seus resultados numéricos. Tal proposta de classificação é intitulada pelo autor como Modelo Neural de Múltiplos Canais (MONEMC), visto que esse tem como base e inspiração um modelo formal de neurônios intitulado Modelo de Gerstner-Galves-Löcherbach, cuja principal característica é sua estocasticidade intrínseca[7, 8]. Além disso, será descrito a aprendizagem do dicionário léxico e uma simples comparação entre: o modelo desenvolvido, o método Bayesiano e a classificação do próprio ser humano. Vale ressaltar que esse algoritmo faz parte de um projeto vencedor da segunda edição do *Mining Lab-Nexa (2018)* na categoria de Inteligência Artificial, tendo como um dos propósitos inferir e prever comportamentos característicos do mercado financeiro através da análise de textos relacionados.



**Parte I**

**Econofísica**



Nesta parte serão apresentados os conceitos básicos relativos a Econofísica, em especial, aos que se referem a distribuição de renda e indicadores de desigualdade. O capítulo 3 intitulado Teoria Cinética e Distribuição de Renda no Brasil parte dessa dissertação refere-se ao artigo “*Kinetic Theory and Brazilian Income Distribution, 2019*”, estudo realizado pelo mesmo autor dessa qualificação e seu orientador e aprovado para publicação na revista “*Physica A: Statistical Mechanics and its Applications - Journal - Elsevier*”[9].

O capítulo 4 será dedicado à apresentação de um novo modelo para a distribuição de riqueza partindo do princípio da concentração de riqueza. Esse último, envolvendo agentes interagentes segregados em duas classes distintas conforme os seus respectivos processos de geração de riqueza.



## 2 Introdução a Econofísica e aos modelos de agentes para distribuição de riqueza

### 2.1 Econofísica e a distribuição de renda e riqueza

Na economia moderna desproporções na distribuição de renda e riqueza ao longo das sociedades representam fontes de instabilidade social. Desse fato, pode-se entender o motivo do tópico em questão receber tamanho destaque em se tratando de pesquisas relativas à Econofísica, em especial, a partir do século XIX com as ideias de Vilfredo Pareto[2].

Um dos primeiros marcos no estudo da distribuição de renda ocorreu quando o economista Vilfredo Pareto em seu *Cours d'Économie Politique* postulou que a distribuição assintótica de renda (classe mais rica) em todos os países e em todos os tempos deveria respeitar um comportamento de lei de potencia[10]. Esta hipótese foi verificada empiricamente ao longo dos anos que a sucederam em diversos contextos e regiões do mundo[11, 12, 13, 9]. Em 1931, Gibrat propôs um modelo de renda como um processo multiplicativo aleatório que resulta em uma distribuição estacionária do tipo log-normal[14]. Embora a lei do Gibrat forneça bons resultados para a renda média, a distribuição log-normal não é estacionária [15].

Yakovenko e outros pesquisadores, através de uma análise semi-empírica, desenvolveram a Teoria da Econofísica de "Duas Classes", que afirma que a distribuição de probabilidade da riqueza/renda/dinheiro deveria ser dividida em duas partes. Desta forma, a distribuição de renda é aproximadamente exponencial (distribuição de Boltzmann-Gibbs) e as regiões assintóticas de alta renda obedecem a uma Lei de potencia[16, 17, 18] Há também outras distribuições que quando combinadas com a lei de Pareto se mostraram bem sucedidas na caracterização da riqueza, bem como, a divisão da população em três classes [19, 20]. Esses casos particulares refletem um consenso geral de que a lei de Pareto é de fato aplicável à fração populacional de alta renda, enquanto o resto da sociedade é governado por leis não universais, como funções exponenciais, log-normal e Gompertz[2]

Em outra linha de ação, alguns econofísicos foram tentados a estudar a evolução de uma única variável, chamada renda, dinheiro ou riqueza, baseados em processos estocásticos[21, 22], dentre os quais destacamos os trabalhos de Solomon. Ele e colaboradores, valendo-se do o modelo Lotka-Volterra Generalizado (LVG) com competição autocatalítica, foram capazes de derivar a distribuição de Pareto com o expoente

insensível à mudanças em alguns dos seus parâmetros. Além disso, também conseguiram fornecer uma interpretação robusta para o modelo, baseado-se no contexto econômico[23, 24, 25]. Depois disso, Huang generalizou o modelo LVG para incluir riqueza negativa[26] e Willis forneceu alguns resultados empíricos associados à teoria econômica padrão[27].

Apesar de enfrentar algumas críticas pelas simplificações excessivas comuns em muitos modelos de econofísica, onde a diversidade de bens é reduzida a uma única variável - riqueza, dinheiro ou renda[28] — esforços recentes foram feitos para refinar essas interpretações, especialmente em termos de distinguir entre riqueza e renda do trabalho. Berman, em sua pesquisa, aborda essas distinções e propõe um modelo baseado em três fontes primárias de capital que geram riqueza: Renda do Trabalho, Renda de Capital e Mudança no Valor do Capital[29].

## 2.2 Modelos de distribuição de renda e riqueza

### 2.2.1 Lei de Pareto

Vilfredo Pareto cientista político, sociólogo e economista italiano em seu livro *Cours d'économie politique*[10] publicado em 1897 formulou uma lei de distribuição de riqueza associada ao número de indivíduos  $N$  que possuem uma renda maior do que  $m$ :

$$N = \frac{A}{(m + b)^\alpha} \quad (2.2.1)$$

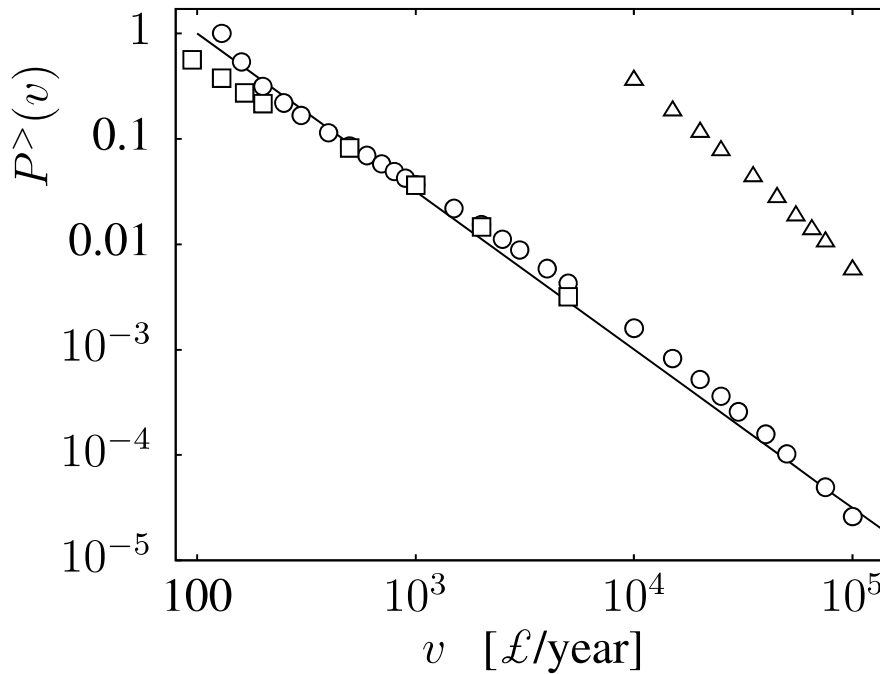
onde  $b$  é uma constante fixada como 0 na grande maioria dos estudos subsequentes e  $\alpha$  representa o popularmente conhecido expoente de Pareto, esse último assumindo valores entre 1 e 2 na grande maioria das sociedades[30, 18, 9]. Pareto ainda postulou que o comportamento referido seria válido apenas a partir de um certo valor de corte, i.e., para a parcela mais rica da sociedade. Pesquisas recentes apontam que essa porção representa cerca de 5~10% da população total[30, 15].

$$P(m' > m) \equiv Prob(\text{renda} > m) \sim m^{-\alpha}, \quad m \rightarrow \infty \quad (2.2.2)$$

a lei de potencia representada pela eq. 2.2.2 ao longo dos anos que a sucedeu foi e vem sendo ,constantemente, confirmada empiricamente em diferentes sociedades e períodos da história.



Figura 2.2.1 – Exemplo de distribuição de renda



Fonte: SLANINA, FRANTIESEKSlanina,, 2014.

Figura 2.2.1: Distribuição de renda acumulada complementar na Grã-Bretanha a partir de receitas fiscais. Os símbolos denotam anos 1801 ( $\square$ ), 1911-1912 ( $\triangle$ ) e 1918-1919 ( $\circ$ ). Para a linha indicada  $m^{-1.5}$ . Observe-se que os pontos para os anos 1801 e 1918-1919 recaem essencialmente na mesma curva.

### 2.2.2 Modelo de Duas Classes

O Modelo de Duas Classes proposto por Yakovenko e demais pesquisadores tem apresentado grande êxito no estudo da distribuição de renda e desigualdade social[17, 31, 13, 15, 32].

A economia é um grande sistema com milhões de agentes participantes, por isso é um alvo promissor para aplicações de Mecânica Estatística. Existe um quantidade conservada na economia? Nesse sentido os autores argumentam que tal quantidade conservada é o dinheiro  $m$ . Isto posto, partindo do conceito fundamental da conservação do dinheiro, uma analogia com a teoria cinética dos gases é proposta quando o dinheiro  $m$  é substituído pela energia  $\varepsilon$ . Na teoria cinética clássica, a distribuição de Boltzmann-Gibbs(DBG)  $\overline{f(\varepsilon)}$  é a mais provável para um gás em equilíbrio de  $N$  elementos contidos em uma caixa. Essa consideração é independente da forma da partícula ou dos detalhes sobre as interações que ocorrem durante as colisões elásticas (conservação de energia) e por esse motivo a DBG é universal. Portanto, uma função

inicial de distribuição de dinheiro arbitrária deve atender às condições:

$$\sum_{i=1}^I \eta_i = N \quad (2.2.3)$$

$$\sum_{i=1}^I m_i \eta_i = M \quad (2.2.4)$$

onde  $I$  é o número de classes de renda,  $m_i$  é a renda média da  $i$ -ésima classe,  $\eta_i$  é o número de agentes da  $i$ -ésima classe,  $N$  o número total de agentes e  $M$  é o dinheiro total.

Para satisfazer as duas condições anteriores, a distribuição de dinheiro  $P(m)$  deve ser exponencial, assim como a distribuição de equilíbrio de Boltzmann-Gibbs. Matematicamente após alguns passos:

$$P(m) = C \exp(-\lambda m_i) \quad (2.2.5)$$

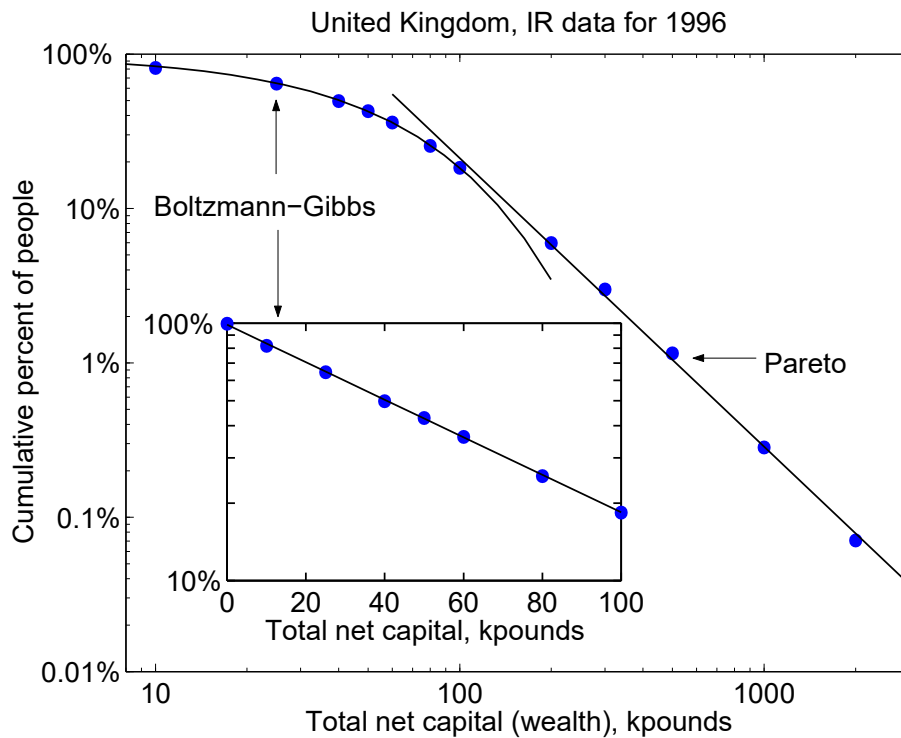
onde  $C$  é uma constante e  $\lambda$  é um multiplicador Lagrange, Assim, normalizando  $\int_0^\infty P(m) dm = 1$  e calculando a renda média  $\langle m \rangle = \int_0^\infty m P(m) dm$ . Finalmente, podemos escrever a distribuição de renda pessoal como:

$$P(m) = \frac{1}{T} \exp(-m/T) \quad (2.2.6)$$

note que a “temperatura”  $T$  é igual ao dinheiro média [15, 31, 32].

Assim o modelo de Duas Classe propõem que a distribuição do dinheiro  $m$  de uma população é dividida em duas classes: Lei de Pareto ( $m \rightarrow \infty$ ) e Distribuição de Boltzmann-Gibbs.

Figura 2.2.2 – Exemplo de distribuição de dinheiro: Modelo de Duas Classes

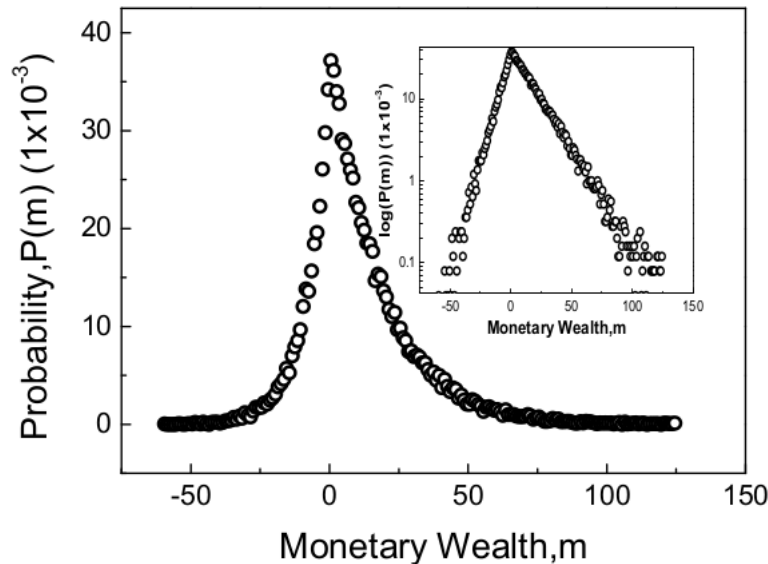


**Fonte:** Yakovenko, V.M., 2009.

Figura 2.2.2: Distribuição acumulada complementar de probabilidade de dinheiro em kpounds no Reino Unido mostrada em escala log-log (painel principal) e log-linear (inset). Os pontos representam os dados das receitas, e as linhas sólidas são os ajustes exponencial (Boltzmann-Gibbs) e lei de potência (Pareto).

Num contexto mais amplo, quando passamos a considerar não somente a renda do trabalho e outras fontes, mas também os débitos associados, naturalmente, devemos ampliar as proposições para incluir valores negativos (ver figura 2.2.3).

Figura 2.2.3 – Exemplo de distribuição de dinheiro: Modelo de Duas Classes com dinheiro negativo



Fonte: Yakovenko, V.M, 2009.

Figura 2.2.3: Densidade de probabilidade do dinheiro numa situação em que os débitos são considerados. Os pontos representam dados vindos de uma simulação.

Para Yakovenko, quando os agentes tomam empréstimos de bancos considerados reservatórios de dinheiro, seu saldo monetário (dinheiro positivo) aumenta, mas eles também adquirem uma obrigação de dívida (dinheiro negativo), mantendo o patrimônio líquido total inalterado.

Essa dinâmica altera a condição de limite do sistema  $m \geq 0$ , permitindo saldos negativos de patrimônio líquido e resultando em uma distribuição de dinheiro que nunca se estabiliza, com a riqueza se distribuindo ilimitadamente entre valores extremamente positivos e negativos. Apesar dessa modificação nas condições de limite, uma lei de conservação generalizada do patrimônio líquido ainda se aplica, com algumas pessoas se tornando mais ricas à custa do endividamento de outras[33].

### 2.2.3 O Modelo Lotka-Volterra Generalizado

O sistema LVG no contexto econômico descreve a evolução da riqueza dos  $N$  indivíduos  $W_i(t)$  a cada passo de tempo discreto  $t$ :

$$\begin{aligned}
 W_i(t+1) &= (1 + \lambda(t))W_i(t) + \sum_{(j \neq i)}^N G_{ij}W_j(t) - \sum_{(j \neq i)}^N G_{ji}W_i(t) - \sum_{(j \neq i)}^N C_{ij}W_i(t)W_j(t) \\
 W_j(t+1) &= W_j(t), \quad , j = 1, \dots, N; \quad j \neq i.
 \end{aligned}
 \tag{2.2.7}$$

No caso simples, onde todos os agentes interagem entre si na mesma taxa,  $G_{ij} = G/N$  e  $C_{ij} = C/N$ . Devido a essa escolha, reescrevemos 2.2.7:

$$\begin{aligned} W_i(t+1) &= (1 + \lambda(t))W_i(t) + G\overline{W}(t) - CW_i(t)\overline{W}(t) \\ W_j(t+1) &= W_j(t), \quad j = 1, \dots, N; \quad j \neq i. \end{aligned} \quad (2.2.8)$$

Na equação anterior, Solomon estabelece a aproximação  $\lambda(t) = \lambda(t) - G$ .

O processo multiplicativo  $\lambda$  é extraído de uma função de probabilidade  $P(\lambda)$  e é expresso por:

$$\lambda(t) = \bar{\lambda} + \eta(t). \quad (2.2.9)$$

No contexto econômico,  $\lambda$  representa a riqueza originada devido ao investimento no mercado de ações. O termo  $\eta(t)$  é um processo estocástico de ruído branco independente, normalmente distribuído  $\mathcal{N}(0, D)$ , com as propriedades usuais. Além disso,  $\bar{\lambda}$  é o valor médio de  $P(\lambda)$  que é independente de  $t$  e  $i$ , e  $D$  é o seu desvio padrão dado por:

$$D = \overline{\lambda^2} - \bar{\lambda}^2. \quad (2.2.10)$$

Os segundo e terceiro termos de 2.2.7 descrevem as trocas entre firmas ou investidores  $i$  e  $j$ , ou seja, a quantidade de riqueza que os agentes  $j$  gastam comprando a produção e serviços do agente  $i$  ou ganham vendendo os seus.

O último termo  $\sum_{(j \neq i)}^N C_{ij}W_i(t)W_j(t)$  é um fator de saturação devido ao tamanho econômico finito que controla o crescimento geral da riqueza no sistema. No entanto, ele quebra a simetria sob reescala de riqueza e, por essa razão, trunca o comportamento da lei de potência de Pareto para grandes riquezas[25].

Além disso, outro ponto importante na dinâmica LVG é a atualização assíncrona, ou seja, apenas um  $W_i$  é escolhido aleatoriamente para atualização. Além disso, um ganho em  $W_i$  não requer necessariamente uma perda imediata pelos outros  $W_j$  durante interações binárias. Em termos econômicos, uma negociação pode ser mutuamente benéfica, mesmo quando os mercados não são perfeitamente eficientes.

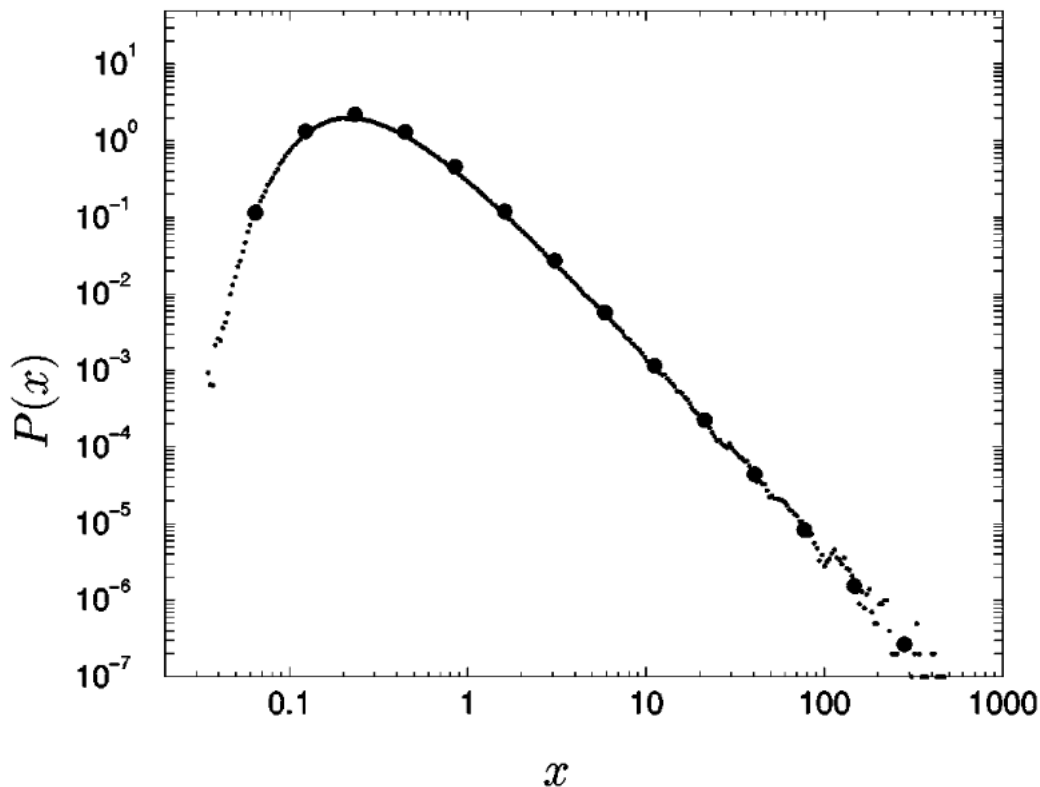
Para se estudar a distribuição de probabilidade de equilíbrio do modelo LVG, devemos utilizar a riqueza normalizada  $x$ :

$$x = \frac{W_i}{\overline{W}} \quad (2.2.11)$$

Aplicando esta mudança de variáveis e a partir das equações do modelo, encontramos que a distribuição segue uma lei de potência com um termo exponencial (ver figura 2.2.4), dado por:

$$P(x) = x^{-1-\alpha} \exp\left(\frac{1-\alpha}{x}\right) \quad (2.2.12)$$

Figura 2.2.4 – Modelo LVG



Fonte: Malcai, O., 2002.

Figura 2.2.4: Densidade de probabilidade da riqueza no modelo LVG. Os pontos menores representam dados vindos de uma simulação e, os maiores da equação teórica.

O expoente de Pareto  $\alpha$  depende de um único parâmetro  $\frac{G}{D}$ , que é a razão do coeficiente global de crescimento pelo desvio das flutuações. Isto é expresso como:

$$\alpha = 1 + \frac{2G}{D} \quad (2.2.13)$$

Com isso, podemos analisar como a riqueza se distribui entre os indivíduos e sob quais condições a distribuição de Pareto é mantida ou não no contexto do modelo LVG.

#### 2.2.4 O Modelo Duplo Pareto

O pesquisador Alexis Akira Toda[34] propõe um modelo para a dinâmica da renda que gera uma distribuição estacionária dupla de Pareto. A evolução do logaritmo da renda,  $X$ , é modelada como um movimento browniano com deriva, representado pela equação diferencial estocástica:

$$dX(t) = -\lambda \text{sgn}(X(t))dt + \sigma dB(t), \quad (2.2.14)$$

onde  $\text{sgn}(X(t))$  é a função sinal,  $\lambda > 0$  é uma constante positiva que representa a força da retração em direção à média,  $\sigma$  é a volatilidade do processo  $B(t)$  que é um Movimento Browniano Padrão.

A equação de Fokker-Planck associada, descrevendo a evolução temporal da densidade de probabilidade  $p$  para  $X$ , é dada por:

$$\frac{\partial p}{\partial t} = \lambda \text{sgn}(X) \frac{\partial p}{\partial X} + \frac{\sigma^2}{2} \frac{\partial^2 p}{\partial X^2}.$$

Sob condições estacionárias, encontramos a distribuição  $p(x)$ :

$$p(x) = \frac{\lambda}{2} e^{-\lambda|x|}, \quad (2.2.15)$$

onde  $\alpha = \frac{2\lambda}{\sigma^2}$ . Esta distribuição é conhecida como a distribuição de Laplace. Quando retornamos à variável original de renda através da exponenciação de  $X$ , obtemos a distribuição dupla de Pareto para renda  $y$ :

$$p(y) = \begin{cases} \frac{\alpha}{2} y^{-1-\alpha} & \text{para } y > 1, \\ \frac{\alpha}{2} y^{\alpha-1} & \text{para } 0 < y \leq 1. \end{cases}$$

O expoente de Pareto  $\alpha$  é diretamente proporcional à volatilidade  $\sigma$  e inversamente proporcional à constante de retração  $\lambda$ , ilustrando como a distribuição de renda pode ser ajustada por esses parâmetros.

Este modelo é significativo por sua capacidade de explicar as características empíricas das distribuições de renda observadas em dados econômicos, proporcionando uma estrutura para a análise da desigualdade de renda e das dinâmicas econômicas subjacentes.





## 3 Teoria cinética e distribuição de renda no Brasil

### 3.1 Modificações no Modelo de Duas Classes

Neste capítulo apresentaremos uma investigação sobre a distribuição de renda pessoal brasileira utilizando dados da Pesquisa Nacional por Amostra de Domicílios (PNAD), uma pesquisa anual disponibilizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) que fornece características gerais da população do país. Usando a base de dados da PNAD, também confirmamos a eficácia de um modelo semi empírico que concilia a de Pareto para pessoas de alta renda e a distribuição de Boltzmann-Gibbs para o resto da população, o Modelo de Duas Classes. Assim partindo das equações:

$$P(m) = \frac{A}{m^{(1+\alpha)}} \quad (3.1.1)$$

$$P(m) = \frac{1}{T} \exp(-m/T) \quad (3.1.2)$$

expressões essas que representam as Leis de Pareto e de Boltzmann-Gibbs, respectivamente, propomos então a inserção de um parâmetro  $m_l$  representando o corte entre as distribuições de probabilidade. Como poderá ser visto ao longo do texto, esse valor poderá conjuntamente a renda média  $\langle m \rangle$ , o expoente de Pareto  $\alpha$  e o índice de Gini  $G$  exercer o papel de indicador de desigualdade social. Assim:

$$P(m) = \frac{1}{T} \theta(m_l - m) \exp(-m/T) + \frac{A\theta(m - m_l)}{m^{(1+\alpha)}} \quad (3.1.3)$$

na equação acima  $\theta$  é a função de Heaviside,  $m_l$  é o valor de corte e  $A$  uma constante empírica. Ainda, por conveniência, derivamos a distribuição acumulada complementar de probabilidade a partir de 3.1.3  $P(m' \geq m)$  :

$$P(m' \geq m) = \int_m^{\infty} P(m) dm \quad (3.1.4)$$

assim,  $P(m' \geq m) = \exp(-m/T) - \exp(-m_l/T) + \frac{Am_l^{-\alpha}}{\alpha}$  se  $m \leq m_l$  e  $P(m' \geq m) = \frac{Am^{-\alpha}}{\alpha}$  se  $m \geq m_l$ . Reconhecendo que na grande maioria dos casos  $\alpha \sim 1.5$  e obviamente  $\frac{m_l}{T} > 1$ , nós estimamos que:

$$P(m' \geq m) \cong \begin{cases} \exp(-m/T) & m \leq m_l \\ Bm^{-\alpha} & m \geq m_l \end{cases} \quad (3.1.5)$$

isto é, a probabilidade de uma pessoa possuir uma renda individual maior ou igual a  $m$ .

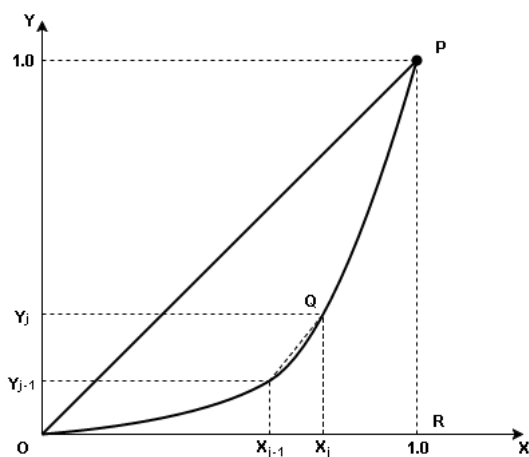
## 3.2 Indicadores de desigualdade social

Por meio do modelo semi-empírico anterior, podemos tomar três parâmetros úteis para medir a desigualdade de renda entre subgrupos de uma mesma população ou até mesmo comparar diferentes populações, o índice de Pareto, a “temperatura” e o valor de corte  $m_j$ . O índice de Pareto está diretamente associado às pessoas mais ricas e, portanto, é limitado por esse intervalo. Ainda assim,  $\alpha$  torna-se muito conveniente na realização de comparações e até mesmo análises de evolução temporal, como pode ser visto em alguns estudos[16, 35, 11]. Nesse sentido, aplicando log em ambos os lados da expressão 3.1.1, temos:  $\log(P(m)) = \log(A) - (1 + \alpha) \log(m)$ , onde  $A$  é uma constante. Assim, podemos notar que um gráfico em escala log-log da Lei de Pareto possui um comportamento linear como pode ser observado 2.2.1. Como consequência, quanto maior o valor de  $\alpha$  menor será a desigualdade de renda entre os ricos.

A “temperatura” e o valor de corte podem nos fornecer uma boa perspectiva sobre a desigualdade de renda entre subgrupos pertencentes a uma determinada população; no entanto, devido às taxas de câmbio entre a moeda base e as moedas estrangeiras, a inflação e outros aspectos complexos, o rendimento médio e o limiar de corte entre as distribuições sem um tratamento adequado não podem contribuir muito mais do que isso.

Além do modelo proposto, o coeficiente de Gini (G) tem sido o instrumento mais utilizado para medir a desigualdade de renda na literatura[36, 37].

Figura 3.2.1 – Curva de Lorenz



Fonte: Próprio autor.

Figura 3.2.1: Curva de Lorenz em escala normalizada:  $X$  percentual acumulado de pessoas e  $Y$  é o percentual acumulado da renda total.

Na figura 3.2.1 a simples relação  $Y = X$  representa a situação onde a renda é perfeitamente distribuída entre a população, i.e, os 20% “mais pobres” possuem o mesmo percentual de renda e assim sucessivamente.[36]. Uma outra função comumente conhecida como curva de Lorenz representa o comportamento das sociedades

reais. Isto posto, o coeficiente de Gini é obtido pelo cálculo da razão entre a área limitada pela linha diagonal e a Curva de Lorenz (OPQ) e, a área do triângulo (OPR). Em outras palavras,  $G = \frac{AreaOPQ}{AreaOPR}$ . Como resultado, o problema chave passa a ser o de se estimar o valor da área abaixo da curva de Lorenz. Numericamente, esse resultado pode ser obtido a partir de sucessivas aproximações trapezoidais  $S_j$ . Note que para cada par  $(X, Y)$  a baixo da curva de Lorenz podemos construir  $N-1$  trapézios[38]. Onde o  $j$ th trapézio possui base  $Y_j$  e  $Y_{j-1}$  e altura  $(X_j - X_{j-1})$ . Portanto,

$$S_j = \frac{(Y_j + Y_{j-1})(X_j - X_{j-1})}{2} \quad (3.2.1)$$

Em vista da equação anterior, podemos definir graficamente o coeficiente de Gini pela fórmula de Brown:

$$G = 1 - 2 \sum_j^{N-1} S_j = 1 - \sum_j^{N-1} (Y_j + Y_{j-1})(X_j - X_{j-1}) \quad (3.2.2)$$

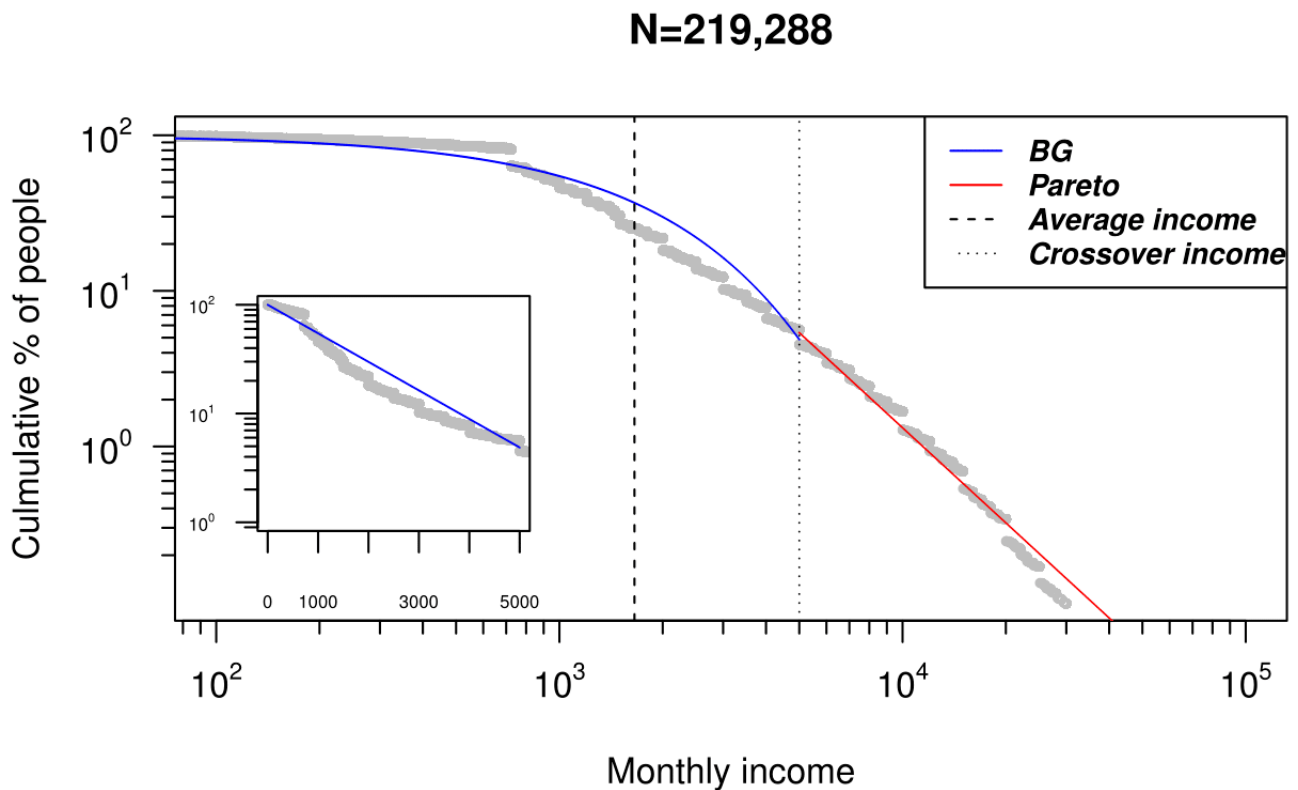
Embora o coeficiente de Gini contribua positivamente para a avaliação da desigualdade de renda, ele também apresenta algumas limitações. A desvantagem mais evidente está relacionada com a sua alta sensibilidade à mudanças, especialmente na classe média.

### 3.3 Distribuição de renda no Brasil

Nesta seção, investigaremos a distribuição da renda pessoal brasileira usando microdados da Pesquisa Nacional por Amostra de Domicílios (PNAD), disponibilizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE). A PNAD é uma pesquisa anual que estuda aspectos gerais da sociedade brasileira quanto a trabalho, renda, educação e outros [39]. Para a realização da análise, extraímos dos dados da PNAD o valor da renda mensal de todas as fontes para pessoas com idade superior a 10 anos ( $m$ ) e também filtramos esse valor para quatro subconjuntos: pretos / indígenas / pardos, brancos / amarelos, homens e mulheres. Depois que obtivemos a variável renda, rejeitamos as pessoas sem renda, assim como os valores ausentes no novo subconjunto. Desta forma, uma parte dos resultados do trabalho é sobre a população economicamente ativa.

No caso específico do ano de 2014, computamos 362.625 casos, dos quais, seguindo os critérios mencionados anteriormente, trabalhamos com 219.288 deles. Inicialmente, para fornecer evidências da eficácia do modelo descrito na seção 5.1, calculamos a distribuição acumulada da renda para toda a população brasileira, como pode ser visto na figura a seguir.

Figura 3.3.1 – PNAD 2014



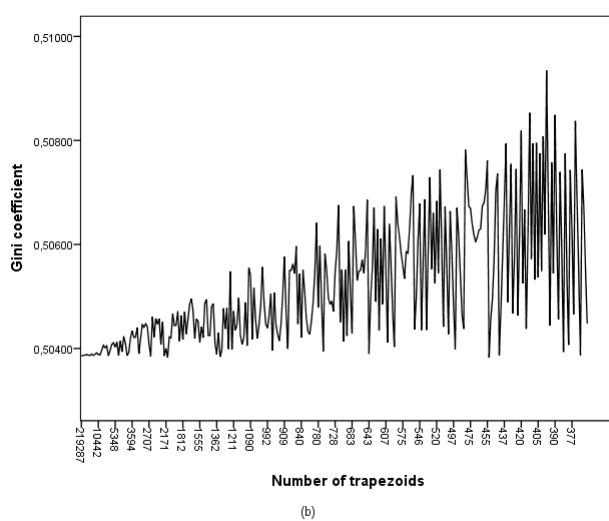
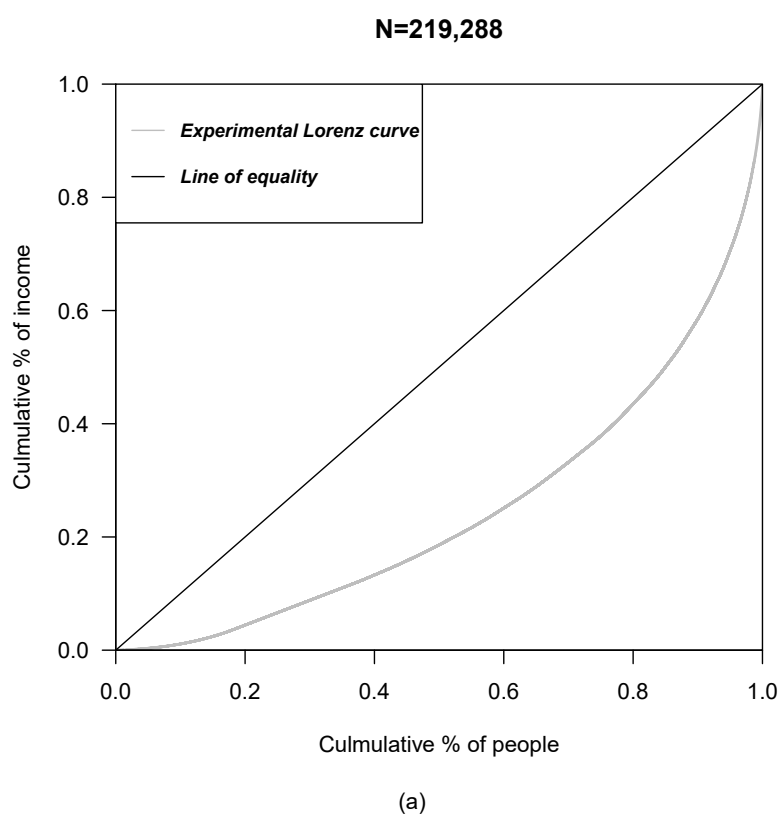
Fonte: Próprio autor.

Figura 3.3.1: Distribuição acumulada de probabilidade da renda em escala log-log. Os pontos cinzas representam os valores extraídos da PNAD 2014 e a linha sólida corresponde ao modelo expresso pela equação 3.1.5, onde a região de lei de potencia é diferenciada pela cor vermelha representando quase 5% da população. O restante do gráfico (linha azul) refere a região com DBG. Os parâmetros destacados na figura são a renda média  $R\$ 1655 \pm 6$  e o valor de corte  $R\$ 5000$ . Ainda obtivemos  $\alpha = 2.181 \pm 0.011$ .

Para o índice de Pareto obteve-se o valor de 2.181 com erro padrão de 0,011 , além disso, encontrou-se um intervalo de confiança de 95% de 0,024 em torno dessa medida. Esses valores foram calculados seguindo o procedimento *bootstrap*, onde utilizamos metade do número de casos, selecionados aleatoriamente, na região mais rica (5% da população total).

Para completar o conjunto das medidas de desigualdade para a população brasileira, calculou-se o índice de Gini a partir da fórmula de Brown e do método da extrapolação oscilatória convergente[40]. Através dessa metodologia uma estimativa do grau de incerteza associado ao índice de Gini tornou-se possível. Assim,  $G = 0.504$  com um erro praticamente nulo para a população total através dos dados da PNAD 2014.

Figura 3.3.2 – Curva de Lorenz e Coeficiente de Gini: PNAD 2014



Ainda, procedimento de análise similar foi desenvolvido para duas diferentes dicotomias da população brasileira. A primeira divisão refere-se a um recorte por sexo (homem e mulher) e a segunda por cor/raça (preto/pardo/indígena e branco/amarelo) como é expresso na tabela 3.3.1.

Tabela 3.3.1 – PNAD de 2014

Grupo subgrupo	Sexo/Gênero		Raça/Cor	
	Homem	Mulher	BA	PPI
coeficiente de Gini	0.497	0.497	0.519	0.457
Temperatura (R\$)	1940 ± 9	1359 ± 6	2144 ± 11	1278 ± 5
Índice de Pareto	2.172 ± 0.012	2.247 ± 0.020	2.306 ± 0.017	2.187 ± 0.012
Valor de Corte (R\$)	6000	4010	7000	3510

Fonte: Próprio autor.

Tabela 3.3.1: Parâmetros de desigualdade para cada subgrupo da população no ano de 2014 onde BA significa brancos e amarelos e PPI pretos, pardos e indígenas. Para a construção da tabela adotou-se o mesmo critério aplicado a população total. Desse modo, fixou-se com boa precisão o valor de corte  $m_l$  como sendo o menor valor de renda mensal para os 5% mais ricos de cada subgrupo da população.

Como pode ser visto na Tabela acima, para o grupo de gênero, não há diferença entre o coeficiente de Gini medido para homens e mulheres. Para os 5% mais ricos, observa-se uma pequena melhoria na distribuição de renda para as mulheres economicamente ativas se a compararmos com a distribuição de renda dos homens. Essa pequena melhora é estatisticamente significativa ao considerarmos o intervalo de confiança de 95% de 0,024 em torno do índice de Pareto de ambos: homens e mulheres. No entanto, as diferenças entre homens e mulheres são mais expressivas na avaliação dos valores da renda mensal, já que tanto a renda de corte quanto a “temperatura” são muito maiores para os homens. Esse contraste entre os parâmetros de renda de homens e mulheres pode ser melhor compreendido se notarmos que no ano de 2014, através dos dados da PNAD, cerca de 51,5% dos entrevistados eram mulheres, mas apenas 40% do total da renda mensal pertencia a elas.

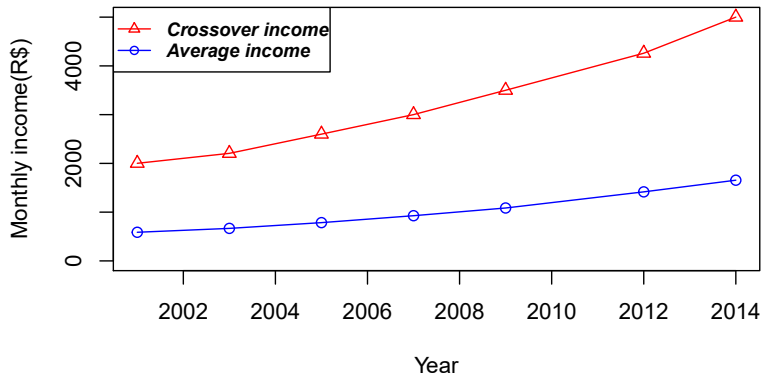
Quanto a separação da população por cor, nota-se que a desigualdade em geral é menor para o subgrupo PPI. Mas, agregando-se a esse resultado a diferença entre as rendas médias dos subgrupos PPI e BA que representa quase 68% da renda média mensal PPI, e o contraste apresentado no valor de corte  $m_l$  nos mesmos subgrupos, podemos reconhecer que o menor coeficiente de Gini e maior índice de Pareto calculado para PPI reflete o fato de que as pessoas pertencentes a esse subconjunto estão concentradas nas classes mais pobres, seja na região da BG ou na região de Pareto. Para o grupo separado por cor, 57,5% dos entrevistados da PNAD são autoidentificados como pretos, indígenas ou mestiços, enquanto cerca de 44% da renda mensal total está concentrada nesse subgrupo.

### 3.4 Séries temporais dos coeficientes de desigualdade

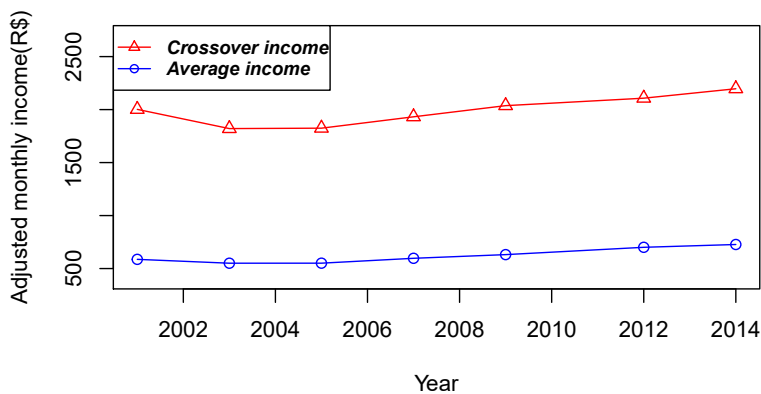
Para ampliar a compreensão sobre a distribuição de renda no Brasil, nesta seção investiga-se a dinâmica dos parâmetros de desigualdade ao longo dos anos 2001-2014. Para isso a mesma fonte de dados PNAD é utilizado e ainda, são respeitados os critérios de limpeza de dados, procedimento estatístico e formação de grupos previamente adotados para a PNAD de 2014.

A primeira observação interessante é sobre a região de Pareto, que permaneceu relativamente constante nos anos selecionados, variando entre 4% a 6% da população total com os maiores valores de renda. Por essa razão, o valor de corte  $m_l$  é fixado como a renda mensal mais baixa dos 5% mais ricos. O próximo resultado refere-se à série temporal dos valores obtidos para a temperatura (renda média) e o valor de corte da população total, ambos medidos em moeda local (Real) sem correção inflacionária (3.4.1.a) e quando tais parâmetros de desigualdade são transformados em renda real por ajuste feito através do IPCA (Índice de Preços ao Consumidor), vide figura 3.4.1.b.

Figura 3.4.1 – Série temporal da renda média e Valor de Corte



(a)



(b)

**Fonte:** Próprio autor.

Figura 3.4.1: (a) Série temporal para a renda média e o valor de corte calculado através dos anos: 2001, 2003, 2005, 2007, 2009, 2012, 2014. (b) Série temporal da renda média e do valor de corte ajustados pelo IPCA.

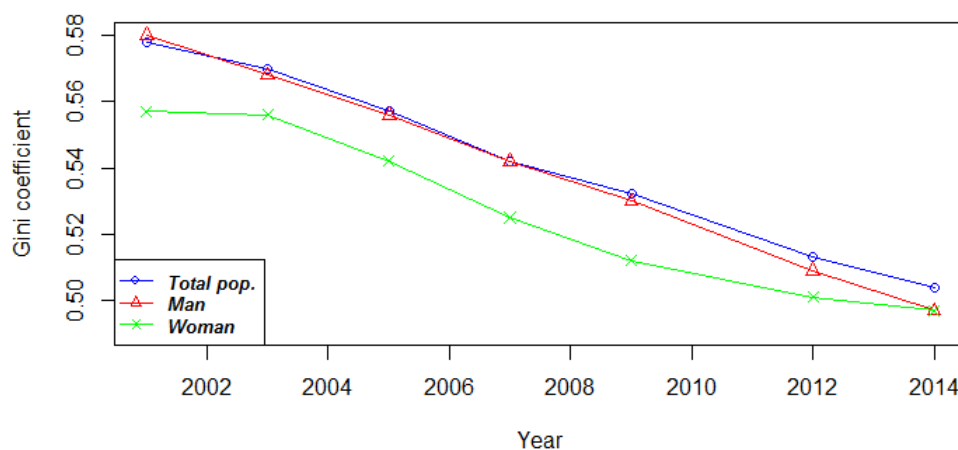
Nesse contexto, analisando 3.4.1, pode-se observar que a diferença percentual entre a “temperatura” e o valor de corte no ano de 2001 correspondeu a 241% da renda média. Para o último ano estudado (2014) essa diferença percentual diminuiu para 202%. Quando os valores do rendimento médio e da renda de corte são ajustados pelo IPCA, pode-se perceber um aumento percentual na renda média de 24%, enquanto que para o valor de corte esse aumento foi de, aproximadamente, 10% entre 2001 e 2014. Essa indicativa sugere, fortemente, que a distribuição de renda melhorou ao longo deste período de tempo.

Doutro modo, concentrando-se na série temporal do coeficiente de Gini, estimamos para o seu valor para mesmos subconjuntos de antes: sexo (homem, mulher) e

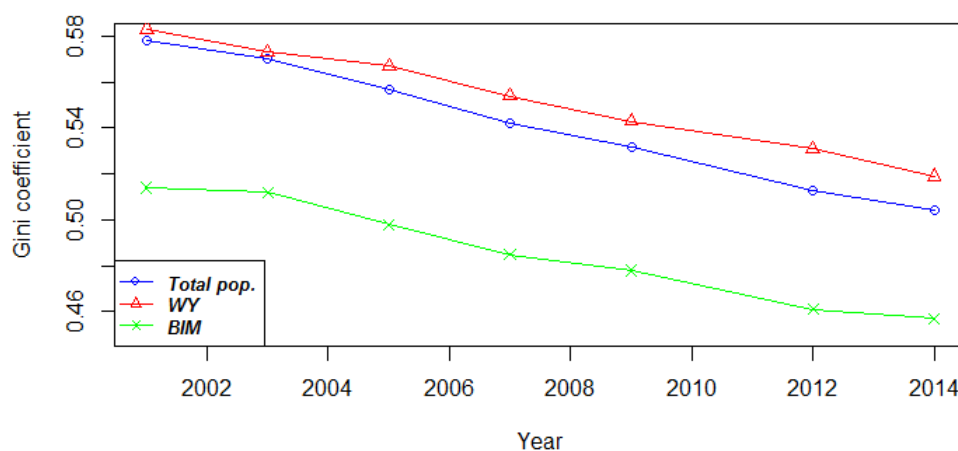


cor (PPI, BA).

Figura 3.4.2 – Série temporal Coeficiente de Gini



(a)



(b)

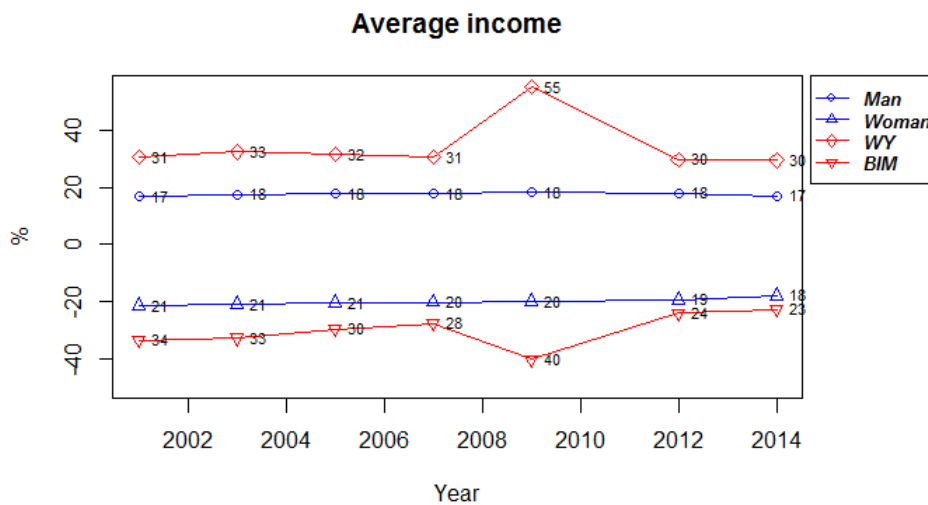
Fonte: Próprio autor.

Figura 3.4.2: (a) Coeficiente de Gini para a população total dividida por gênero. (b) Coeficiente de Gini para a população total dividida por cor.

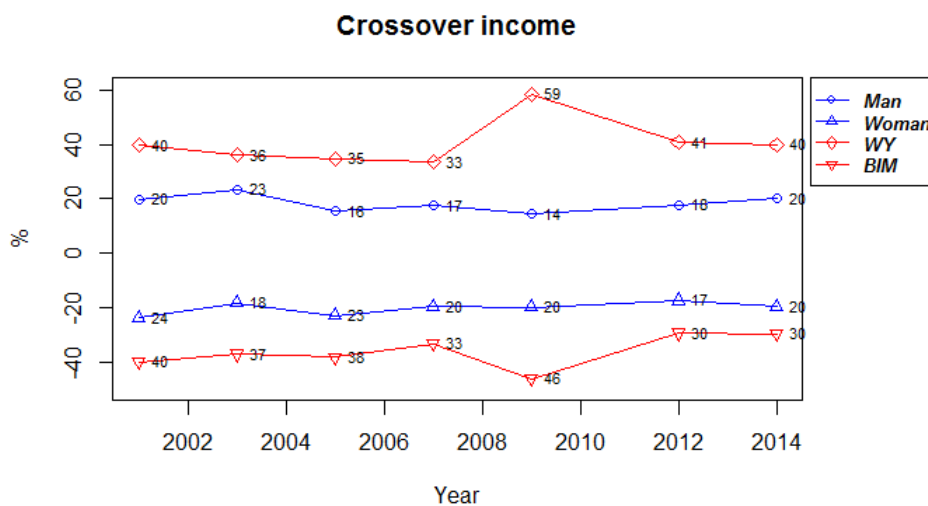
Para todos os subgrupos considerados uma melhoria significativa do coeficiente de Gini nestes 14 anos é observada. Em específico para a população total, este parâmetro de desigualdade mudou de 0,578 para 0,504, um desenvolvimento de aproximadamente 13%. Esses resultados sugerem uma diminuição da desigualdade da renda brasileira. Uma das possíveis razões para essa melhoria é um aumento no número de pessoas na classe de renda média, onde o valor do coeficiente de Gini é mais sensível a mudanças. Além disso, como pode ser verificado em 3.4.2,

os subgrupos historicamente desfavorecidos (mulher e PPI) apresentam um melhor coeficiente de Gini. No caso de gênero, a diferença entre o coeficiente de Gini ao longo dos anos selecionados foi decrescente e, em 2014, tornou-se zero. Portanto, outra questão pertinente foi levantada: como o dinheiro é distribuído entre esses subgrupos? Para responder a essa pergunta, avalia-se a diferença percentual da renda média e da renda de corte para cada subgrupo em comparação com a população total.

Figura 3.4.3 – Diferença Percentual Relativa a População Total



(a)



(b)

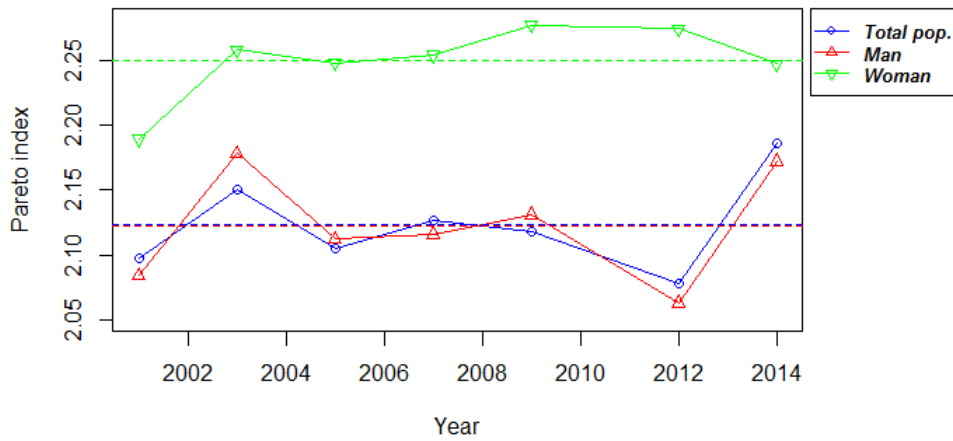
Fonte: Próprio autor.

(a) diferença percentual na renda média em relação a população total. O erro padrão adotado para o gráfico é de 2% obtido pelo cálculo do erro propagado para cada medida onde escolhe-se o maior valor de incerteza entre todos os erros. (b) diferença percentual no valor de corte em relação a população total. O erro padrão é também de aproximadamente 2%.

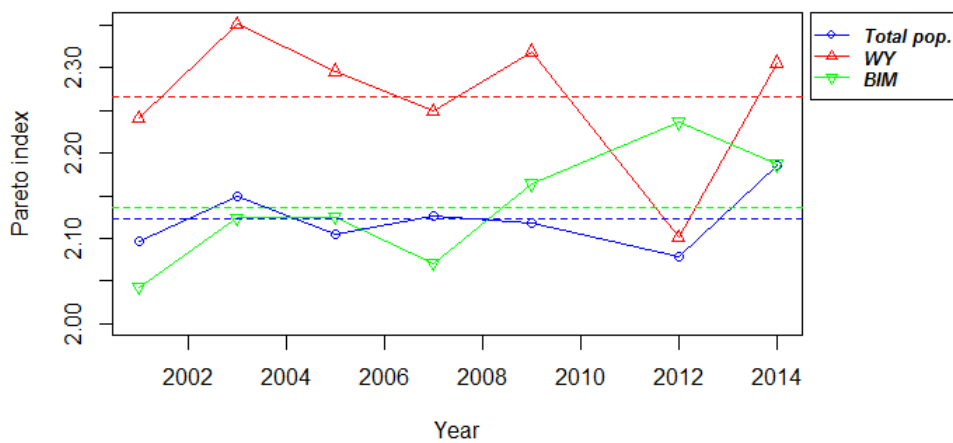
Na Figura 3.4.3 é indicado que a desigualdade de renda é mais expressiva quando a população é dividida por cor e também que a diferença percentual é maior para o valor de corte do que para a renda média. Embora seu comportamento similar, a renda média e o valor de corte não possuem correlação forte. Focando a dicotomia de gênero (linhas azuis), não há mudança substancial em direção a uma melhor igualdade de renda para os parâmetros estudados, se levarmos em conta os erros padrões. Em outras palavras, apesar da redução ao longo do tempo do coeficiente de Gini, o fato de que as mulheres brasileiras ganharem em média 40% menos que os homens permaneceu inalterado nos anos investigados. Quando o povo brasileiro é separado por cor / raça (linhas vermelhas), as diferenças percentuais para renda média e valor de corte em relação à população total são maiores. Brancos e amarelos ganham cerca de 65% a 50% mais que pretos, pardos e indígenas. No entanto, para o subgrupo PPI, podemos notar uma redução de 10% de sua diferença percentual em comparação com a população total tanto na renda média como no valor de corte nos anos selecionados.

Outra limitação do coeficiente de Gini é que ele não contém informações sobre rendimentos absolutos nacionais ou pessoais[41]. Assim, mesmo com um melhor coeficiente de Gini para subconjuntos historicamente desfavorecidos, através da Figura acima, podemos ver que esses subgrupos desfavorecidos da população brasileira têm uma renda mensal muito menor do que os outros. Por esta razão, é possível observar que a estrutura social de homens que ganham mais que mulheres e BAs ganharem mais do que PPIs permaneceu ao longo da série temporal.

Figura 3.4.4 – Série Temporal Índice de Pareto



(a)



(b)

Fonte: Próprio autor.

(a) Índice de Pareto para a população total dividida por gênero. (b) Índice de Pareto para a população total dividida por cor. As linhas horizontais representam as médias dos coeficientes de Pareto nos quais cada subconjunto da população flutua.

A partir do expoente de Pareto são investigados os valores assintóticos da distribuição de renda, compreendendo cerca dos 5% da população que possuem as maiores rendas mensais. A inserção da Figura 3.4.4 mostra que, para quase todos os subgrupos da população total, o índice de Pareto flutua em torno das suas respectivas médias. Essa tendência é indicada pelas linhas tracejadas horizontais. Para o grupo dividido por sexo e para as cores BA o coeficiente de inclinação tem magnitude  $10^{-3}$ . Doutro modo, um pequeno aumento é observado no grupo PPI, onde o coeficiente de

inclinação é  $10^{-2}$ .

Assim, para o grupo de gênero aplicando-se um teste T não pareado com correção de Welch, a média do coeficiente de Pareto ao longo dos anos selecionados é significativamente maior para as mulheres do que para os homens. Isso representa uma melhor distribuição de renda para as mulheres mais ricas. Para a separação por cor, a situação foi invertida: o índice de Pareto para BA é estatisticamente maior.



## 4 Modelo Econômico com Concentração de Riqueza (MECR)

Muitos trabalhos na área indicam que a distribuição de probabilidade da renda (riqueza) deve ser dividida por classes. Alguns desses estudos ainda tentam combinar a distribuição de Pareto que pode descrever regiões assintóticas de alta renda com uma ou mais funções de distribuição de probabilidade para o resto da população (por exemplo, distribuição exponencial, Gompertz, Normal ou Galton). Esta análise empírica nos últimos anos obteve grande sucesso quando aplicada em diferentes tipos de dados extraídos diretamente de pesquisas domiciliares, pesquisas de imposto de renda ou outras fontes similares. Do mesmo modo, a modelagem desse fenômeno tem se concentrado, principalmente, em dinâmicas estocásticas onde o comportamento de cauda pesada está associado a região assintótica da distribuição de equilíbrio do sistema, ou ainda, em processos de troca com leis universais [34, 22, 21, 35, 32, 24].

Assim, este capítulo será dedicado à exploração de um novo modelo para a distribuição de riqueza. Esse baseia-se no princípio da concentração de riqueza e, envolve agentes interagentes segregados em duas classes distintas conforme os seus respectivos processos de geração de riqueza. Decorrente dessa separação de processos, duas dinâmicas distintas são derivadas para uma mesma população e uma distribuição de equilíbrio definida por partes é alcançada.

### 4.1 Fundamentação do modelo

Apesar de algumas críticas relacionadas às simplificações presentes em vários modelos de econofísica, onde a diversidade de bens é traduzida em uma única variável chamada riqueza, dinheiro ou renda [28]. Nesse modelo também será utilizado a variável  $W$  (riqueza/renda/dinheiro), mas no sentido de que essa representa uma quantidade simbólica envolvida nas transações de todos os tipos de mercadorias. Assim, a diversidade e a preferência dos agentes por certos tipos de bens será simplisticamente relacionada ao número de processos (meios pelos quais o agente pode adquirir seus bens preferidos) que cada grupo de indivíduos pode participar. Para o MECR, o número de pessoas que compõem uma população  $N$  a qualquer momento é dividido em duas classes, de acordo com a riqueza e a diversidade de processos de produção de riqueza que lhes é permitido. Para ser mais preciso, dizemos que um agente interativo pertence à classe rica (classe alta) no tempo  $t$  quando o mecanismo de disparo é ativado ( $X_i(n) = 1$ ), e, conseqüentemente, sua dinâmica é baseada em certos processos de produção de riqueza. Inversamente, agentes para os quais o mecanismo

de disparo não foi ativado são interpretados como pertencentes a outra classe (classe baixa). Ainda, através dos resultados obtidos nas simulações (apresentados na seção 4.2), assumimos que para um conjunto específico de parâmetros do modelo existe um corte de riqueza  $W_c(t)$  entre estas duas classes. Isto é, nessas condições, podemos simplificar a frase anterior e dizer que um agente interativo pertence à classe alta quando sua riqueza total é maior do que uma certa riqueza de corte  $W_c(t)$ . Portanto, para um tempo fixo,

$$\frac{N_l(t)}{N} + \frac{N_u(t)}{N} = 1 \quad (4.1.1)$$

onde  $N_l(t)$  e  $N_u(t)$  são o número total de agentes na classe baixa e na classe alta, respectivamente.

Uma visão geral de alguns estudos empíricos sobre a distribuição de riqueza indicam que 3-10% de uma dada população pertencem à classe rica seguindo a distribuição de Pareto [30, 42, 43, 11, 18].

Então, para derivar a distribuição de probabilidade relativa à riqueza pessoal  $P(W)$ , toma-se o tempo discreto em intervalos iguais  $\tau > 0$ , ou seja,  $t = n\tau$ ,  $n \in \mathbb{N}$ . Com isso em mente, a riqueza pessoal dependente do tempo de  $i$  é denotada por  $W_i(n)$  e a riqueza média é denotada por:  $\overline{W(n)} = (W_1(n) + W_2(n) + \dots + W_N(n))/N$ . Assim, é fácil mostrar que  $\overline{W(n)}$  pode ser escrita como a soma ponderada de  $\overline{W_u(n)}$  (riqueza média da classe alta) e  $\overline{W_l(n)}$  (riqueza média da classe baixa):

$$\overline{W(n)} = \frac{N_u(n)}{N} \overline{W_u(n)} + \frac{N_l(n)}{N} \overline{W_l(n)}. \quad (4.1.2)$$

Para expressar a concentração de riqueza, como primeiro princípio do modelo, assumimos duas hipóteses : (a) O número de agentes na classe baixa  $N_l$  é próximo ao da população total,  $N \sim N_l$ ; (b) a média da riqueza da classe alta, no equilíbrio, e muito maior que a da classe complementar ( $\overline{W_u(n)} \gg \overline{W_l(n)}$ ) e, então, a última equação pode ser aproximada por,

$$\begin{aligned} \overline{W(n)} &= \overline{W_u(n)} \left( \frac{N_u(n)}{N} + \frac{N_l(n)}{N} \frac{\overline{W_l(n)}}{\overline{W_u(n)}} \right) \simeq \overline{W_u(n)} \left( \frac{N_u(n)}{N_l(n)} + \frac{\overline{W_l(n)}}{\overline{W_u(n)}} \right) \simeq \\ &\simeq \overline{W_u(n)} \left( \frac{N_u(n)}{N_l(n)} + \epsilon \right) \simeq \frac{N_u(n)}{N_l(n)} \overline{W_u(n)} = \omega(n) \overline{W_u(n)} \end{aligned}$$

onde negligenciamos as flutuações temporais no parâmetro  $\epsilon = \frac{\overline{W_l(n)}}{\overline{W_u(n)}}$ . Portanto, definimos a densidade relativa  $\omega$ :

$$\omega(n) \equiv \frac{\overline{W(n)}}{\overline{W_u(n)}} \simeq \left( \frac{N_u(n)}{N_l(n)} + \frac{\overline{W_l(n)}}{\overline{W_u(n)}} \right) \simeq \frac{N_u(n)}{N_l(n)}. \quad (4.1.3)$$



Isso significa que a mudança na riqueza média por intervalo de tempo  $t$ :  $\overline{\Delta W} = \overline{W(n+1)} - \overline{W(n)}$ , em uma aproximação de primeira ordem, pode ser escrita como:

$$\overline{\Delta W} \approx \omega \overline{\Delta W_u} + O(2) \quad (4.1.4)$$

consequentemente, a variação média da riqueza de uma dada população está intimamente relacionada com os agentes mais ricos. Essa abordagem restringe o modelo a sociedades com concentração de riqueza. Do ponto de vista prático, como vários estudos apontam, muitos dos países respeitam essa condição [44, 45, 46].

A concentração de riqueza, como se tornará mais claro ao longo do texto, não está associada apenas à retenção maciça de capital e bens mas, também, à diversidade de processos de geração de riqueza que um determinado grupo de agentes pode acessar.

Cada agente  $i$  que compõe essa população é aleatoriamente organizado em uma rede uniforme e pode interagir (ou não) com outros agentes  $j$ 's. I.e., de certa forma, há pessoas que estão em posição de influenciar a riqueza dos outros. Portanto, assumimos que a evolução temporal da riqueza pessoal é descrita por:

$$W_i(n+1) = \begin{cases} W_i(n) - \frac{H}{\omega(n)} \text{sgn}(W_i(n)) \overline{W(n)} + \xi_i(n) \overline{W(n)} + \frac{W_i(n)}{W_u(n)} \overline{\Delta W_u}, & \text{if } X_i(n) = 0 \\ W_i(n) + \lambda_i(n) W_i(n) + \sum_{j=1}^{N_u(n)} G_{ij} \frac{W_i(n) W_j(n)}{W_u(n)} & \text{if } X_i(n) = 1 \end{cases} \quad (4.1.5)$$

por simplicidade tomamos  $\tau = 1$ .

O modelo descreve a evolução da riqueza individual a partir de diferentes fontes:

- Renda Social: O termo  $\frac{H}{\omega(n)} \text{sgn}(W_i(n)) \overline{W(n)}$  com  $H > 0$ , um parâmetro livre e,  $\text{sgn}(W_i(n))$  sendo a função sinal. Esse é um termo de controle, que representa dois mecanismos: quando o agente tem uma riqueza positiva, ele representa a quantidade de recursos que o agente gasta para se manter como um membro de uma dada sociedade, incluindo impostos, aluguel, empréstimos, etc. No entanto, observe que quando o agente atinge um deficit em sua riqueza, este termo se torna positivo devido à função sinal. Isso representa subsídios ou assistência financeira que o indivíduo recebe em situações de pobreza acentuada, como a renegociação de suas dívidas e programas sociais.
- Renda Individual: A expressão  $\xi_i(n) \overline{W(n)}$  envolve  $\xi$ 's como um processo estocástico de ruído branco independente e normalmente distribuído  $\mathcal{N}(0, \sqrt{\frac{\sigma}{\omega}})$ , com as propriedades usuais. Além disso, o desvio-padrão  $\sqrt{\frac{\sigma}{\omega}}$  da gaussiana possui o parâmetro livre  $\sigma$  e depende, também, da densidade relativa  $\omega$ . Assim, o termo representa um valor financeiro ganho (ou perdido) pelo indivíduo através de

seu esforço produtivo e negociações igualmente provável de bens usuais (em pequenas quantidades) no tempo  $t$ . Isso inclui salários e ordenados, participação nos lucros, trabalho autônomo, comissões e bônus, entre outros.

- Efeito de Transbordamento (Trickle-Down Effect): Relembrando a equação 4.1.5,  $\frac{W_i(n)}{W_u(n)} \overline{\Delta W_u}$ , onde avaliamos ,

$$\begin{aligned} \overline{\Delta W_u} &= \frac{1}{N_u(n)} \sum_{i=1}^{N_u(n)} \Delta W_u = \frac{1}{N_u(n)} \sum_{i=1}^{N_u(n)} W_i(n+1) - W_i(n) \\ \overline{\Delta W_u} &= \frac{1}{N_u(n)} \sum_{i=1}^{N_u(n)} (\lambda_i(n) W_i(n) + \sum_{j=1}^{N_u(n)} G_{ij} \frac{W_i(n) W_j(n)}{W_u(n)}) \end{aligned} \quad (4.1.6)$$

para  $N_u(n)$  agentes em um dado passo  $n$ . Este termo encapsula o impacto direto da classe alta na evolução econômica da classe baixa. Ele pode ser interpretado como todos os fatores pelos quais um aumento na riqueza dos mais ricos pode influenciar os mais pobres. Como por exemplo, um aumento de investimentos e consumo pela classe mais rica, resultando em mais empregos e salários mais altos; um aumento da receita governamental e a criação de programas de auxílio financeiro. Por essa razão, o sistema é atualizado em dois passos: primeiro os mais ricos e depois os mais pobres, uma vez que a renda dos mais ricos precisa ser definida a priori. Tal fenômeno implica que a variação da riqueza da classe baixa é condicionada pela classe rica.

- Renda Individual da Classe Alta: o processo multiplicativo  $\lambda_i(n) W_i(n)$ , onde os  $\lambda$ 's são extraídos de uma função de probabilidade gaussiana  $\mathcal{N}(m, \sqrt{\omega})$ . Aqui temos algumas diferenças com relação ao termo de renda individual associado à classe baixa. A primeira, refere-se à média do processo gaussiano, que neste caso ainda que muito próxima, é diferente de zero ( $m \neq 0$ ). E a outra, refere-se ao termo estar multiplicado pela renda do indivíduo e não pela renda média. Essas modificações tentam refletir que, as atividades recorrentes mais lucrativas são realizadas pelos indivíduos mais ricos, ou seja, seus salários e ganhos são maiores. Além disso, em atividades geradoras de riqueza mais alta, como investimentos e empreendedorismo, em geral, elas são mais arriscadas (multiplicação pela renda do indivíduo).
- Renda da Rede: O termo para a classe alta  $\sum_{j=1}^{N_u(n)} G_{ij} \frac{W_i(n) W_j(n)}{W_u(n)}$  representa a transferência de riqueza com pesos estatísticos  $G_{ij}$  entre um único agente  $i$  e outros membros dentro de uma rede da classe alta. No contexto da geração de riqueza, este componente da dinâmica busca representar de maneira muito simplificada todas as fontes de riqueza que dependem de outros agentes pertencentes à classe mais rica (grandes quantias de dinheiro). Além disso, ilustra que as interações

financeiras entre os agentes  $i$  e  $j$  são amplificadas à medida que suas riquezas individuais excedem a média da classe alta,  $\overline{W_u(n)}$ , resultando em uma contribuição proporcionalmente maior para a transferência de riqueza geral dentro deste estrato socioeconômico. Isso sugere um sistema onde os mais ricos não apenas têm maior influência econômica, mas também se beneficiam de maneira mais marcante das interações financeiras, reforçando a dinâmica de acumulação de riqueza entre os já ricos.

Outro ponto importante na atualização da dinâmica é que as interações são mutuamente benéficas, ou seja, um ganho em  $W_i$  não requer necessariamente uma perda imediata pelos outros  $W_j$ 's durante interações binárias[25]. Em termos econômicos, uma negociação pode ser mutuamente benéfica mesmo quando os mercados não são perfeitamente eficientes[28].

Os mecanismos de geração de riqueza da classe baixa podem gerar uma situação de riqueza negativa ou endividamento, principalmente devido aos termos de renda social e renda individual. O termo de renda social pode tanto reduzir a riqueza de um indivíduo através de custos como impostos e aluguel quanto fornecer assistência financeira em casos de riqueza negativa. Por outro lado, o termo de renda individual reflete a variabilidade dos ganhos e perdas financeiras individuais, que podem levar a uma situação de endividamento quando os resultados das negociações ou esforços produtivos são desfavoráveis. Juntos, esses termos destacam como fatores sociais e individuais interagem para influenciar a possibilidade de riqueza negativa ou endividamento entre a classe baixa.

A variável  $X_i(n) \in [0, 1]$  é um valor lógico Booleano. A condição  $X_i(n) = 1$  indica que o agente  $i$  inicia uma dinâmica de evolução temporal alternativa envolvendo um termo de troca  $G_{ij}$  com o agente  $j$  e outros mecanismos para variar sua riqueza, como um processo multiplicativo. Nesse sentido, estabelecemos que o principal fator de troca é restrito aos agentes da classe alta.

O mecanismo habitual de reset para potenciais em modelos de neurônios introduz uma divergência crítica do modelo GGL tradicional e o presente modelo[7]. Enquanto no domínio neural, refletindo o comportamento de neurônios do tipo "integra-e-dispara", o potencial  $W$  reseta para zero (ou para um valor basal  $W_R$ ) após um pico, tal reset é impraticável no paradigma econômico. Como observado, neste modelo, a função de disparo está relacionada à separação entre as classes econômicas e suas respectivas dinâmicas de evolução.

Para  $X_i(n) = 1$  está associada uma função de probabilidade  $\Phi(W_i(n))$  onde o potencial da rede é a própria riqueza pessoal. Então, sob essas premissas, a atividade

da rede pode ser medida pela densidade de interações  $\rho_{(n)}(X_j(n) = 1)$ :

$$\rho_{(n)} = \frac{1}{N_u(n)} \sum_{j=1}^{N_u(n)} X_j(n) \quad (4.1.7)$$

onde a densidade de interações por passo de tempo é obtida pela densidade de probabilidade de unidades similares com potenciais  $W(n)$  em um intervalo arbitrário  $dW$  no tempo  $t$ , ou seja,  $P_u(W(n))dW(n)$ .

$$\rho_{(n)} = \int \Phi(W(n))P_u(W(n))dW(n) \quad (4.1.8)$$

onde a variável  $X_i(n) \in [0, 1]$  é booleana.

Por exemplo, pode-se escolher a função de disparo como uma função degrau em  $W_c(n)$  e definir esse valor como limite inferior para a riqueza da população mais rica (~90%), onde neste caso,  $N_u/N$  é constante e,  $W_c(n) \equiv p_{90}(W_i(n))$  representa o percentil 90 dos valores de riqueza. Nessa situação a função de disparo é estritamente determinística. Em um cenário mais realista, também, pode-se definir a função de disparo como uma sigmóide e permitir que a densidade de agentes na classe alta varie ao longo do tempo.

Ao considerarmos que a distribuição dos pesos  $G_{ij}$  tem uma variância finita e média  $G/N_u$  usamos a aproximação de campo médio que consiste na seguinte estimativa:

$$\begin{aligned} W_i(n) + \lambda_i(n)W_i(n) + \sum_{j=1}^{N_u} G_{ij} \frac{W_i(n)W_j(n)}{\overline{W_u(n)}} &= \\ &= W_i(n) + \lambda_i(n)W_i(n) + GW_i(n) \end{aligned} \quad (4.1.9)$$

Então, aplicando os resultados prévios na equação anterior, pode-se reescrever 4.1.5 em termos da sua aproximação de campo médio:

$$W_i(n+1) = \begin{cases} W_i(n) - \frac{H}{\omega(n)} \text{sgn}(W_i(n)) \overline{W(n)} + \xi_i(n) \overline{W(n)} + \frac{W_i(n)}{\overline{W_u(n)}} \overline{\Delta W_u}, & \text{se } X_i(n) = 0 \\ W_i(n) + \lambda_i(n)W_i(n) + GW_i(n) & \text{se } X_i(n) = 1 \end{cases} \quad (4.1.10)$$

A eq. 4.1.10 representa a dinâmica de campo médio do modelo proposto. Dessa separação, a possibilidade de se estudar riquezas negativas também é alcançada, visto que outros modelos estocásticos como a própria equação de Lotka-Volterra generalizada possuem, intrinsecamente, valores de riqueza mínimo estritamente positivos. O "efeito de riqueza negativa", nessa situação, significa que alguns dos agentes mais pobres gastam mais do que recebem, ou ainda, que possuem uma dívida que excede seu total de ativos. Apesar de recorrente, ainda existem poucas pesquisas domiciliares e modelos econofísicos que levam em conta esse fenômeno. Alguns relatórios indicam que 14% dos domicílios nos EUA têm riqueza estritamente negativa[47].

Do mesmo modo, observa-se que não há estado estacionário para eq. 4.1.10,  $\overline{W}_u(t) = \overline{W}_u(0)e^{(m+\omega)t}$ . Em outras palavras, permitimos a produção de riqueza em cada passo de tempo. Nesse sentido, também não há conservação de riqueza e o sistema não pode ser interpretado como um regime termodinâmico. Conseqüentemente, devemos reescrever a última equação em termos da riqueza relativa:

$$V_i(n) = \frac{W_i(n)}{\overline{W}_u(n)} \quad (4.1.11)$$

assim, usando a regra da cadeia:  $\Delta V_i \cong \frac{1}{\overline{W}_u(n)} \Delta W_i - \frac{W_i(n)}{\overline{W}_u(n)^2} \Delta \overline{W}_u$ , caracterizamos a variação relativa da riqueza em um passo  $n$  como:  $\Delta V_i = V_{i(n+1)} - V_i(n)$ , para valores grandes de  $N$ . Então se  $X_i(n) = 0$ :

$$\begin{aligned} \Delta V_i &= \frac{1}{\overline{W}_u(n)} \left[ -\frac{H}{\omega(n)} \text{sign}(W_i(n)) \overline{W}(n) + \xi_i(n) \overline{W}(n) + \frac{W_i(n)}{\overline{W}_u(n)} \Delta \overline{W}_u \right] - \frac{W_i(n)}{\overline{W}_u(n)^2} \Delta \overline{W}_u = \\ &= \omega(n) \xi_i(n) - \text{sgn}(V_i(n)) H \end{aligned}$$

onde usamos a equação 4.1.3 para simplificar a expressão. Para  $X_i(n) = 1$ :

$$\begin{aligned} \Delta V_i &= \frac{1}{\overline{W}_u(n)} [\lambda_i(n) W_i(n) + G W_i(n)] - \frac{W_i(n)}{\overline{W}_u(n)^2} \left[ \frac{1}{N_u(n)} \sum_{i=1}^{N_u(n)} (\lambda_i(n) W_i(n) + G W_i(n)) \right] = \\ &= \frac{1}{\overline{W}_u(n)} [\lambda_i(n) W_i(n) + G W_i(n)] - \frac{W_i(n)}{\overline{W}_u(n)^2} [m \overline{W}_u(n) + G \overline{W}_u(n)] = \\ &= \frac{1}{\overline{W}_u(n)} \lambda_i(n) W_i(n) - \frac{W_i(n)}{\overline{W}_u(n)} m = V_i(n) \lambda_i(n) - m V_i(n) \end{aligned}$$

Finalmente,

$$\Delta V_i = \begin{cases} \omega(n) \xi_i(n) - \text{sgn}(V_i(n)) H, & \text{if } X_i(n) = 0 \\ V_i(n) \lambda_i(n) - m V_i(n) & \text{if } X_i(n) = 1 \end{cases} \quad (4.1.12)$$

Portanto, ao analisar 4.1.12, derivada da aproximação de campo médio, vemos que a dinâmica da variável normalizada  $V$  não depende do termo de troca  $G$ .

Nas próximas três subseções, trabalharemos em detalhes com cada um dos dois grupos de riqueza formados (classe alta e classe baixa) para calcular resultados analíticos.

### 4.1.1 Classe baixa

Para agentes com  $X_i(n) = 0$ , a partir da equação 4.1.12:  $\Delta V_i = \omega(n)\xi_i(n) - \text{sgn}(V_i(n))H$ , onde  $\omega(n)$  ajusta a escala de impacto do ruído no sistema, podemos observar que eles não interagem diretamente na rede da classe alta devido ao termo principal de atividade comercial desaparecer. Além disso, vale ressaltar que a evolução da riqueza relativa  $V$  não é condicionada pela dinâmica da classe alta, então temos um conjunto de  $N_l$  equações estocásticas idênticas e desacopladas com dois parâmetros livres:  $H, \sigma$ . Desse modo, no contínuo, temos:

$$dV = -\text{sgn}(V)Hdt + \sqrt{\sigma\omega(t)}d\xi. \quad (4.1.13)$$

Para esse subgrupo de agentes  $i = 1, \dots, N_l$ , usando a formulação de Fokker-Planck, podemos descrever a evolução temporal da distribuição de probabilidade da riqueza relativa da classe inferior  $P_l(V, t)$ :

$$\frac{\partial P_l(V, t)}{\partial t} = \text{sgn}(V)H \frac{\partial}{\partial V} [P_l(V, t)] + \frac{\sigma}{2} \frac{\partial^2}{\partial V^2} [\omega(t)P(V, t)]. \quad (4.1.14)$$

Portanto, para a solução estacionária  $\partial P_l(W)/\partial t = 0$ , considerando que no equilíbrio  $\lim_{t \rightarrow \infty} \omega(t) = \frac{\bar{W}}{\bar{W}_u}$ , temos a seguinte solução,

$$P_l(V) = Ce^{-\frac{2H}{\sigma\omega}|V|} \quad V < V_c \quad (4.1.15)$$

Para a região em que  $V < V_c$ , a distribuição de riqueza pode ser descrita por uma distribuição semelhante à Laplaciana, aplicável aos agentes dentro da classe econômica mais baixa. Aqui,  $C$  atua como um fator de normalização. Essa expressão acomoda valores negativos de riqueza e exibe decaimento exponencial em ambos os lados a partir do centro.

Ainda, se considerarmos valores positivos ( $W > 0$ ) e voltarmos à variável  $W$ , obtemos como caso particular :  $P(W, H = \frac{1}{2}c, \sigma = 1c) = Ce^{-\frac{W}{\bar{W}}}, c = \text{constante}$ . Formulação essa que é a mesma utilizada nos estudos de Yakovenko e outras análises empíricas que ajustam com precisão uma variedade de distribuições observadas de riqueza (e renda)[13, 15, 9, 17, 32, 18].

### 4.1.2 Classe alta

Para a classe mais rica da sociedade, trabalhamos exclusivamente com agentes  $i = 1, \dots, N_u$  definidos quando o mecanismo de disparo é ativado,  $X_i(n) = 1$ , em um passo arbitrário  $n$ .

Para isso, observamos que a equação  $\Delta V_i = V_i(n)\lambda_i(n) - mV_i(n)$  corresponde a um processo estocástico multiplicativo geral  $V_i(n+1) - V_i(n) = A(V(n))\lambda_i(n) +$

$B(V(n))$ , onde, por meio de uma mudança conveniente de variáveis, o ruído multiplicativo é transformado em ruído aditivo:

$$dv = \frac{1}{A(V(t))} dV \quad (4.1.16)$$

portanto, usando a regra de diferenciação de segunda ordem,  $\Delta v \simeq \frac{dv}{dV} \Delta V + \frac{1}{2} \frac{d^2 v}{dV^2} \Delta V^2$ , obtemos:

$$\Delta v \simeq \frac{1}{A} \Delta V + \frac{1}{2} \frac{d}{dV} \left( \frac{1}{A} \right) (A\lambda + B)^2$$

em seguida, aproximamos o termo quadrático fazendo a média sobre o termo de ruído,  $E(\lambda^2) = \text{var}(\lambda) + [E(\lambda)]^2 = \omega + m^2$ , onde.

$$\Delta v \simeq \frac{B}{A} - \frac{1}{2} \frac{dA}{dV} \left( \frac{B^2}{A^2} + \frac{2Bm}{A} + \omega + m^2 \right)$$

Agora, consideramos um processo onde  $\omega \gg B^2/A^2$ , que para o caso em questão, equivale a  $\omega \gg m^2$  para  $m < 1$ . Assim, derivamos uma equação de Langevin discreta:

$$\Delta v = \lambda + J(v) \quad (4.1.17)$$

onde a “drift force”  $J(v)$  é dada por,

$$J(v) \simeq \frac{B}{A} - \frac{\omega}{2} \frac{dA}{dV} \quad (4.1.18)$$

Isso leva a uma distribuição de probabilidade assintótica da seguinte forma[48, 49, 50, 25]:

$$P_u(V) = e^{\frac{2}{\omega} \int^v J(v') dv'} = \frac{1}{A^2(v')} e^{\frac{2}{\omega} \int^v \frac{B(v')}{A^2(v')} dv'}. \quad (4.1.19)$$

No caso particular do MECR, temos:  $A = V$  e  $B = -mV$ . Portanto, a distribuição de probabilidade estacionária da riqueza relativa  $V$  levando em consideração a dinâmica 4.1.12 é:

$$P_u(V) = C' V^{-1-\alpha}, \quad V \geq V_c \quad (4.1.20)$$

com,

$$\alpha = 1 + \frac{2m}{\omega} \quad (4.1.21)$$

onde  $C'$  é uma constante determinada pelos parâmetros do sistema.

A distribuição de probabilidade de riqueza relativa da classe alta está alinhada com a descoberta de Vilfredo Pareto, seguindo uma lei de potência ( $\sim V^{-1-\alpha}$ ) pura[51, 10], com um índice de Pareto  $\alpha$  maior que um.

Ainda, é importante observar que, mesmo diante de condições altamente instáveis que causam dinâmicas globais não estacionárias, a equação 4.1.20 sugere a persistência de leis de potência estáveis com expoentes constantes ao longo do tempo. Esse expoente  $\alpha$ , diferentemente do visto em outras modelagens do mesmo fenômeno[22, 26, 24], tem uma dependência direta com a densidade relativa de agentes do sistema  $\omega$ , tal como, há na distribuição de probabilidade da classe baixa. Esse efeito, como será melhor explorado na seção de resultados, tem o papel de ajustar as funções de densidade de probabilidade ao número de agentes de cada classe. Uma vez que essa seleção é feita pela variável  $X_i(n) \in [0, 1]$  associada, unicamente, a função de disparo e seu potencial.

### 4.1.3 Condições de contorno do MECR

Finalmente, para completar as soluções, devemos encontrar os valores de  $C$  e  $C'$ . Para isso, vamos utilizar as condições de normalidade e continuidade:

$$\int_{-\infty}^{\infty} P(V) dV = 1 \quad (4.1.22)$$

$$P_l(V_c) = P_u(V_c) \quad (4.1.23)$$

de 4.1.23 temos

$$C' = C e^{-\frac{2H}{\sigma\omega}|V_c|} V_c^{1+\alpha} \quad (4.1.24)$$

e de 4.1.22 isolamos  $C$

$$C = \left( \int_{-\infty}^{V_c} e^{-\frac{2H}{\sigma\omega}|V|} dV + e^{-\frac{2H}{\sigma\omega}|V_c|} V_c^{1+\alpha} \int_{V_c}^{\infty} V^{-1-\alpha} dV \right)^{-1} =$$

$$C = \frac{\alpha}{\alpha K_l + e^{-\frac{2H}{\sigma\omega}|V_c|} V_c} \quad (4.1.25)$$

Ainda, há uma singularidade em zero em  $V^{-1-\alpha}$ , logo, o valor de corte  $V_c$  precisa ser positivo. E por esse motivo definimos,

$$K_l = \frac{\sigma\omega}{2H} (2 - e^{-\frac{2H}{\sigma\omega}V_c}) \quad (4.1.26)$$

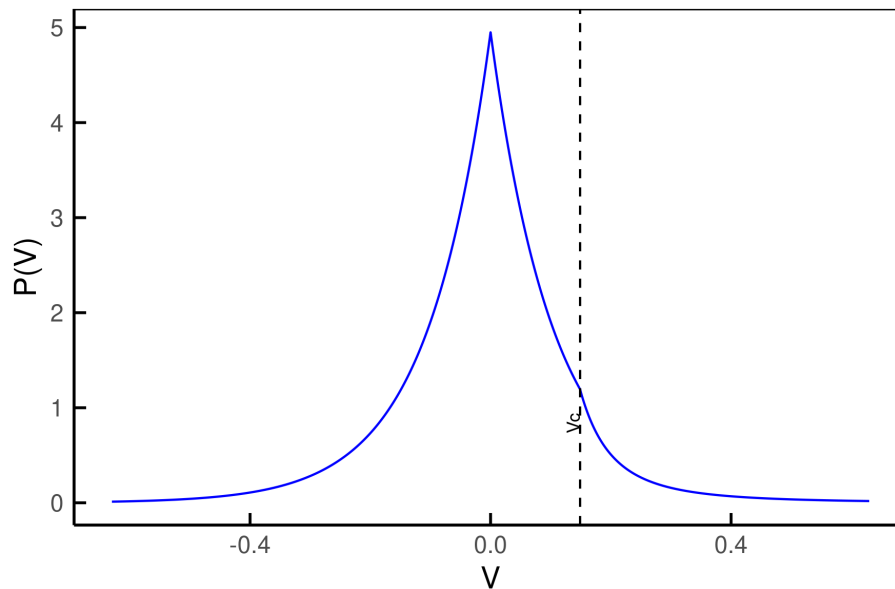
e por consequência,

$$P(V) = \begin{cases} C e^{-\frac{2H}{\sigma\omega}|V|} & \text{se } V < V_c \\ C' V^{-1-\alpha} & \text{se } V \geq V_c \end{cases} \quad (4.1.27)$$



onde essa ultima expressão representa a função densidade de probabilidade da riqueza relativa  $V$ . Quanto a sua forma, em uma comparação simplista, essa seria uma distribuição que combina as conhecidas: distribuição de Laplace  $P(x|\mu, b) = \frac{1}{2b}e^{-\frac{|x-\mu|}{b}}$ , muito usada para descrever fenômenos que têm uma tendência central, com  $\mu$  sendo o parâmetro de dispersão e  $b > 0$  o parâmetro de localização; e a conhecida distribuição de Pareto Tipo I,  $P(x|x_{min}, \alpha) = \frac{\alpha x_{min}^\alpha}{x^{\alpha+1}}$ . Dessa associação, podemos considerar 4.1.27 uma distribuição do tipo lapacliana-pareto-1, onde seu gráfico é visto na figura abaixo. Nessa mesma linha, note que para rendas negativas seu decaimento é exponencial  $\sim e^{\frac{2H}{\sigma\omega}}$  e para valores positivos existe um ponto de corte  $V_c$ , onde há uma mudança na taxa de decaimento (não diferenciável em  $V_c$ ).

Figura 4.1.1 – MECR densidade de probabilidade



**Fonte:** Próprio autor.

Figura 4.1.1: Função densidade de probabilidade do MECR  $P(V)$ . Onde a linha vertical representa o corte entre as distribuições  $P_l(V)$  e  $P_u(V)$ .

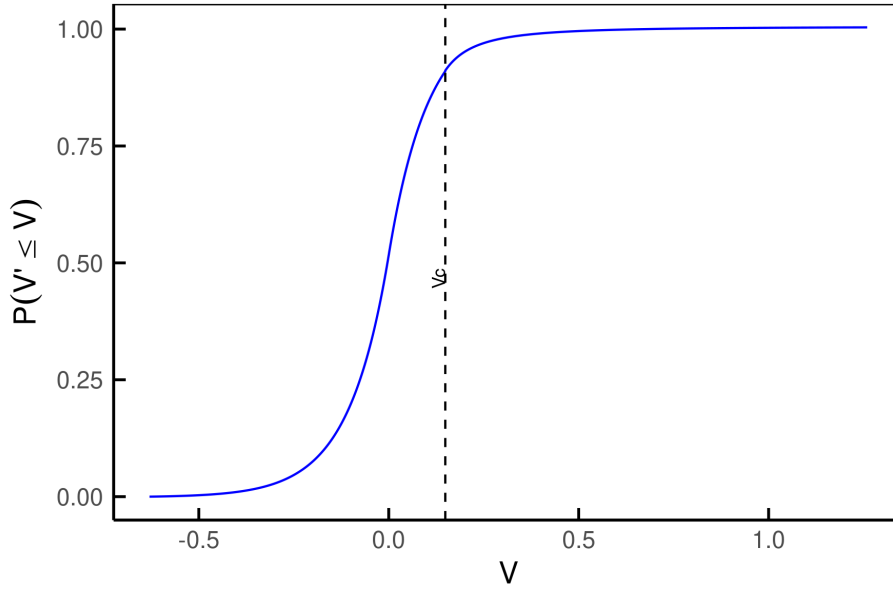
Do mesmo modo, podemos derivar sua distribuição acumulada de probabilidade:

$$P(V \leq V') \equiv \int_{-\infty}^{V'} P(V) dV. \quad (4.1.28)$$

Para seu calculo, temos :

$$P(V' \leq V) = \begin{cases} C \frac{\sigma\omega}{2H} (1 + e^{-\frac{2HV_c}{\sigma\omega}} - e^{-\frac{2HV}{\sigma\omega}}) & se \quad V < V_c \\ C \frac{\sigma\omega}{2H} + \frac{V_c^{-\alpha} - V^{-\alpha}}{\alpha} C' & se \quad V \geq V_c \end{cases} \quad (4.1.29)$$

Figura 4.1.2 – MECR densidade de probabilidade acumulada



**Fonte:** Próprio autor.

Figura 4.1.2: Função densidade de probabilidade acumulada do MECR  $P(V' \leq V)$ . Onde a linha vertical representa o corte entre as distribuições  $P_l(V)$  e  $P_u(V)$ .

Portanto, a partir de 4.1.29 de imediato obtemos, também, a distribuição de probabilidade acumulada complementar, uma vez que,  $P(V' \geq V) \equiv \int_V^\infty P(V)dV = 1 - P(V' \leq V)$ .

Dado que a função de disparo determina a proporção  $\frac{N_u}{N_l} = \omega$ , a densidade de probabilidade 4.1.29 será ajustada para tal configuração, mesmo que os parâmetros livres do modelo,  $H, \sigma, m$  permaneçam inalterados, pois, como já mencionado, ambas as distribuições são dependentes de  $\omega$ . Assim, naturalmente, para que 4.1.22 e 4.1.23 permaneçam válidas o valor de corte  $V_c$ , que determina a quebra entre as distribuições, também será alterado,  $V_c = f(\omega, \sigma, H, m)$ . Logo, de 4.1.3:

$$\omega = \frac{N_u}{N_l} \simeq \frac{\int_{V_c}^{\infty} P_u(V)dV}{\int_{-\infty}^{V_c} P_l(V)dV} = \frac{C' \frac{V_c^{-\alpha}}{\alpha}}{C \frac{\sigma\omega}{2H} (2 - e^{-\frac{2HV_c}{\sigma\omega}})} = \frac{C e^{-\frac{2H}{\sigma\omega} V_c} V_c^{1+\alpha} \frac{V_c^{-\alpha}}{\alpha}}{C \frac{\sigma\omega}{2H} (2 - e^{-\frac{2HV_c}{\sigma\omega}})} = \frac{e^{-\frac{2H}{\sigma\omega} V_c} V_c}{\frac{\alpha\sigma\omega}{2H} (2 - e^{-\frac{2HV_c}{\sigma\omega}})}$$

$$\omega = \frac{V_c}{\frac{\alpha\sigma\omega}{2H} (2e^{\frac{2HV_c}{\sigma\omega}} - 1)} \quad (4.1.30)$$

na equação transcendental acima, estamos interessados em valores de  $V_c$  próximos de zero, por esse motivo, expandimos a expressão anterior até segunda ordem:

$\frac{V_c}{\frac{\alpha\sigma\omega}{2H} (2e^{\frac{2HV_c}{\sigma\omega}} - 1)} \sim -\frac{2}{\alpha} \left(\frac{2H}{\sigma\omega}\right)^2 V_c^2 + \frac{1}{\alpha} \left(\frac{2H}{\sigma\omega}\right) V_c + O(3)$ , e assim obtemos:

$$-\frac{2}{\alpha} \left(\frac{2H}{\sigma\omega}\right)^2 V_c^2 + \frac{1}{\alpha} \left(\frac{2H}{\sigma\omega}\right) V_c = \omega \quad (4.1.31)$$

um polinômio de segunda ordem com concavidade para baixo tendo suas raízes dada por :

$$V_c = \frac{H\sigma\omega \pm \sigma\omega\sqrt{H^2 - 8\alpha\omega H^2}}{8H^2}. \quad (4.1.32)$$

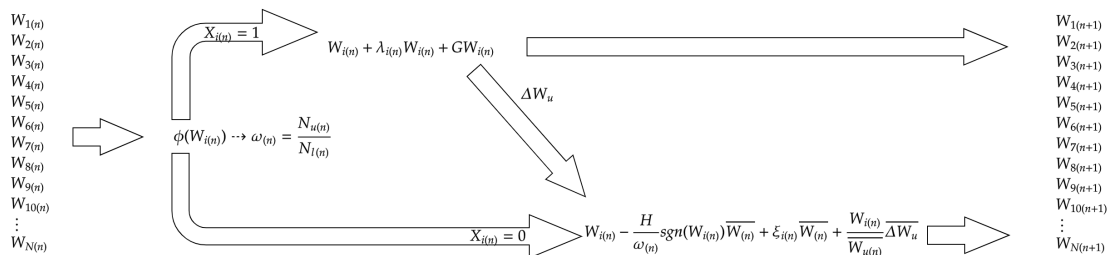
A expressão anterior, dentro das condições do modelo ( $N_l \gg N_u, \overline{W}_u \gg \overline{W}_l$  e  $V_C \sim 0$ ), fornece-nos as valores possíveis de  $V_c$  para que 4.1.3 seja válida.

## 4.2 MECR simulações e resultados

Nesta seção, apresentaremos os resultados obtidos a partir de simulações do modelo. Todos as simulações computacionais aqui apresentadas foram implementadas com a linguagem de programação R[52], fixando-se a semente aleatória 2024.

Para a tarefa consideramos uma população inicial composta por  $N$  agentes, que, por opção, assume uma distribuição uniforme com riqueza  $W$  entre  $-10^3$  e  $10^3$ , sem perda de generalidade. O procedimento de atualização de um passo  $n + 1$  para  $W$  seguirá a dinâmica de campo médio dada pela equação 4.1.10 e ocorrerá em três etapas: primeiro, será estimado a função de ativação  $\Phi(W_i(n))$  e, por consequência, as populações  $N_l(n)$  e  $N_u(n)$  serão definidas; depois a classe alta, e finalmente a classe baixa. Ainda, a simulação será replicada 100 vezes para que as medições sejam feitas com um intervalo de confiança de 95%.

Figura 4.2.1 – Diagrama de um passo de simulação



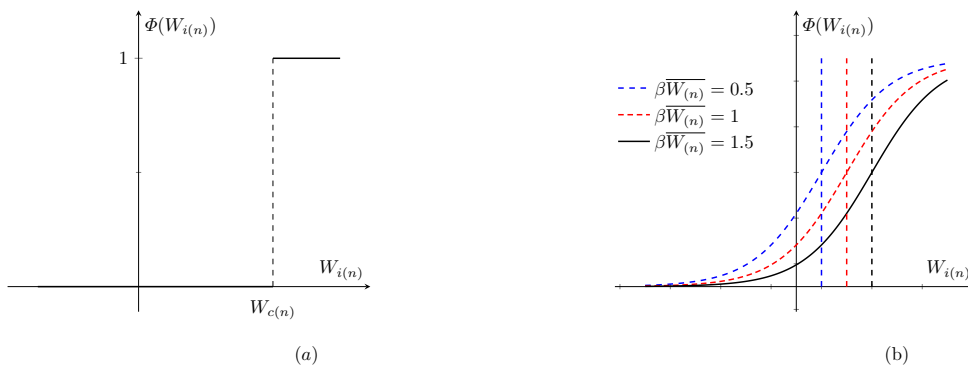
Fonte: Próprio autor.

Figura 4.2.1: Diagrama ilustrando as três etapas de atualização de um passo de simulação MECR.

Além disso, assumimos que a simulação alcançou o equilíbrio ( $\frac{dP(V,t)}{dt} = 0$ ) caracterizado pela estabilidade das propriedades de interesse, a partir de dois critérios: (1) ao atingir um limiar para a correlação entre a população inicial e aquela no passo  $n$  inferior a 0.05; (2) que suas médias móveis para as propriedades de interesse ao longo

do tempo convirjam para valores constantes, i.e.,  $\frac{dN_u}{dt} = 0$  e  $\frac{dV}{dt} = 0$ . Especificamente para as médias, determinamos que o sistema alcançou um estado de equilíbrio quando as variações entre os valores finais da média móvel — calculada utilizando uma janela de 100 passos — ficam abaixo de um limiar estabelecido de 5 do valor médio nesse intervalo. Ao final desse processo, as repetições do último passo serão combinadas para formar uma única população média, considerando todas as repetições. Antes de calcular a renda média individual dos agentes, eles serão ordenados com base em sua riqueza. Por fim, quanto a  $\Phi(W_i(n))$ , investigaremos também dois tipos de funções de ativação: uma sigmóide e outra do tipo função degrau.

Figura 4.2.2 – Funções de disparo

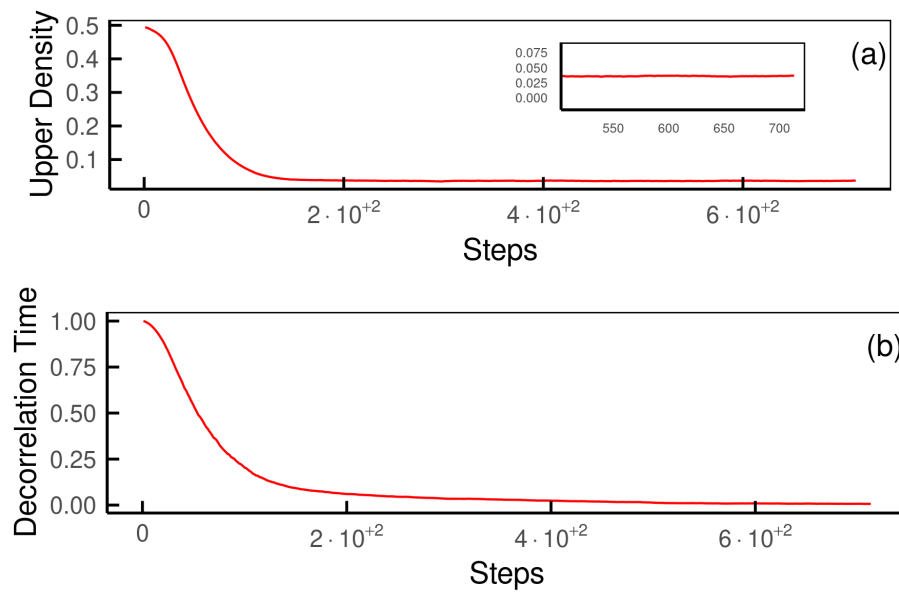


Fonte: Próprio autor.

Figura 4.2.2: (a) Função de disparo do tipo degrau  $\Phi(W_i(n)) = \Theta(W_i(n) - W_{c(n)})$ , onde  $W_{c(n)} \equiv p_X(W_i(n))$  é o percentil  $X$  dos valores de riqueza. (b) Função de disparo do tipo sigmóide  $\Phi(W_i(n)) = \frac{1}{1 + e^{-(W_i(n) - \beta \bar{W}(n))}}$ , onde  $\beta > 0$  é uma constante multiplicativa. Observe que, se  $\beta = 1$ , temos uma sigmóide padrão deslocada para a riqueza média do sistema.

A correlação entre a população inicial e a do passo  $n + 1$  foi calculada a partir do coeficiente de Pearson assim definido:  $Cor(X, Y) = \frac{Cov(X, Y)}{\sqrt{var(X)var(Y)}}$ . Onde  $Cov(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$  é a covariância entre as duas variáveis e  $var(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$  a variância de cada variável. Portanto, quanto mais próximo de zero for essa grandeza significa que as duas populações não dependem linearmente uma da outra.

Figura 4.2.3 – Densidade dos ricos e tempo de decorrelação

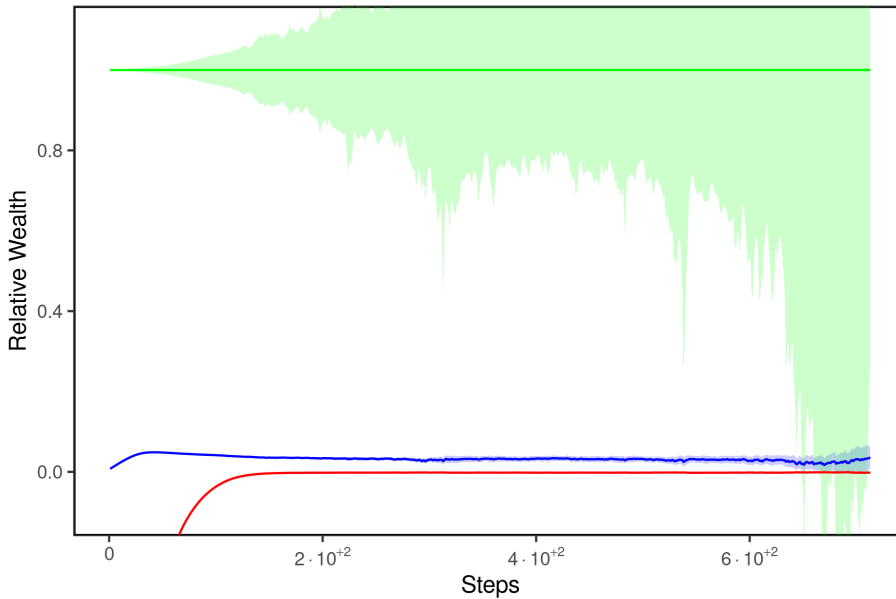


Fonte: Próprio autor.

Figura 4.2.3: (a) densidade de ricos ao longo dos passos de simulação, no equilíbrio,  $\frac{N_u}{N} = 0.037 \pm 0.002$ . Para os parâmetros utilizados temos:  $N = 5000$ ,  $H = 0.00354$ ,  $G = 0.00039$ ,  $\sigma = 0.006255$ ,  $m = 0.0182$ . Função de disparo utilizada do tipo sigmóide com  $\beta = 3$ . (b) tempo de decorrelação para a simulação  $T = 713$  passos de simulação .

Da figura 4.2.3 (a) observamos a convergência da densidade de ricos  $\frac{N_u}{N}$  após o equilíbrio quando a função de ativação utilizada é do tipo sigmóide. Nessa simulação, obtivemos  $\bar{W} = 2.46 \times 10^9$ ,  $\bar{W}_u = 6.91 \times 10^9$  e  $\bar{W}_l = -1.50 \times 10^7$ . Onde  $\frac{\bar{W}}{\bar{W}_u} = 0.036$ , assume valor próximo ao encontrado para  $\frac{N_u}{N} = 0.039$  e dentro do intervalo de confiança, confirmando a equação 4.1.2 e os limites do modelo. Como já mencionado na construção teórica, ainda que  $W$  não possua equilíbrio, com suas médias assumindo valores que tendem ao infinito, quando analisamos a evolução temporal da riqueza relativa  $V$ , um estado estacionário é encontrado ( $\frac{dV}{dt} = 0$ ), como pode ser visto na figura 4.2.4.

Figura 4.2.4 – Evolução temporal da riqueza relativa média



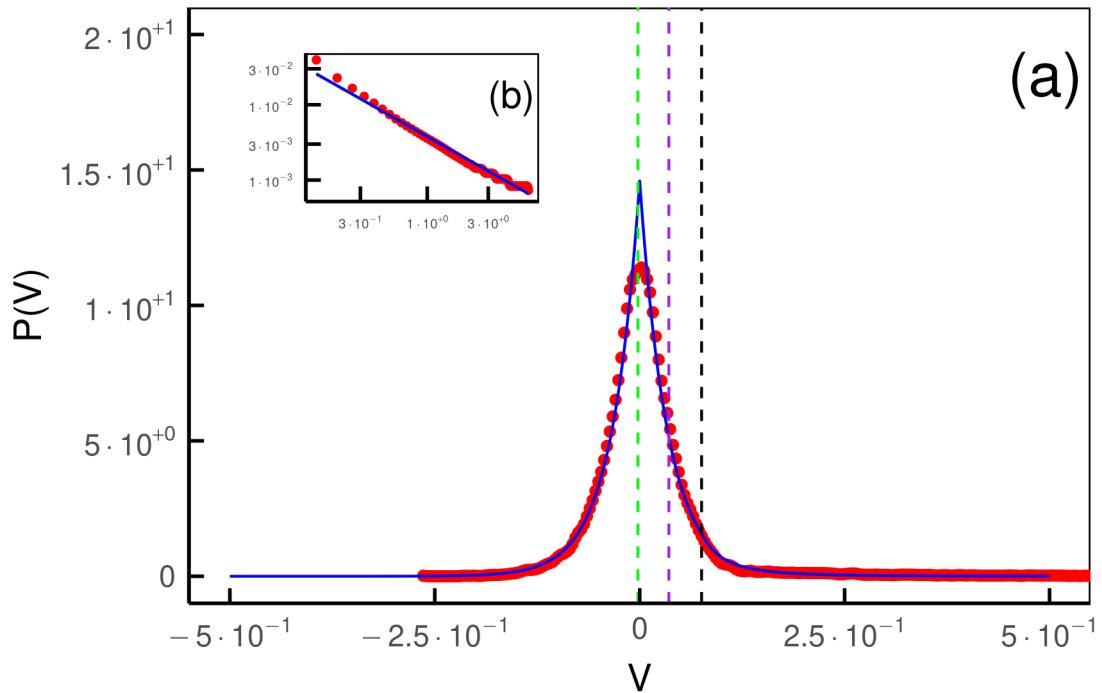
**Fonte:** Próprio autor.

Figura 4.2.4: Evolução temporal da riqueza relativa média  $\bar{V}$ . Onde a linha azul representa  $\bar{V}_{(n)}$ , a linha vermelha  $\bar{V}_{l(n)}$  e a linha verde  $\bar{V}_{u(n)}$ .

Para os parâmetros utilizados temos:  $N = 5000$ ,  $H = 0.00354$ ,  $G = 0.00039$ ,  $\sigma = 0.006255$ ,  $m = 0.0182$ . Função de disparo utilizada do tipo sigmóide com  $\beta = 3$ .

A comparação de  $P(V)$  teórico e computacional demonstrada na figura 4.2.5 sugere uma boa aderência entre ambas a partir de uma binagem adequada, mesmo quando os parâmetros são dispostos para que a densidade de ricos seja muito pequena. Nesse ponto, obtivemos um valor do teste KS(B.0.1) para as amostras de  $D = 0.09$ , através da comparação das suas FDAs. Nos dois casos, modelo teórico e simulação, uma lei de potência acentuada é observada. Para a simulação é estimado a partir da linearização da transformação  $\log - \log$  dos dados um índice de Pareto de  $\alpha = 1.96$ , valor muito próximo a predição teórica de  $\alpha = 1 + \frac{2m}{\omega} = 1.94$ . Ainda, como esperado da teoria, os dados da simulação estão exponencialmente distribuídos ao redor de um pico em zero, onde estão concentradas as riquezas de boa parte da população total.

Figura 4.2.5 – Densidade de probabilidade MECR+sigmóide vs. simulação

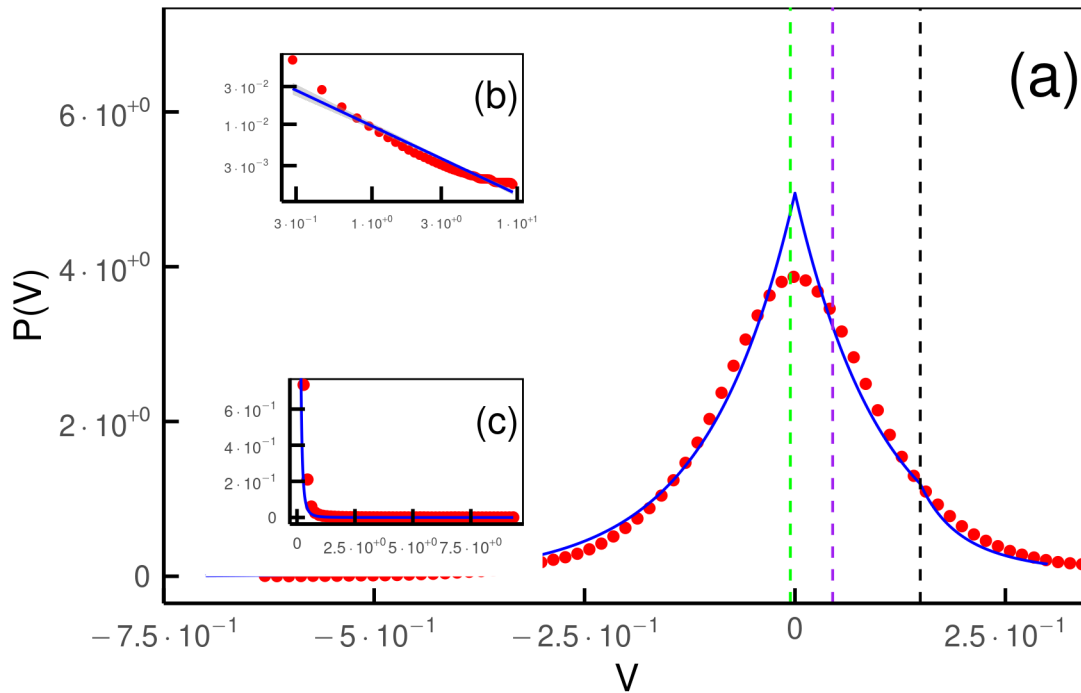


**Fonte:** Próprio autor.

Figura 4.2.5: Densidade de probabilidade MECR versus simulação computacional. Para os parâmetros utilizados temos:  $N = 5000$ ,  $H = 0.00354$ ,  $G = 0.00039$ ,  $\sigma = 0.006255$ ,  $m = 0.0182$ . Função de disparo utilizada do tipo sigmóide com  $\beta = 3$ . O tempo de equilíbrio foi de  $T = 713$  passos de simulação. (a) No gráfico, linhas tracejadas verticais: riqueza relativa média da classe baixa  $\bar{V}_l$  (verde) e riqueza relativa média  $\bar{V}$  (roxo); pontos vermelhos são os pontos da simulação; curva azul, modelo teórico MECR. (b) Linearização em escala log-log da FDAC dos agentes da classe alta para cálculo do coeficiente de Pareto  $\alpha = 1.96$ .

Um ponto importante a ser enfatizado é que não há no sistema memória de quem são os agentes da classe alta/baixa nos passos anteriores. Esse processo de divisão em classes está associado somente a função de disparo  $\Phi(W_i(n))$  e a própria riqueza. Assim, toda a dinâmica da riqueza estará condicionada a esse mecanismo (processo não-Markoviano). Isto posto, podemos utilizar o mesmo conjunto de parâmetros para o modelo e obter resultados distintos, ao variarmos a função de disparo. Para ilustrar, alterando  $\beta$  para 2, alcançamos uma densidade de ricos de  $\sim 7\%$ , utilizando os mesmos parâmetros de 4.2.5. Desse fato, alguns pontos do desenvolvimento teórico emergem, a densidade relativa  $\omega = \frac{N_u(n)}{N_l(n)} + \frac{\bar{W}_l(n)}{\bar{W}_u(n)} \simeq \frac{N_u}{N_l} \simeq \frac{\bar{W}}{\bar{W}_u}$  presente tanto na definição do índice de Pareto 4.1.21 de  $P_u(V)$  como na taxa de decaimento da exponencial de  $P_l(V)$ , tem o papel de adequar a distribuição às modificações de escala do sistema, sobretudo, advindas da função de disparo. Na mesma linha, não serão todo conjunto de parâmetros àqueles acessíveis ao modelo, na medida que: (1) a renda de corte precisa ser positiva,  $V_c > 0$ ; (2) a separação entre as classes deve estar presente,  $\min(V_u) \gtrsim \max(V_l)$ ; (3) a relação 4.1.3 deve ser respeitada.

Figura 4.2.6 – Densidade de probabilidade MECR+degrau vs. simulação

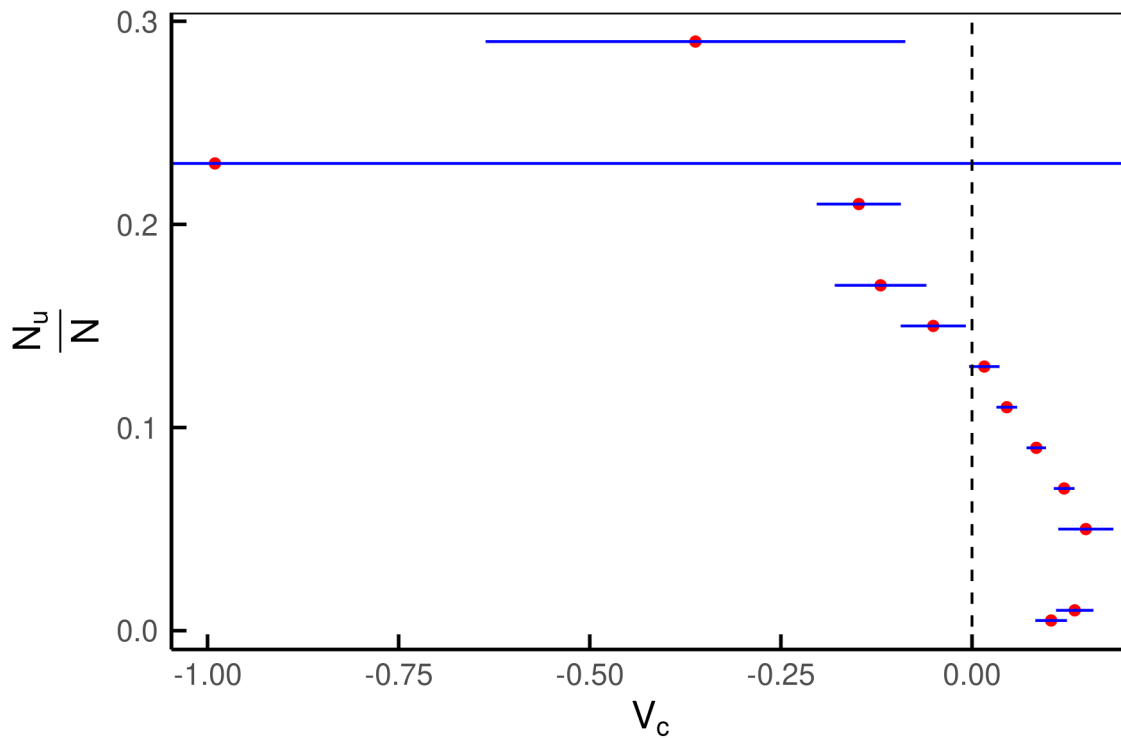


Fonte: Próprio autor.

Figura 4.2.6: Densidade de probabilidade MECR versus simulação computacional. Para os parâmetros utilizados temos:  $N = 5000$ ,  $H = 0.0215$ ,  $G = 0.000920$ ,  $\sigma = 0.0855$ ,  $m = 0.0234$ . Função de disparo utilizada do tipo degrau com  $W_c = p_{95}(W)$ . O tempo de equilíbrio foi de  $T = 331$  passos de simulação. (a) No gráfico, linhas tracejadas verticais: riqueza relativa média da classe baixa  $\bar{V}_l$  (verde); riqueza relativa média  $\bar{V}$  (roxo); riqueza de corte  $V_C$  (preto); riqueza relativa média da classe alta  $\bar{V}_u$  (amarelo). Os pontos vermelhos são os pontos da simulação; curva azul, modelo teórico MECR. (b) Linearização em escala log-log da FDAC dos agentes da classe alta para cálculo do coeficiente de Pareto  $\alpha = 1.86$ . (c) Ajuste ampliado para classe alta.

No gráfico 4.2.6, feito com a função de disparo do tipo degrau com o valor de corte fixado no percentil 0.95. Observa-se que ao alterarmos os parâmetros para que a densidade de ricos seja maior, conseguimos explicitar o comportamento de lei de potencia da distribuição conjunta. Assim a função densidade de probabilidade deixa de ser uma laplaciana "pura" e se aproxima muito do exposto na figura 4.1.1. A concordância com os valores da simulação ainda é muito boa, em especial, para a parte da classe alta. O coeficiente de Pareto teórico é de 1.89 frente ao da simulação de  $\alpha = 1.86$ . Já para a relação entre as médias temos  $\frac{\bar{W}}{\bar{W}_u} = 0.045$ , valor próximo ao esperado de 0.052.

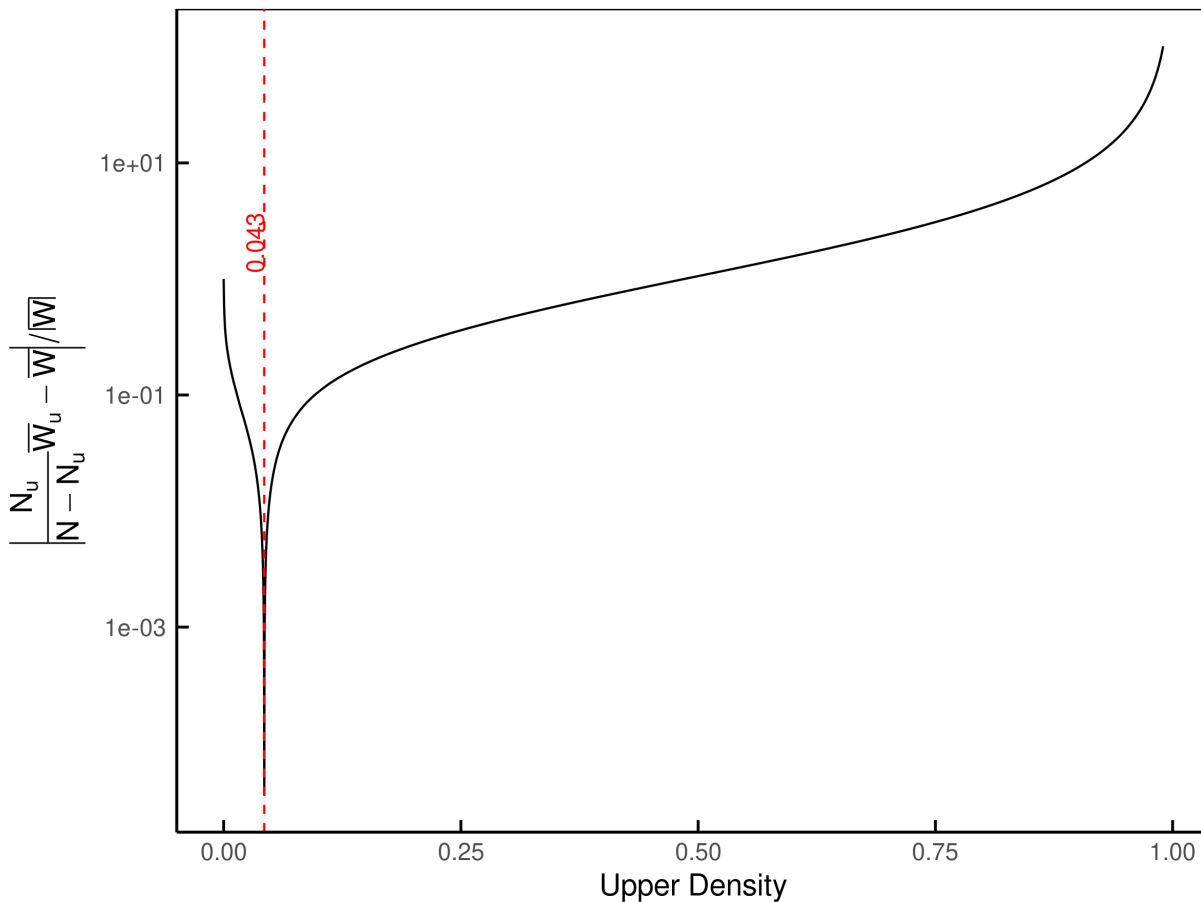


Figura 4.2.7 – Variação da proporção de ricos e Renda de Corte  $V_C$ 

**Fonte:** Próprio autor.

Figura 4.2.7: Variação da densidade de ricos a partir da função de disparo determinística. Para os parâmetros utilizados temos:  $N = 5000$ ,  $H = 0.0215$ ,  $G = 0.000920$ ,  $\sigma = 0.0855$ ,  $m = 0.0234$ . Função de disparo utilizada do tipo degrau com  $p_{99.5}(W)$ ,  $p_{99}(W)$ ,  $p_{95}(W)$ ,  $p_{93}(W)$ ,  $p_{91}(W)$ ,  $p_{89}(W)$ ,  $p_{87}(W)$ ,  $p_{85}(W)$ ,  $p_{83}(W)$ ,  $p_{79}(W)$ ,  $p_{77}(W)$ ,  $p_{71}(W)$ . As linhas horizontais azuis representam o intervalo de confiança de  $V_C$ . A partir de  $p_{85}(W)$ ,  $V_C$  passa a ser negativo quebrando a continuidade do modelo.

Ao utilizar a função determinística de ativação do tipo degrau, temos a facilidade de ajustar as proporções  $N_u$  e  $N_l$  de maneira mais controlada. Da figura 4.2.7 é possível perceber que ao variarmos a densidade de ricos sem alterar os demais parâmetros, ou seja, aumentar a proporção de ricos através da função de disparo, resulta em um deslocamento da densidade de probabilidade da classe alta progressivamente para a esquerda. Esse deslocamento ocorre até que as condições de contorno do modelo deixem de ser atendidas. Além disso, esse ajuste leva a uma diminuição contínua no coeficiente de Pareto, partindo de 1.89 quando  $W_c = p_{95}(W)$ , para 1.38 quando  $W_c = p_{91}(W)$  mas ainda respeitando as condições do modelo. E após, para 1.27 quando  $W_c = p_{85}(W)$ . Essa observação reforça a ideia de que existe um valor mínimo para o coeficiente de Pareto, que deve ser necessariamente maior do que um.

Figura 4.2.8 – Método para determinação de  $V_C$ 

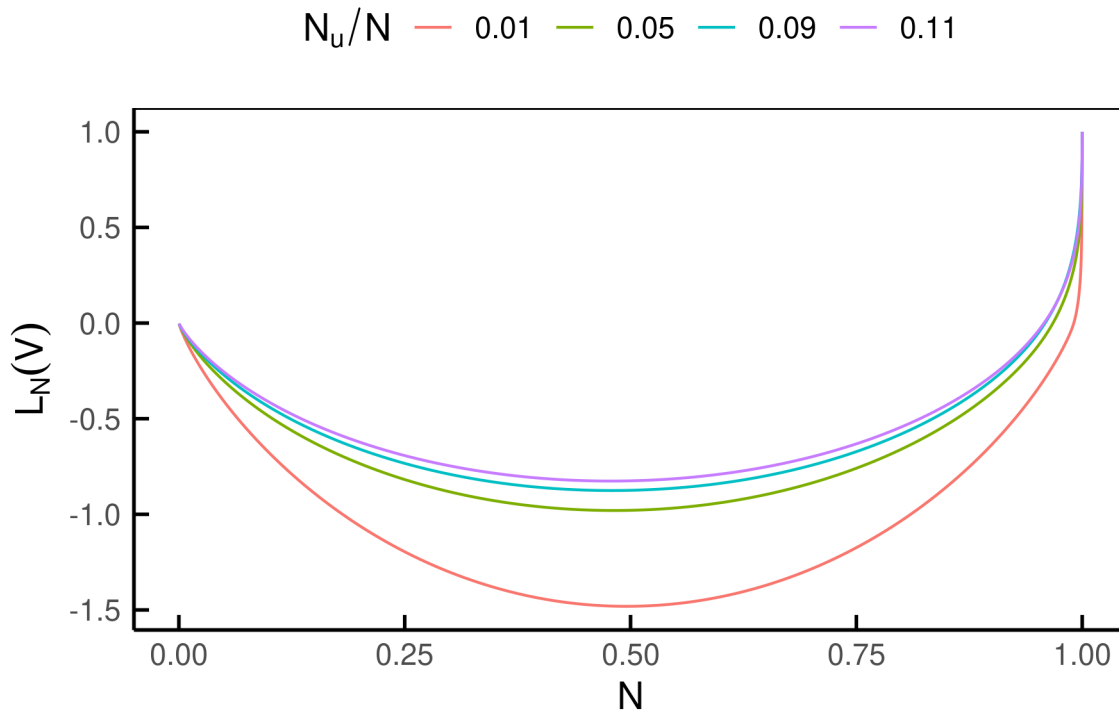
Fonte: Próprio autor.

Figura 4.2.8: Método computacional para obtenção de  $V_c$ . Linha tracejada vertical vermelha  $\frac{N_u}{N}$ . Para os parâmetros utilizados temos:  $N = 5000$ ,  $H = 0.0115$ ,  $G = 0.00033$ ,  $\sigma = 0.0655$ ,  $m = 0.0101$ . Função de disparo utilizada do tipo degrau com  $p_{95}(W)$ .

Um dos grandes desafios de se trabalhar com distribuições de renda/riqueza definidas por partes é o de determinar o valor de corte entre ambas. Muitos trabalhos propõem técnicas sofisticadas para tal tarefa, por exemplo, a minimização das diferenças entre dados empíricos e teóricos através de algoritmos de otimização, como o 'Particle Swarm Optimization'[18]. Nesse contexto, de 4.1.3 podemos propor um método computacional, iterativo e de grande simplicidade, para a estimação de  $V_c$  e  $\frac{N_u}{N}$  para um conjunto de dados arbitrários de riqueza. Para isso, basta analisarmos que, de acordo o MECR:  $\frac{N_u}{N-N_u} \overline{W}_u - \overline{W} \sim 0$ , ou ainda,  $\frac{N_u}{N-N_u} \overline{W}_u - \overline{W} < \epsilon$ . Onde devemos tornar  $\epsilon$  tão pequeno quanto possível. Assim, para uma dada população de  $N$  agentes, ordenando-os por riqueza e iterando o número de agentes da classe rica de 2 até  $N$  conseguimos determinar o menor valor de  $\epsilon$  possível e, por consequência, definir  $V_c = \min(W_u), \frac{N_u}{N}$  (vide apêndice C). Na figura 4.2.8 é possível observar os resultados da iteração, de onde obtivemos um erro inferior a 10% para a obtenção da densidade de ricos. Já para o valor de corte  $V_c$ , temos uma divergência um pouco maior, onde

na simulação obtivemos  $V_c^{sim} = 0.011$  e a partir do método  $V_c^{alg} = 0.050$ . Ainda, da equação teórica 4.2.7 obtemos para as raízes de  $V_C$  os valores  $[0.062, 0.013]$ , o que confere para esses resultados um intervalo de boa confiabilidade.

Figura 4.2.9 – Curva de Lorenz



**Fonte:** Próprio autor.

Figura 4.2.9: Curva de Lorenz  $L_N$  para diferentes densidades de ricos. Para os parâmetros utilizados temos:  $N = 5000$ ,  $H = 0.0215$ ,  $G = 0.000920$ ,  $\sigma = 0.0855$ ,  $m = 0.0234$ . Função de disparo utilizada do tipo degrau.

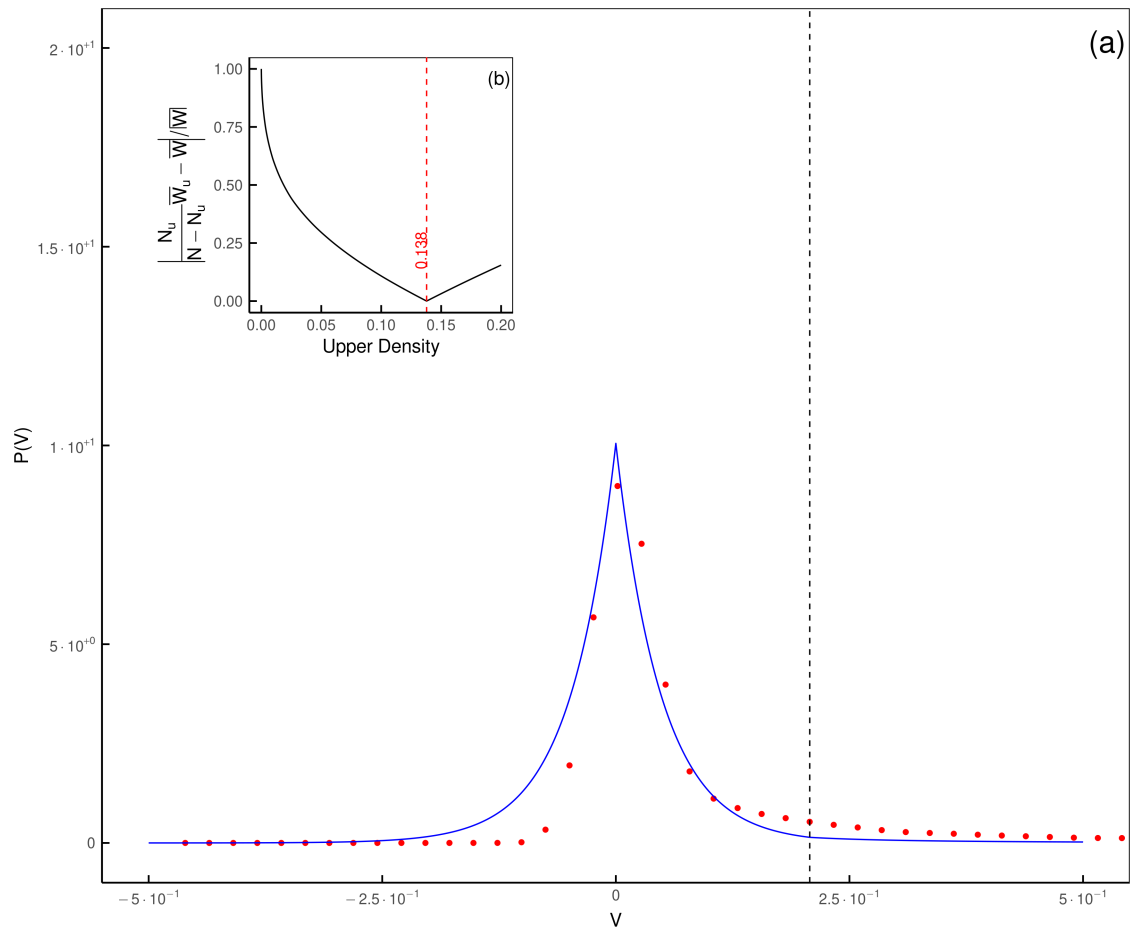
As variações decorrentes unicamente da função de disparo têm o potencial de modificar a distribuição de valores, conforme ilustrado na figura 4.2.9. A análise da curva de Lorenz revela que a diminuição na densidade de indivíduos mais ricos e por consequência, aumento no índice de Pareto ( $\alpha \sim \frac{1}{\omega}$ ) sem a alteração de parâmetros, implica em um aumento na desigualdade da população. Quando a proporção da população na classe alta é de 1%, a curva se afasta mais acentuadamente da linha de igualdade quase desde o início, indicando que a desigualdade é muito pronunciada em toda a distribuição da população. A desigualdade continua aumentando à medida que nos movemos para a direita do gráfico (ou seja, em direção a uma população maior). Portanto, sem modificar os mecanismos de geração de riqueza para um grupo específico de agentes e apenas aumentando a quantidade de ricos (aqui definidos como agentes que seguem uma distribuição de Pareto), estaríamos, de fato, diminuindo a disparidade de riqueza na população em questão.

Para exemplificar a aplicação do novo modelo, utilizamos uma versão condensada dos microdados do DINA (Distribuição Nacional de Contas), disponibilizada por Zucman, que permite sua divulgação pública [53]. Assim, as análises subsequentes estarão restritas a esse conjunto reduzido de dados, limitando sua abrangência. O DINA é uma metodologia inovadora que integra dados de distribuição de renda e riqueza com as contas nacionais tradicionais. Ela oferece uma visão mais completa da distribuição econômica ao combinar informações detalhadas sobre a distribuição da renda e da riqueza com os agregados macroeconômicos das contas nacionais dos Estados Unidos. O processamento desses microdados seguiu a metodologia desenvolvida por Piketty, Saez e Zucman[54], que foi posteriormente atualizada por Saez e Zucman em 2020[55]. A partir dessa abordagem conseguimos derivar a variável riqueza para nossas análises.

Com os dados tratados em mãos, estabelecemos a seguinte metodologia para análise e ajuste da curva teórica de riqueza aos dados empíricos:

1. Obtemos o valor de corte e a densidade de ricos dos dados empíricos tendo como base a abordagem desenvolvida neste texto (veja a figura 4.2.8);
2. Com a separação definida, aplicamos o algoritmo de otimização L-BFGS-B (vide Apêndice D) para estimar os melhores valores dos parâmetros( $H, \sigma, m$ ) a partir da minimização do erro quadrático médio entre a curva teórica e os dados do DINA.

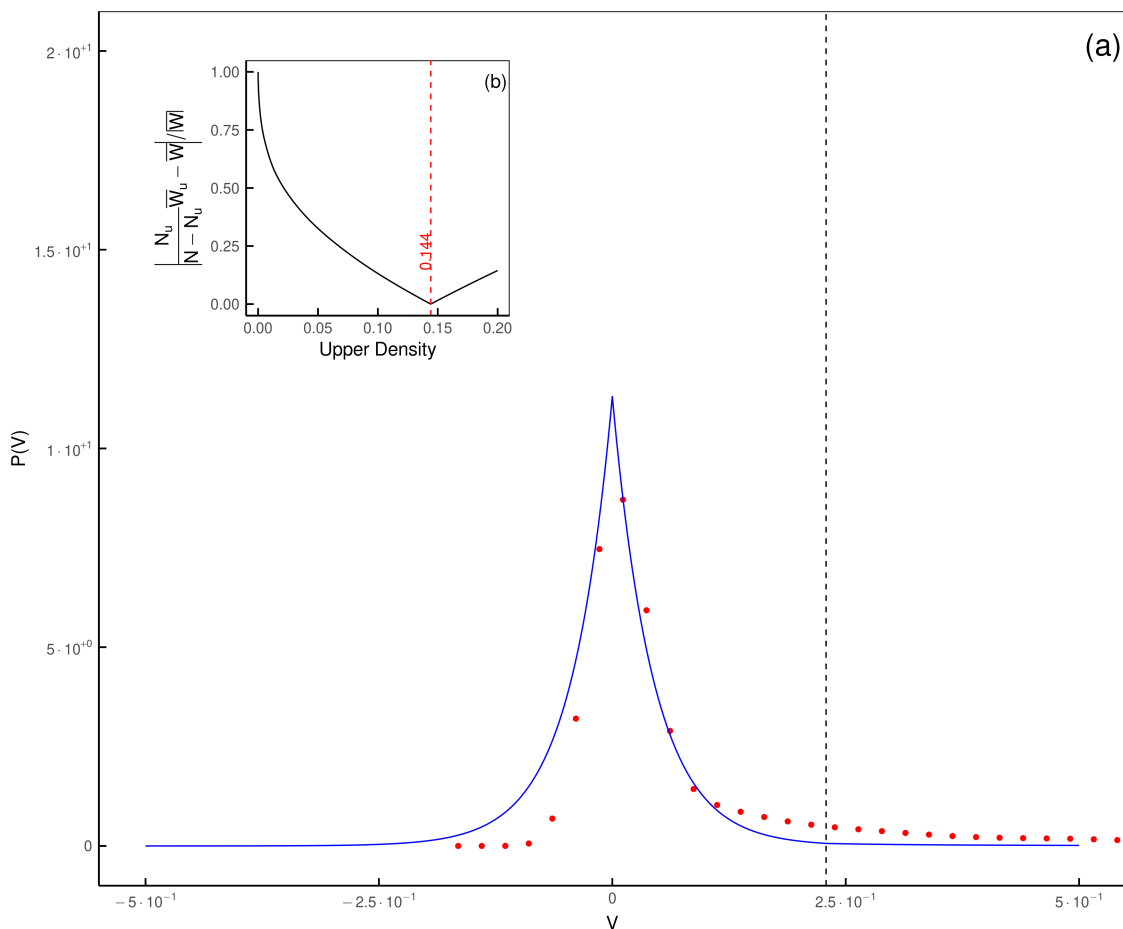
Figura 4.2.10 – DINA 2009



**Fonte:** Próprio autor.

Figura 4.2.10: (a) Dados DINA 2009 vs. curva teórica MECR. Os parâmetros encontrados a partir da otimização foram  $H = 0.040$ ,  $\sigma = 0.024$ ,  $m = 0.001$ . Erro quadrático médio de 0.19 e R-quadrado ajustado de 0.93. (b) Método para obtenção de  $N_u/N$  e  $V_C$ , de onde obtivemos:  $N_u/N = 0.138$  e  $V_C = 0.21$ .

Figura 4.2.11 – DINA 2019



Fonte: Próprio autor.

Figura 4.2.11:(a) Dados DINA 2019 vs. curva teórica MECR. Os parâmetros encontrados a partir da otimização foram  $H = 0.32$ ,  $\sigma = 0.17$ ,  $m = 0.001$ . Erro quadrático médio de 0.18 e R-quadrado ajustado de 0.93. (b) Método para obtenção de  $N_u/N$  e  $V_C$ , de onde obtivemos:  $N_u/N = 0.144$  e  $V_C = 0.23$ .

Nas figuras 4.2.10 e 4.2.11 temos o ajuste do modelo para os dados do DINA de 2009 e 2019. A partir do ajuste com base na minimização do erro quadrático médio conseguimos o valor do R-quadrado ajustado de 0.93 para ambos os anos, o que indica que, aproximadamente, 93% da variância da variável dependente pode ser explicada pelo modelo. Isso sugere que o modelo tem um excelente poder explicativo para os dados, considerando o número de preditores incluídos.

## **Parte II**

# **Mineração de Textos**





Nesta parte será apresentado uma visão geral sobre os conceitos basilares da análise de sentimentos em textos. Além disso, apresentaremos em detalhes o modelo para análise de textos e classificação de sentimentos desenvolvido. Vale ressaltar que esse algoritmo faz parte de um projeto vencedor da segunda edição do *Mining Lab-Nexa* (2018) na categoria de Inteligência Artificial, tendo como um dos propósitos inferir e prever comportamentos característicos do mercado financeiro através da análise de textos relacionados.

Sobre essa formulação, ressalta-se que todo o algoritmo computacional, como também, o dicionário base foram desenvolvidos para a língua inglesa. Isto posto, as categorias de classificação escolhidas foram: raiva (*anger*), aversão (*disgust*), medo (*fear*), analítico (*analytical*), surpresa (*surprise*), alegria (*joy*), tristeza (*sadness*) e confiança (*confident*). Essa atribuição foi adotada tendo como base outros classificadores de sentimentos como o IBM Watson, o Qemotion e o Sentiment. Além disso, a partir desse último classificador, uma biblioteca de inferência Bayesiana para a linguagem de programação R, devolvemos o dicionário base agregando outras palavras e seus sentimentos manualmente.

Nas subseções seguintes, também, será descrito a aprendizagem da IA (Inteligência Artificial), ou seja, do dicionário léxico e uma simples comparação entre: o modelo desenvolvido, o método Bayesiano e a classificação do próprio ser humano.



# 5 Análise de sentimento

## 5.1 Análise de sentimento: uma visão geral

Na era da internet a Mineração de Textos (*Text Mining*) é um campo de pesquisa que tem ganhado notoriedade como consequência da grande quantidade de dados produzidos, sobretudo, através das interações em redes sociais, pesquisas em saúde, notícias, transações bancárias, etc. Para se ter uma ideia, estima-se que até 2020 esse volume de dados atinja a marca de 40 zettabytes<sup>1</sup>. Ainda, sabe-se que boa parcela desses dados( em sua maioria não estruturados) são armazenados na forma de texto, devido a facilidade de processamento e otimização de armazenamento[56]. Isto posto, o objetivo da Mineração de Textos é o de processar dados não estruturadas (textuais) visando a obtenção de informações importantes através da busca de padrões e tendências.

Dentro desse escopo, uma técnica chamada de Análise Sentimento (AS) ou Mineração de Opinião (MO) se destaca devido a sua ampla aplicabilidade: classificação de opiniões de um modo geral, análise de satisfação, métricas de rede social, etc[4]. Sucintamente, a AS visa determinar a atitude de um falante ou escritor com respeito a algum assunto estabelecido ou a polaridade contextual geral ou reação emocional a um documento, interação ou evento. Essa atitude pode ser um julgamento, estado afetivo (isto é, o estado emocional do autor ou falante), ou a comunicação emocional pretendida (isto é, o efeito emocional pretendido pelo autor ou interlocutor)[57].

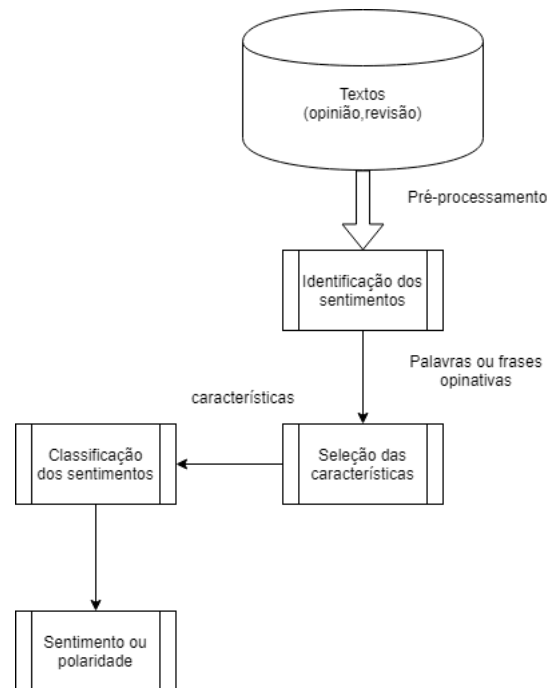
A AS, tecnicamente, pode ser considerada um processo de classificação em três níveis[5]:

1. Documento: Classifica a opinião expressa no documento quanto a sua polaridade (positiva ou negativa) ou sentimento;
2. Sentença: Classifica o sentimento expresso em cada sentença, ou seja, em partes menores do documento;
3. Aspecto: Visa classificar o sentimento em relação a aspectos específicos das entidades (palavras). Ou seja, classifica as diversas opiniões do sujeito da ação em relação a diferentes entidades (palavras).

---

<sup>1</sup> 1 ZB = 10<sup>21</sup> bytes

Figura 5.1.1 – Processo de Análise de Sentimento



**Fonte:** o próprio autor.

Figura 5.1.1: Adaptado de MEDHAT, W.; HASSAN, A.; KORASHY, H, 2014.

Na figura 5.1.1 está representado uma arquitetura geral para algoritmos dedicados a AS.

O pré-processamento textual é uma das chaves para a maioria dos algoritmos de Mineração de Textos. Nessa etapa são aplicadas algumas técnicas cruciais para o bom desempenho do classificador[56]:

1. Tokenização: Tokenização (*do inglês tokenization*) é o processo de particionar o texto em pedaços menores (palavras/frases) chamadas de *tokens*;
2. Filtragem: Refere-se a remoção de algumas palavras e caracteres do documento, tais como: pontuação, nomes de pessoas e locais, preposições, conjunções, etc;
3. Lematização: A lematização é o processo, efetivamente, de deflexionar uma palavra para determinar o seu lema (as flexões chamam-se *lexemas*);
4. *Stemming*: representa a tarefa de se tomar o radical da palavra.

Vale ressaltar que nem todos os algoritmos de AS, necessariamente, utilizam todas as técnicas de pré-processamento citadas. Uma vez que a partir delas são assumidas várias simplificações. Por exemplo, algoritmos que na etapa de pré-processamento utilizam a técnica de Lematização, automaticamente, desprezam a importância das flexões (gênero e grau) de uma mesma palavra.

Ainda na Figura 5.1.1, tipicamente, para a etapa de identificação de termos opinativos são utilizados dicionários léxicos de palavras e frases com suas respectivas

polaridades (ou sentimentos). Há vários métodos sofisticados para a identificação desses termos, onde muitos levam em consideração o contexto o qual esses estão inseridos[58, 59].

Uma vez selecionados os termos opinativos, como próximo passo temos a seleção das características. Essa é realizada, sobretudo, através da: (a) Contagem de frequência: contagem das palavras (individualmente ou um conjunto delas) e suas frequências de ocorrência; (b) ponderação binária, para cada palavra do texto atribui-se 1 se estiver contida no dicionário léxico e 0 para o outro caso. Também, é possível a inserção de pesos de frequência para identificar a importância do termo; (c) Partes do discurso: Busca por adjetivos como termos opinativos[5].

Em posse das características do texto (*bag of features*), já na etapa de classificação, são utilizados métodos estatísticos (*Machine Learning*) para a classificação e posterior aprendizagem (supervisionada, semi-supervisionada, não-supervisionada ou por reforço) do algoritmo, dentre os quais destacaremos as mais comuns: Naive Bayes (NB) e Máxima Entropia (ME) [60].

### 5.1.1 Métodos estatísticos de classificação

Para uma sucinta explanação sobre alguns dos principais métodos estatísticos de classificação aplicados a AS, denotaremos  $F = \{f_1, f_2, \dots, f_m\}$  como sendo o conjunto de  $m$  pré-definidas características que podem aparecer em um texto. Além disso, temos que  $n_i(e)$  representa o número de ocorrências da característica  $f_i$  no texto  $e$ . Assim, cada texto (documento) pode ser representado pelo vetor:  $\vec{e} = (n_1(e), n_2(e), \dots, n_m(e))$ [60].

#### 5.1.1.1 Naive Bayes

O classificador NB e suas variações devido a boa performance e precisão associados a simplicidade são, entre todos os probabilísticos, os mais populares[56, 61, 62].

Em geral, em classificadores do tipo NB atribui-se a classe  $c$  (polaridade, sentimento, entre outros) para um texto  $e$  a partir da função  $c^* = \operatorname{argmax}_c(P(c|e))$ , onde  $P(c|e)$  é dado pela teorema de Bayes:

$$P(c|e) = \frac{P(c)P(e|c)}{P(e)} \quad (5.1.1)$$

Na equação (5.1.1), por simplicidade, em grande parte das vezes, omite-se  $P(e)$  na valoração da probabilidade condicional. Assim, os termos restantes são estimados a partir da suposição de que as  $f_i$ 's características são condicionalmente independentes, i.e.,

$$P_{NB}(c|e) = \frac{P(c) \prod_{i=1}^m P(f_i|c)^{n_i(e)}}{P(e)}. \quad (5.1.2)$$

Apesar da simplicidade e da forte suposição de independência das probabilidades, o algoritmo de Naive Bayes continua sendo uma ótima opção, especialmente, como primeiro tratamento devido a sua fácil implementação.

### 5.1.1.2 Máxima Entropia

O conceito de entropia máxima originário do formalismo da Mecânica Estatística e de problemas que envolvem Teoria da Informação é uma técnica alternativa a NB e que, em algumas vezes, é mais eficaz na classificação e mineração de textos[63]. Nesse contexto, assumimos que a probabilidade condicional  $P(c|e)$  segue uma distribuição exponencial da forma:

$$P(c|e) = \frac{1}{Z(e)} \exp\left(\sum_i \lambda_{i,c} F_{i,c}(t, e)\right) \quad (5.1.3)$$

onde a função  $F_{i,c}$  responsável pela relação entre a característica  $f_i$  e a classe  $c_i$  é frequentemente definida como uma função degrau do tipo:

$$F(e, c') = \begin{cases} 1 & n_i(t) > 0 \quad e, c' = c \\ 0 & n_i(t) \leq 0 \end{cases} \quad (5.1.4)$$

Importante salientar que ao contrário do método de Naive Bayes, a Máxima Entropia não faz suposições sobre as relações entre as características e, portanto, pode obter melhor precisão na medida que os pressupostos de independência condicional não precisam ser cumpridos.

Os parâmetros  $\lambda'_{i,c}$ s são pesos estatísticos a serem estimados. Esses indicam o quão fortemente a característica  $f_i$  e a classe  $c$  estão relacionadas. Os valores de tais parâmetros são definidos de modo a maximizar a entropia da distribuição induzida (daí o nome do classificador) sujeita à restrição de que os valores esperados das funções *feature/class* em relação ao modelo são iguais aos valores esperados em relação aos dados de treinamento. Por fim, a expressão  $Z(e)$  representa um fator de normalização [60].

## 6 Modelo neural de múltiplos canais para análise de sentimento

### 6.1 Fundamentos teóricos do MONEMC

No modelo de AS proposto, um texto  $e$ , objeto de classificação, é dividido em  $n_i$  agrupamentos de palavras sequenciais  $p_i$ . Para ser mais preciso, dado um texto  $e = \{n_1, n_2, \dots, n_r\}$ , esse será formado por um número total de  $N$  sentenças, onde cada sentença  $n_i$  agirá como um agente interagente na dinâmica a ser descrita. Ainda, adota-se que a partição  $n_i$  é construída a partir dos caracteres textuais que indicam: fim de uma ideia, pausa em um discurso e separação de orações presentes no texto  $e$ , ou seja, os sinais de pontuação. Como consequência, um conjunto  $n_i$  possui um número restrito de palavras  $L^i$ , podendo essas estarem (ou não) associadas a algum sentimento, ou seja:

$$e = \{ \{p_{n_1}^1, p_{n_1}^2, \dots, p_{n_1}^{L_1}\}, \{p_{n_2}^1, p_{n_2}^2, \dots, p_{n_2}^{L_2}\}, \{p_{n_3}^1, p_{n_3}^2, \dots, p_{n_3}^{L_3}\}, \dots, \{p_{n_r}^1, p_{n_r}^2, \dots, p_{n_r}^{L_r}\} \} \quad (6.1.1)$$

A subdivisão do texto  $e$  em subconjuntos de trechos ( $n_i$ ) que por sua vez serão subdivididos em outros subconjuntos de palavras sequenciais ( $p_i$ ) são as bases para a sistematização do algoritmo. Assim, dada a palavra  $p_i^j$  essa poderá ser relacionada a uma ou mais classes de sentimentos  $C = \{c_1, c_2, \dots, c_{A-1}, c_A\}$  através de  $P(p_{n_i}^j \in c_m)$ , a probabilidade da palavra  $p_i^j$  remeter a uma classe de sentimento  $c_m$ .

Nesse sentido, como elemento estruturante da classificação, construiu-se um dicionário léxico  $D(T)$  contendo um total de 2048 palavras únicas formando 2115 relações da forma:

$$p_i^j \rightarrow (c_m, P(p_{n_i}^j \in c_m)) \quad (6.1.2)$$

A critério de simplificação, inferimos que as palavras contidas no dicionário inicial  $D(T)$  remetessem a no máximo três classes onde, arbitrariamente, atribui-se:  $P(p_{n_i}^j \in c_m) = 1/3$ ,  $P(p_{n_i}^j \in c_m) = 1/2$  ou  $P(p_{n_i}^j \in c_m) = 1$ , respectivamente, respeitando-se a condição de normalização.

Tabela 6.1.1 – Exemplo de dicionário base  $D(T)$ 

Words	Sentiment	Probability
preference	fear	1
pride	joy	1/2
pride	confident	1/2
trusting	confident	1
remorse	sadness	1
stung	anger	1
sublime	surprise	1
timorously	fear	1
tiny	analytical	1

Fonte: o próprio autor.

Tabela 6.1.1: Dicionário base construído a partir da biblioteca Sentiment da linguagem R.

Para a classificação foram fixadas oito classes ( $A = 8$ ) ou canais, baseando-se em outros classificadores de sentimentos como o IBM Watson, o Qemotion e o Sentiment: raiva (*anger*), aversão (*disgust*), medo (*fear*), analítico (*analytical*), surpresa (*surprise*), alegria (*joy*), tristeza (*sadness*) e confiança (*confident*).

Após realizado a etapa de pré-processamento (tokenização e filtragem) do texto contíguo a estruturação acima referida, infere-se que cada agente  $n_i$  possui um potencial de significado  $\overrightarrow{V_{n_i}(t)}$  distribuído em  $A$  canais que representam as classes ou sentimentos escolhidos para o analisador de sentimento:

$$\overrightarrow{V_{n_i}(t)} = \sum_{m=1}^8 V_{n_i}(t) \hat{c}_m \quad (6.1.3)$$

Na sequência, afim de adotar o tratamento estocástico, o tempo é discretizado em intervalos iguais  $\tau > 0$ , i.e.,  $t\tau, t\tau \in \mathbb{N}$ .

Além disso, em  $t = 0$  o valor do potencial dos múltiplos canais  $c_m$  é fixado através da relação (6.1.2). Em outras palavras, a relação entre as palavras contidas no dicionário léxico  $D(T)$  e seus sentimentos, onde:  $V_{n_i}^{c_m}(0) = \sum_j P(p_{n_i}^j \in c_m)$  representa a soma das probabilidades de todas as palavras rotuladas do trecho  $n_i$  pertencerem ao sentimento  $c_m$ . Para as demais palavras  $p_{n_i}^j \notin c_m$ , as que não pertencem a nenhuma classe preestabelecida, o potencial nulo e a classe "unknown" são definidas e garantimos que essas entidades não possuem influência sobre a dinâmica do sistema.

Para o problema, escolheu-se uma rede regular colunar onde cada agente  $n_i$  é disposto em um nó  $i$  de forma sequencial ( $n_i = 1, 2, 3, \dots$ ). Desse modo, inferimos que o peso estatístico da ligação bidirecional entre os agentes  $n_i$  e  $n_v$  ( $n_i \neq n_v$ ) representado por  $G_{iv} = G_{vi}$  é independente do tempo e segue a formulação:

$$G_{iv} = e^{-(n_i - n_v)^2} \quad (6.1.4)$$



note que a (6.1.4) representa a distância euclidiana entre dois nós da rede colunar e, essa decresce exponencialmente de acordo com o aumento da distancia entre os agentes orientados na rede, o que significa dizer que a influencia mutua entre trechos mais distantes é menor.

Além disso, de forma similar ao modelo GGL o mecanismo de disparo (acionamento) das interações é associado a variável booleana  $X_{n_v}^{c_m}(t)$  onde a probabilidade de disparo  $P(X_{n_v}^{c_m}(t) = 1)$  é calculada a partir de uma *função de disparo* (FD)  $\Phi_m(V_{n_i}(t)\hat{c}_m = V_{n_i}^{c_m}(t))$ . Essa última, dependente do potencial de significado de cada canal do agente  $n_i$ , isoladamente. Para esse caso, assume-se que, matematicamente, a FD é representada por uma função sigmóide em cada um dos seus canais:

$$\Phi_{c_m}(V_{n_i}^{c_m}(t)) = \frac{1}{1 + e^{-V_{n_i}^{c_m}(t)}} \quad (6.1.5)$$

Portanto, permite-se que os agentes interajam com seus semelhantes dispostos em rede através dos  $A$  canais simultaneamente.

Doutro modo, diferentemente do mecanismo de disparo presente em grande maioria dos modelos de neurônio, onde há a presença de um potencial de reset ( $V_R$ ), em que  $V(t) = V_R$ , após um disparo. No proposto modelo esse tipo de comportamento de inspiração biológica é omitido. Desse modo, define-se a dinâmica de  $V_{n_i}^m(t)$ , isto é, a evolução temporal do potencial de ativação em cada canal  $c_m$  como:

$$\begin{aligned} V_{n_i}^{c_m}(t+1) &= \mu_{n_i}^{c_m} V_{n_i}^{c_m}(t) + \\ &+ \frac{1}{N-1} \sum_{n_v \neq n_i}^{N-1} G_{n_i n_v} (V_{n_v}^{c_m}(t) - V_{n_i}^{c_m}(t)) \theta(V_{n_v}^{c_m}(t) - V_{n_i}^{c_m}(t)) X_{n_v}^{c_m}(t) - \\ &- \frac{1}{N-1} \sum_{n_v \neq n_i}^{N-1} G_{n_v n_i} (V_{n_i}^{c_m}(t) - V_{n_v}^{c_m}(t)) \theta(V_{n_i}^{c_m}(t) - V_{n_v}^{c_m}(t)) X_{n_i}^{c_m}(t) + \\ &+ \frac{1}{A-1} \sum_{c_m \neq c_q}^{A-1} g_{n_i} (V_{n_i}^{c_m}(t) - V_{n_i}^{c_q}(t)) \theta(V_{n_i}^{c_m}(t) - V_{n_i}^{c_q}(t)) X_{n_i}^{c_m}(t) - \\ &- \frac{1}{A-1} \sum_{c_m \neq c_q}^{A-1} g_{n_i} (V_{n_i}^{c_q}(t) - V_{n_i}^{c_m}(t)) \theta(V_{n_i}^{c_q}(t) - V_{n_i}^{c_m}(t)) X_{n_i}^{c_q}(t) \end{aligned} \quad (6.1.6)$$

Particularmente, em (6.1.6) no primeiro termo está representado a influência das condições iniciais na dinâmica. Na medida que a variável aleatória  $\mu_{n_i}^{c_m}$  para o canal  $c_m$  e agente  $n_i$  é independente do tempo e segue uma distribuição exponencial:

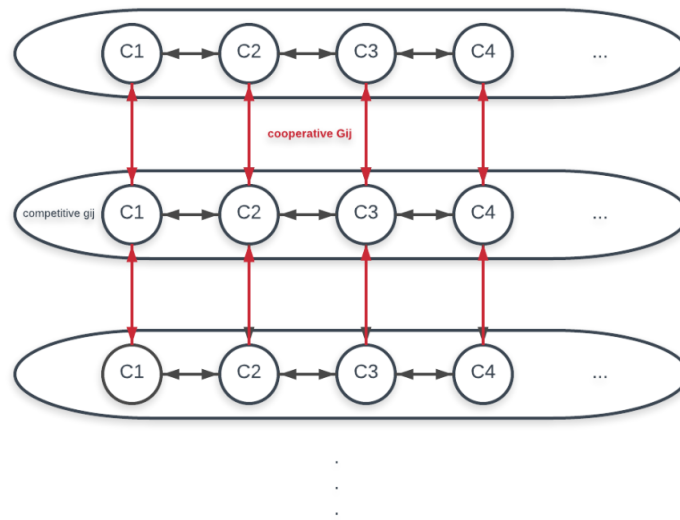
$$\rho(\mu_{n_i}^{c_m}) = 1 - e^{-\lambda^{c_m} V_{n_i}^{c_m}(0)} \quad (6.1.7)$$

onde  $\lambda^{c_m}$  é a taxa inicial de potencial por canal  $\lambda^{c_m} = \begin{cases} \frac{\overline{V_{n_i}^{c_m}(0)}}{\overline{V(0)}} & \text{se } \overline{V(0)} > 0 \\ 0 & \text{se } \overline{V(0)} \leq 0 \end{cases}$ .

note que  $\overline{V(t)}$  representa a média do potencial  $V_{n_i}^{c_m}(t)$  sobre todos os agentes e canais no tempo  $t$ :

$$\overline{V(t)} = \frac{\sum_{c_m} \sum_{n_i} V_{n_i}^{c_m}(t)}{N.A}. \quad (6.1.8)$$

Figura 6.1.1 – Diagrama MONEMC



Fonte: o próprio autor.

Figura 6.1.1: MONEMC diagrama ilustrativo.

Nos demais termos de (6.1.6), observa-se dois tipos de interações características: a interação entre diferentes sentenças (agentes) e, a que ocorre entre canais de um mesmo agente (classes). Em ambas as interações há a presença de um mecanismo do tipo diodo representado pela função de Heaviside  $\theta$ . No segundo e terceiro termos associados as interações na rede, impõem-se que as trocas entre agentes acontecem de forma independente para cada canal  $c_m$ . Nesse caso, o mecanismo de diodo obriga que o fluxo de potencial seja sempre positivo,  $\Delta V_{n_i n_v}^{c_m} > 0$ , ou seja, uma interação estabelecida entre o par  $(n_i, n_v)$  para um específico canal  $c_m$  possuirá, obrigatoriamente, um sentido de fluxo de potencial, onde o neurônio de maior potencial perderá uma parcela desse valor para seu vizinho em um regime de *cooperação*.

No quarto e quinto termos, onde são representados a interação intra-neurônio  $n_i$  e entre canais  $(c_m, c_q)$  temos a situação invertida: uma situação de *competição*. Essa representada pelo fluxo de potencial que acontece, obrigatoriamente, do canal de menor potencial para o de maior potencial. Vale ressaltar que esse último mecanismo é o grande responsável pela convergência do classificador. Ainda, definimos :

$$g_{n_i} = e^{\frac{L_{n_i}}{\sum_{n_i} L_{n_i}}} \quad (6.1.9)$$

onde  $L_{n_i}$  representa o número de palavras contidas na sentença  $n_i$ .

Assim, a classificação de um texto  $e$  está associada ao número de disparos global da rede  $\sum_{n_i} X_{n_i}^{c_m}(t')$ . Ou seja, a probabilidade de um texto  $e$  ser classificado como da classe  $c_m$  em uma única simulação  $s$  será calculado a partir da média temporal do número de disparos do neurônio  $n_i$  na classe  $c_m$  após um certo tempo de "equilíbrio"  $t_{eq}$  (representado por  $\langle X_{n_i}^{c_m}(t') \rangle$ ) em uma simulação  $s$ . Naturalmente, esse valor será

somado sobre todos os neurônios e todas as classes como pode ser visto na expressão a seguir:

$$P_s(e \in c_m) = \frac{\sum_{n_i} \langle X_{n_i}^{c_m}(t') \rangle_s}{\sum_{c_m} \sum_{n_i} \langle X_{n_i}^{c_m}(t') \rangle_s} \quad (6.1.10)$$

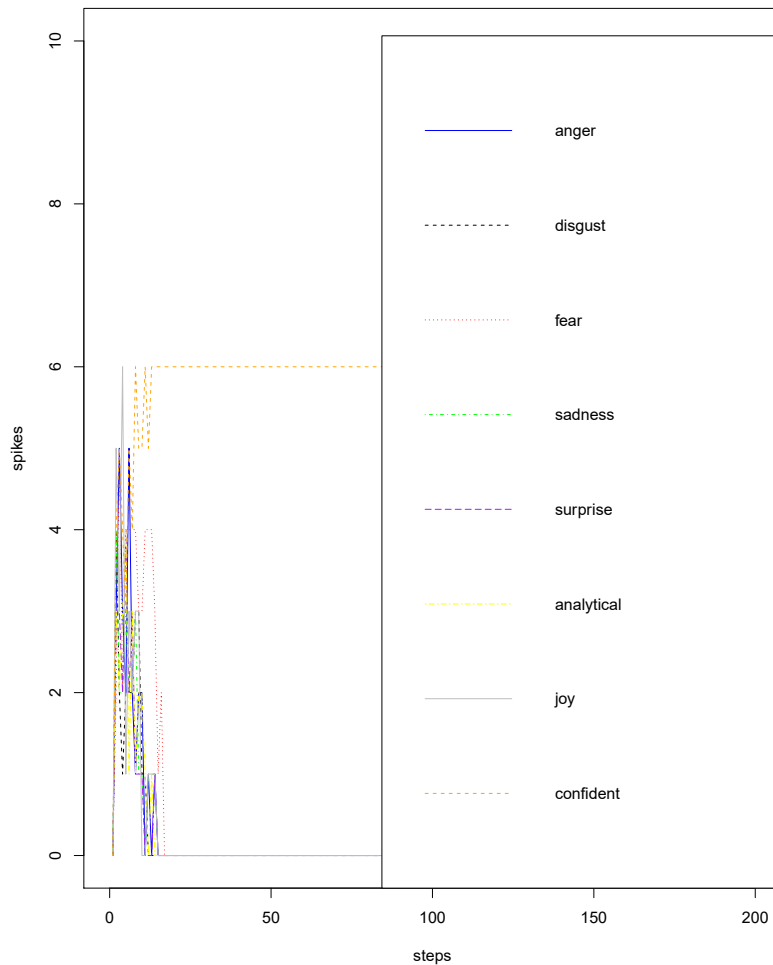
onde  $t' = t_{eq} - t \mid t' > t_{eq}$ .

Portanto, para um conjunto de  $S$  simulações de um mesmo texto  $e$  temos que :

$$P(e \in c_m) = \frac{\sum_s^S P_s(e \in c_m)}{S}. \quad (6.1.11)$$

A expressão acima, chave para o modelo, faz a ligação entre o número de disparos da rede e a classificação do texto.

Figura 6.1.2 – Exemplo de classificação



**Fonte:** o próprio autor.

Figura 6.1.2: MONEMC com  $t = 100$  e  $S = 1$  para o texto  $e =$  "You gain strength, courage, and confidence by every experience in which you really stop to look fear in the face. You are able to say to yourself, 'I lived through this horror. I can take the next thing that comes along.'" (Eleanor Roosevelt).

Na figura 6.1.2 um exemplo de funcionamento do modelo proposto onde obteve-se como resultado  $P(e \in \text{confident}) = 1$ . Também desse simples exemplo é notória a rápida convergência do número de disparos ( $t_{eq} \sim 20$ ) em função dos passos de simulação.

## 6.2 Resultados e trabalhos futuros

Para a obtenção de um indicador da efetividade do modelo apresentado, utilizou-se um conjunto de 61 *twitters* onde esses foram pré-classificados por 5 pessoas (pós-graduandos e graduandos do curso de Física da Universidade Federal de Santa Catarina) de acordo com as 8 classes possíveis: raiva (*anger*), aversão (*disgust*), medo (*fear*), analítico (*analytical*), surpresa (*surprise*), alegria (*joy*), tristeza (*sadness*) e confiança (*confident*). Assim, uma medida de precisão pode ser definida  $\rho \in [0, 1]$  baseada na concordância do classificador com a decisão por votação das pessoas sobre o sentimento predominante em cada *twitter*:

$$\rho = \frac{\#acertos}{\#total \ twitters} \quad (6.2.1)$$

A escolha dos dados (61 *twitters*) para esse teste foi feita tendo como fonte a base Sentiment140 [64] um conjunto de 1.6 milhões de *twitters*. Ainda, levamos em consideração apenas textos que possuíssem em seu corpo as expressões: “*i feel*”, “*i am feeling*”, “*i’m feeling*”, “*i dont feel*”, “*I’m*”, “*Im*”, “*I am*”, e “*makes me*” [65] dos quais 61 foram escolhidos e posteriormente classificados.

Além disso, para uma comparação com outro método de classificação o algoritmo NB (ver 5.1.1) foi implementado, valendo-se do mesmo dicionário léxico utilizado no MONEMC.

Tabela 6.2.1 – Twitters e classificação

Twitter	MONEMC	NB	Pessoas
@teena_punjwani ohhhhh my i am glad they where @eresse 's rice cakes or i could have eaten the pack!!! lol	joy	joy	joy
im on msn at this time being im on the computer in my room and its soooo slow	confident	disgust	analytical
im a twitter virgin.....so be gentle with me please	joy	joy	joy
im going to sneak down to the fridge while no one is watching and it the cookie dough yum	analytical	confident	analytical
@zacofficial i’m already following him	analytical	disgust	analytical
im actually planning to watch dancing with the stars this year	analytical	analytical	analytical

Twitter	MONEMC	NB	Pessoas
im graduating tom!!! i can't believe it!! i'm soooo happy	joy	confident	joy
wat a beautiful day yesterday got tired joust right,and today i feel almost great nothing cloud get better today,maybe someone to hug	joy	analytical	joy
@tommcfly wish marvin good luck from me and he's the cutest kitten ever, so the other cats will love him anyhow xx	joy	joy	joy
im dragging daddy round the shops today	joy	disgust	analytical
i think i spoke 2 soond owner of ta cenc called me 2 see if im available to work as a poolgirl maybe i start working this week	analytical	analytical	analytical
felt really special earlier tonight... my nephew specifically asked for me to put him to bed tonight out of everyone at my bro's house!	joy	anger	joy
just havin a rehearsal, my entrance and model pose is stolen from jeff hardy im feeling so cool	analytical	analytical	joy
@sajina lmao!!! can't wait to see him in g.i. joe though!	sadness	confident	surprise
@trixy98 i feel so bad for you, but i'm really glad school's out. keep up the amazing work, girl	sadness	joy	joy
bed now still havnt figured out what im doing my lazzi's on. think i may have to wing it! later all, wish me luck xx	analytical	analytical	confident
my teeth hurt. im listening to dirty little secrets by the all-american rejects.	sadness	sadness	sadness
@jennifercarey i love boy bangs, but im not sure if i could get used to a guy with the same exact hairstlye as me	confident	confident	joy
day is getting better. having lunch with jen and kevin like old times which makes me very happy	joy	joy	joy
@unklerupert lol... damn drink. i think all the interviewers have been quite pleasant to him so far.	analytical	analytical	analytical
@katdish yes. being a pastor's kid has made me crazy. but loving jesus makes me crazier.	joy	joy	joy
i ate some food and now i feel a little better &lt	surprise	disgust	joy
@harriethawtin aww wow!! makes me feel so special x	surprise	surprise	surprise
@jchavz ok im ready	confident	confident	confident
@mgiraudoofficial ....i was thinking of how you went for a nice walk innyc and im gonna do it too..thanks for the idea!	analytical	joy	analytical

Twitter	MONEMC	NB	Pessoas
a lot stupider people have gone to africa and made it back. she will be fine. im waiting to see doc and help hi	surprise	joy	anger
im dying not having my phone. i am tempted to walk home and get it but im sure that i might not come back to work	surprise	sadness	sadness
@hasket im pretty pumped about the concert, and ill be waiting for deets on the music video thing, thats so sick!	disgust	disgust	joy
@tajim there is no point in commenting on such things everyone knows that we haven't attacked anyone ever!	confident	disgust	anger
@shayda june 15, 2009 haaay i'm shayda. and this is how excited i am that school is over! ... <a href="http://tinyurl.com/n6fsaw">http://tinyurl.com/n6fsaw</a>	joy	joy	joy
@cdr525 it's shit like that that makes me miss you	joy	joy	sadness
im on top of the world... going running	surprise	confident	confident
hehehe just came back few days ago from the greece,bulgaria and spain trip and im off to vegas sunday @heather_paige	joy	disgust	joy
babe. i am there for alot of people. sorry you didnt think so.	sadness	sadness	sadness
@backstreetboys i love you guys! you are the best !!! you are the only one makes me fell like this #bsb	joy	joy	joy
im boreddd. so boredd. thanks paige for making me laughhh	anger	disgust	sadness
why the hell am i always the victim of road rage jerkfucks? dear mr truck driver, fuck you	anger	anger	anger
i took two gravols because i was feeling motion sick. i ended up slept for 9 hours :l i am so not sleping tonight	disgust	disgust	sadness
@shaycarl that hurt a little..how dare you shay carl..im so mad i could...hug you	anger	sadness	sadness
i am excited to see all what god has instore for me	joy	joy	joy
i can say one thing for sure after today: the rich people of la know how to live. i am feeling very inspired to join their club	confident	confident	confident
i love my puppy. he makes me smile	joy	joy	joy
lying down listening to my ipod, realizing that i am the most lucky person alive i truly am the happiest i've ever been in my life!	confident	analytical	confident
@leftwanting my day's been good. omg, i feel like i've neglected ya'll & twitter today, lmao. i'm still trying to figure how i can go.	joy	joy	joy

Twitter	MONEMC	NB	Pessoas
@officialtl makes me more excited to see it. don't let me down! jacob is one of my top faves	joy	joy	surprise
wow, 100 followers! thank you! i feel locked in, like @iamdiddy.	joy	joy	joy
omg im planning a trip to tokyo the month after next im so excited about it!	joy	joy	surprise
@areloz thanks. i sometimes doubt myself when i feel a certain way, and i shouldn't. feelings like those should be cherished.	joy	joy	analytical
conducting an impact assessment with a partner organization. i feel like an aid worker.	joy	joy	analytical
my rabbit is so sweet. i love him so much	joy	joy	joy
truly thankful i am.	analytical	analytical	confident
im defiantly not voting for miley for the choice awards lol	surprise	disgust	analytical
i feel like this is decadent music â™« <a href="http://blip.fm/~8bcog">http://blip.fm/~8bcog</a>	joy	joy	sadness
@favier omg i never thought i look like i feel. thats bad	surprise	surprise	surprise
im a gamer, ive pretty much been a sony fanboy throughout my life, just downloaded lots of vids for new games, gonna be a good year	joy	joy	analytical
@jasminekpk i feel we should throw a global celebration party for lvatt. lvatt ftw.	analytical	disgust	joy
@melody1976 cool...just don't want to piss him off or drive him crazy! but i will continue to give him shit! lol	anger	analytical	anger
just got home from my exam tomrrow i have another one and them friday and im done ! yes babby ! wooo t	confident	analytical	analytical
last night was sad = ( but now i'm happy i am in paris is a beautiful place!	analytical	joy	joy
i am going to write about my recovery from chronic fatigue syndrome today - haven't thought about it much in over a year - yay	sadness	disgust	analytical
i feel like crap ! team miley	anger	joy	anger

Twitter	MONEMC	NB	Pessoas
---------	--------	----	---------

**Fonte:** o próprio autor.

Tabela 6.2.1: Classificação dos *twitters* a partir do modelo MONEMC ( $S = 48$ ,  $t = 80$  e estimamos:  $t_{eq} = 70$ ) e NB.

Isto posto, seguindo 6.2.1 pode-se realizar uma estimativa da precisão geral de cada algoritmo como é apresentado na tabela a seguir:

Tabela 6.2.3 – Comparação NB vs. MONEMC

	NB	MONEMC
$\rho(\%)$	45.9%	52.5%

**Fonte:** o próprio autor.

Tabela 6.2.3: medida da precisão geral dos classificadores NB e MONEMC

Os resultados preliminares indicam que o método de classificação proposto possui uma boa precisão, sendo essa superior ao método de NB mesmo quando esses partem do mesmo dicionário léxico base. Como ponto a se considerar, devido ao MONEMC ser estocástico e de relativa complexidade, seu tempo computacional, consequentemente, é superior se comparado a classificadores ágeis como o de NB.

Como trabalhos futuros o crescimento do dicionário  $D(T)$  torna-se parte fundamental para o aumento da precisão do MONEMC. Uma das técnicas que já comprovaram sucesso nessa tarefa é a *Label Propagation*[66, 67, 68, 69, 70]. Nesse sentido, através dessa abordagem conjunta (*Label propagation* + MONEMC) a aprendizagem do algoritmo passará a ser semi-supervisionada, i.e, utilizará termos já rotulados para a expansão do dicionário base. Nesse mesma linha, para uma análise mais precisa do classificador, mais *twitters* devem ser utilizados para a análise de precisão e outras medidas como *Recall* e *F1 Score* segmentadas por classe devem ser implementadas. Ainda, comparações com outros métodos de análise de sentimentos como: Máquina de vetores de suporte e Máxima Entropia são objeto de interesse.

Um outro ponto a se destacar são os pesos estatísticos  $G_{iv}$  e  $g_{n_i}$ , esses podem ser estudados de forma mais detalhada e até mesmo substituídos se necessário. Do mesmo modo, uma ampla análise da equação 6.1.6 deve ser realizada de modo que suas características principais sejam demonstradas de forma analítica.



## 7 Conclusão

Esse trabalho tem como característica suas duas áreas de atuação distintas: Mineração de Textos e Econofísica; na primeira é apresentada uma visão geral sobre os conceitos basilares da análise de sentimentos em textos. Além disso, apresenta-se em detalhes o modelo para análise de textos e classificação de sentimentos desenvolvido com base no modelo GGL. Sobre essa formulação, ressalta-se que todo o algoritmo computacional, como também, o dicionário base foram desenvolvidos para a língua inglesa. Isto posto, as categorias de classificação escolhidas foram: raiva (*anger*), aversão (*disgust*), medo (*fear*), analítico (*analytical*), surpresa (*surprise*), alegria (*joy*), tristeza (*sadness*) e confiança (*confident*). Os resultados indicam que o método de classificação proposto possui uma boa precisão, sendo essa superior ao método de NB mesmo quando esses partem do mesmo dicionário léxico base. Como ponto a se considerar, devido ao MONEMC ser estocástico e de relativa complexidade, seu tempo computacional, conseqüentemente, é superior se comparado a classificadores ágeis como o de NB.

Como trabalhos futuros o crescimento do dicionário  $D(T)$  torna-se parte fundamental para o aumento da precisão do MONEMC. Uma das técnicas que já comprovaram sucesso nessa tarefa é a *Label Propagation*. Nesse sentido, através dessa abordagem conjunta (Label propagation + MONEMC) a aprendizagem do algoritmo passará a ser semi-supervisionada, i.e, utilizará termos já rotulados para a expansão do dicionário base.

Um outro ponto a se destacar são os pesos estatísticos  $G_{iv}$  e  $g_{n_i}$ , esses podem ser estudados de forma mais detalhada e até mesmo substituídos se necessário. Do mesmo modo, uma ampla análise da dinâmica que envolve os múltiplos canais de potencial deve ser realizada de modo que suas características principais sejam demonstradas de forma analítica.

Quanto aos estudos em Econofísica, o primeiro tem foco na análise da distribuição de renda no Brasil a partir do Modelo de Duas Classes de Yakovenko. Os resultados sugerem uma diminuição da desigualdade de renda brasileira no período selecionado. Outro resultado importante, refere-se ao fato de que os subgrupos historicamente desfavorecidos (Mulheres e pretos / indígenas / pardos), maioria da população, têm uma distribuição de renda mais igualitária. Esses grupos têm, também, uma renda mensal menor que os demais e essa estrutura social permaneceu praticamente inalterada no período de tempo investigado (2001-2014).

O texto final apresenta o Modelo Econômico com Concentração de Riqueza (MECR), um novo modelo proposto pelo autor para a distribuição de riqueza. Este modelo sugere a divisão da população em classes alta e baixa, com mecanismos distin-

tos para a geração de riqueza, resultando em duas dinâmicas segregadas de evolução temporal. No MECR, a separação dos agentes entre as dinâmicas é realizada através de um mecanismo de "integra e dispara", associado a uma função de ativação cujo potencial é a própria riqueza. Uma premissa fundamental deste modelo é a acumulação de riqueza por uma parcela pequena da população. A partir dessa premissa, o modelo desenvolve um construto teórico que culmina em um estado de equilíbrio com uma distribuição de riqueza do tipo Laplace-Pareto-1. As simulações realizadas para testar a dinâmica do modelo demonstram uma concordância significativa com a teoria proposta, reforçando os achados teóricos dentro das condições de contorno estabelecidas.

# Referências

- [1] MANTEGNA, R. N.; STANLEY, H. E. *Introduction to econophysics: correlations and complexity in finance*. [S.l.]: Cambridge university press, 1999.
- [2] SLANINA, F. *Essentials of Econophysics Modelling*. Oxford University Press, 2013. ISBN 9780199299683. Disponível em: <<https://doi.org/10.1093/acprof:oso/9780199299683.001.0001>>.
- [3] CHETVIORKIN, I.; LOUKACHEVITCH, N. Two-step model for sentiment lexicon extraction from twitter streams. In: *Proceedings of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. [S.l.: s.n.], 2014. p. 67–72.
- [4] NISHAJEBASEELI, A.; KIRUBAKARAN, E. A survey on sentiment analysis of (product) reviews. *International Journal of Computer Applications*, Foundation of Computer Science, v. 47, n. 11, p. 36–39, jun 2012.
- [5] MEDHAT, W.; HASSAN, A.; KORASHY, H. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, Elsevier BV, v. 5, n. 4, p. 1093–1113, dec 2014.
- [6] YUE, L. et al. A survey of sentiment analysis in social media. *Knowledge and Information Systems*, Springer, v. 60, p. 617–663, 2019.
- [7] GALVES, A.; LÖCHERBACH, E. Infinite systems of interacting chains with memory of variable length—a stochastic model for biological neural nets. *Journal of Statistical Physics*, Springer Nature, v. 151, n. 5, p. 896–921, mar 2013.
- [8] GERSTNER, W.; HEMMEN, J. L. van. Associative memory in a network of ‘spiking’ neurons. *Network: Computation in Neural Systems*, Informa UK Limited, v. 3, n. 2, p. 139–164, jan. 1992. ISSN 1361-6536.
- [9] SICILIANI, I. D.; TRAGTENBERG, M. Kinetic theory and brazilian income distribution. *Physica A: Statistical Mechanics and its Applications*, Elsevier BV, v. 513, p. 166–174, jan. 2019. ISSN 0378-4371.
- [10] PARETO, V. *Cours d’Économie Politique*. Macmillan. [S.l.]: New York, 1897. II.
- [11] CLEMENTI, F.; GALLETI, M. Power law tails in the italian personal income distribution. *Physica A: Statistical Mechanics and its Applications*, v. 350, p. 427–438, 2005.

- [12] JAYADEV, A. A power law tail in India's wealth distribution: Evidence from survey data. *Physica A: Statistical Mechanics and its Applications*, Elsevier BV, v. 387, n. 1, p. 270–276, jan 2008.
- [13] DRĂGULESCU, A.; YAKOVENKO, V. Evidence for the exponential distribution of income in the USA. *The European Physical Journal B*, Springer Nature, v. 20, n. 4, p. 585–589, apr 2001.
- [14] ARATA, Y. *Income Distribution among Individuals: The effects of economic interactions*. [S.l.], 2013.
- [15] DRĂGULESCU, A.; YAKOVENKO, V. M. Exponential and power-law probability distributions of wealth and income in the united kingdom and the united states. *Physica A: Statistical Mechanics and its Applications*, Elsevier BV, v. 299, n. 1-2, p. 213–221, oct 2001.
- [16] SILVA, A. C.; YAKOVENKO, V. M. Temporal evolution of the “thermal” and “superthermal” income classes in the USA during 1983–2001. *Europhysics Letters (EPL)*, IOP Publishing, v. 69, n. 2, p. 304–310, jan 2005.
- [17] SHAIKH, A.; PAPANIKOLAOU, N.; WIENER, N. Race, gender and the econophysics of income distribution in the USA. *Physica A: Statistical Mechanics and its Applications*, Elsevier BV, v. 415, p. 54–60, dec 2014.
- [18] SANTOS, P. H. dos; SICILIANI, I. D. S.; TRAGTENBERG, M. H. R. Optimal income crossover for a two-class model using particle swarm optimization. *Physical Review E*, American Physical Society (APS), v. 106, n. 3, p. 034313, set. 2022. ISSN 2470-0053.
- [19] SORIANO-HERNÁNDEZ, P. et al. Non-stationary individual and household income of poor, rich and middle classes in mexico. *Physica A: Statistical Mechanics and its Applications*, Elsevier BV, v. 465, p. 403–413, jan 2017.
- [20] FIGUEIRA, F. C.; MOURA, N.; RIBEIRO, M. The gompertz–pareto income distribution. *Physica A: Statistical Mechanics and its Applications*, Elsevier BV, v. 390, n. 4, p. 689–698, feb 2011.
- [21] CALVELLI, M.; CURADO, E. M. F. A wealth distribution agent model based on a few universal assumptions. *Entropy*, v. 25, n. 8, 2023. ISSN 1099-4300. Disponível em: <<https://www.mdpi.com/1099-4300/25/8/1236>>.
- [22] BOUCHAUD, J.-P.; MÉZARD, M. Wealth condensation in a simple model of economy. *Physica A: Statistical Mechanics and its Applications*, Elsevier BV, v. 282, n. 3-4, p. 536–545, jul 2000.

- [23] SOLOMON, S. Generalized Lotka-Volterra (glv) Models and generic emergence of scaling laws in stock markets. 1999. Cond-mat/9901250.
- [24] SOLOMON, S.; RICHMOND, P. Stable power laws in variable economies; Lotka-Volterra implies pareto-zipf. *The European Physical Journal B - Condensed Matter*, Springer Nature, v. 27, n. 2, p. 257–261, may 2002.
- [25] MALCAI, O. et al. Theoretical analysis and simulations of the generalized lotka-volterra model. *Physical Review E*, American Physical Society (APS), v. 66, n. 3, sep 2002.
- [26] HUANG, D.-w. Wealth accumulation with random redistribution. *Physical Review E*, American Physical Society (APS), v. 69, n. 5, p. 057103, maio 2004. ISSN 1550-2376.
- [27] WILLIS, G. Income distribution and income shares: wealth and income distributions explained using generalised lotka-volterra SFC ABM models. *International Review of Applied Economics*, Informa UK Limited, v. 29, n. 6, p. 816–842, jul 2015.
- [28] ANGLIN, P. M. Econophysics of wealth distribution: A comment. In: *New Economic Windows*. [S.l.]: Springer Milan, 2005. p. 229–238.
- [29] BERMAN, Y.; BEN-JACOB, E.; SHAPIRA, Y. The dynamics of wealth inequality and the effect of income distribution. *PLOS ONE*, Public Library of Science (PLoS), v. 11, n. 4, p. e0154196, abr. 2016. ISSN 1932-6203.
- [30] SINHA, S. Evidence for power-law tail of the wealth distribution in india. *Physica A: Statistical Mechanics and its Applications*, v. 359, p. 555–52, 2006.
- [31] YAKOVENKO, V. M. Econophysics, statistical mechanics approach to. In: *Encyclopedia of Complexity and Systems Science*. [S.l.]: Springer New York, 2009. p. 2800–2826.
- [32] LUDWIG, D.; YAKOVENKO, V. M. Physics-inspired analysis of the two-class income distribution in the usa in 1983-2018. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, The Royal Society, v. 380, n. 2224, abr. 2022. ISSN 1471-2962.
- [33] YAKOVENKO, V. M.; ROSSER, J. B. Colloquium: Statistical mechanics of money, wealth, and income. *Rev. Mod. Phys.*, American Physical Society, v. 81, p. 1703–1725, Dec 2009. Disponível em: <<https://link.aps.org/doi/10.1103/RevModPhys.81.1703>>.
- [34] TODA, A. A. Income dynamics with a stationary double pareto distribution. *Physical Review E*, American Physical Society (APS), v. 83, n. 4, p. 046122, abr. 2011. ISSN 1550-2376.

- [35] SOUMA, W. Physics of personal income. 2002. Cond-mat/0202388v1. Disponível em: <cond-mat/0202388v1>.
- [36] MAIO, F. G. D. Income inequality measures. *Journal of Epidemiology Community Health*, BMJ, v. 61, n. 10, p. 849–852, oct 2007.
- [37] GINI, C. *Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche.*[Fasc. I.]. [S.l.]: Tipogr. di P. Cuppini, 1912.
- [38] HOFFMANN, R.; DUARTE, J. The distribution of income in brazil. *Rev. adm. empres*, v. 12, p. 46–66, 1972.
- [39] GEOGRAPHY, B. I. of; (IBGE), S. <https://www.ibge.gov.br/>. Disponível em: <<https://www.ibge.gov.br/>>.
- [40] FARIA, M. S.; BRANCO, N. S.; TRAGTENBERG, M. H. R. Nonuniversal behavior for aperiodic interactions within a mean-field approximation. *Physical Review E*, American Physical Society (APS), v. 77, n. 4, apr 2008.
- [41] WIKIPEDIA. [https://en.wikipedia.org/wiki/Gini\\_coefficient](https://en.wikipedia.org/wiki/Gini_coefficient). Disponível em: <[https://en.wikipedia.org/wiki/Gini\\_coefficient](https://en.wikipedia.org/wiki/Gini_coefficient)>.
- [42] MIMKES, J.; WILLIS, G. Lagrange principle of wealth distribution. In: *New Economic Windows*. [S.l.]: Springer Milan, 2005. p. 61–69.
- [43] CLEMENTI, F.; GALLETI, M. Pareto's law of income distribution: Evidence for germany, the united kingdom, and the united states. *Econophysics Of Wealth Distributions*, p. 3–14, 2005.
- [44] PIKETTY, T.; POSTEL-VINAY, G.; ROSENTHAL, J.-L. Wealth concentration in a developing economy: Paris and france, 1807–1994. *American economic review*, American Economic Association, v. 96, n. 1, p. 236–256, 2006.
- [45] BLANCHET, T.; MARTÍNEZ-TOLEDANO, C. Wealth inequality dynamics in europe and the united states: Understanding the determinants. *Journal of Monetary Economics*, Elsevier, v. 133, p. 25–43, 2023.
- [46] PFEFFER, F. T.; WAITKUS, N. The wealth inequality of nations. *American Sociological Review*, SAGE Publications Sage CA: Los Angeles, CA, v. 86, n. 4, p. 567–602, 2021.
- [47] ARMANTIER LUIS ARMONA, G. D. G. O.; KLAUW, W. van der. *Which Households Have Negative Wealth?* 2016. Disponível em: <<https://www.newyorkfed.org/aboutthefed>>.

- [48] KESTEN, H. Random difference equations and renewal theory for products of random matrices. *Acta Mathematica*, International Press of Boston, v. 131, n. 0, p. 207–248, 1973.
- [49] FUJIMOTO, S.; ISHIKAWA, A.; TOMOYOSE, M. A multiplicative stochastic process deriving the probability distribution in exact form. In: . [S.l.: s.n.], 2010. (221).
- [50] SORNETTE, D.; CONT, R. Convergent multiplicative processes repelled from zero: Power laws and truncated power laws. *Journal de Physique I*, v. 7, n. 3, p. 431–444, 1997.
- [51] MOSHE, L.; SOLOMON, S. New evidence for the power-law distribution of wealth. *Physica A: Statistical Mechanics and its Applications*, v. 242, p. 90–94, 1997.
- [52] R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2021. Disponível em: <<https://www.R-project.org/>>.
- [53] ZUCMAN, G. *Distributional National Accounts (DINA) with an Application to the United States*. 2020. Accessed: 2024-02-04. Disponível em: <<https://gabriel-zucman.eu/usdina/>>.
- [54] PIKETTY, T.; SAEZ, E.; ZUCMAN, G. Distributional national accounts: methods and estimates for the united states. *The Quarterly Journal of Economics*, Oxford University Press, v. 133, n. 2, p. 553–609, 2018.
- [55] SAEZ, E.; ZUCMAN, G. The rise of income and wealth inequality in america: Evidence from distributional macroeconomic accounts. *Journal of Economic Perspectives*, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203-2418, v. 34, n. 4, p. 3–26, 2020.
- [56] ALLAHYARI, M. et al. A brief survey of text mining: Classification, clustering and extraction techniques. *arXiv*, 2017.
- [57] LI, N.; WU, D. D. Using text mining and sentiment analysis for online forums hotspot detection and forecast. *Decision Support Systems*, Elsevier BV, v. 48, n. 2, p. 354–368, jan 2010.
- [58] WILSON, T.; WIEBE, J.; HOFFMANN, P. Recognizing contextual polarity in phrase-level sentiment analysis. In: . Vancouver, CA: [s.n.], 2005. Disponível em: <<https://www.microsoft.com/en-us/research/publication/recognizing-contextual-polarity-in-phrase-level-sentiment-analysis/>>.

- [59] JAIN, T. I.; NEMADE, D. Recognizing contextual polarity in phrase-level sentiment analysis. *International Journal of Computer Applications*, Foundation of Computer Science, v. 7, n. 5, p. 12–21, sep 2010.
- [60] PANG, B.; LEE, L.; VAITHYANATHAN, S. Thumbs up? In: *Proceedings of the ACL-02 conference on Empirical methods in natural language processing*. [S.l.]: Association for Computational Linguistics, 2002.
- [61] MCCALLUM, A.; NIGAM, K. A comparison of event models for naive bayes text classification. *AAAI Technical Report*, 1998.
- [62] PATODKAR, V. N.; I.R, S. Twitter as a corpus for sentiment analysis and opinion mining. *IJARCCCE*, Tejass Publisheers, v. 5, n. 12, p. 320–322, dec 2016.
- [63] NIGAM, K. Using maximum entropy for text classification. In: *In IJCAI-99 Workshop on Machine Learning for Information Filtering*. [S.l.: s.n.], 1999. p. 61–67.
- [64] GO, A.; BHAYANI, R.; HUANG, L. Twitter sentiment classification using distant supervision. *CS224N Project Report, Stanford*, v. 1, n. 12, 2009.
- [65] BOLLEN, J.; MAO, H.; ZENG, X. Twitter mood predicts the stock market. *Journal of computational science*, Elsevier, v. 2, n. 1, p. 1–8, 2011.
- [66] ZHU, X.; GHAHRAMANI, Z. Learning from labeled and unlabeled data with label propagation. Citeseer, 2002.
- [67] ZHU, X. Semi-supervised learning literature survey. Citeseer, 2005.
- [68] BODÓ, Z.; CSATÓ, L. A note on label propagation for semi-supervised learning. *Acta Universitatis Sapientiae, Informatica*, De Gruyter Open, v. 7, n. 1, p. 18–30, 2015.
- [69] YANG, M. et al. Lcct: a semisupervised model for sentiment classification. In: ASSOCIATION FOR COMPUTATIONAL LINGUISTICS (ACL). *Human Language Technologies: The 2015 Annual Conference of the North American Chapter of the ACL*. [S.l.], 2015.
- [70] GIULIANELLI, M. Semi-supervised emotion lexicon expansion with label propagation and specialized word embeddings. *arXiv preprint arXiv:1708.03910*, 2017.
- [71] BROCHINI, L. et al. Phase transitions and self-organized criticality in networks of stochastic spiking neurons. *Scientific Reports*, Springer Nature, v. 6, p. 35831, nov 2016.



# Apêndices



# APPENDIX A – O Modelo de Gerstner-Galves-Löcherbach

O modelo de Gerstner-Galves-Löcherbach (GGL) desenvolvido pelo neurocientista computacional alemão e suíço Wulfram Gerstner no ano de 1991 e aperfeiçoado pelos matemáticos Antonio Galves e Eva Löcherbach no ano de 2012[7, 8] é um modelo estocástico inicialmente proposto para o estudo das redes neurais interagentes, nele os disparos estão condicionados a uma certa probabilidade de ocorrer. Ou seja, cada emissão é associada a uma função chamada de “*firing function*”  $\phi$  interagindo com a evolução do sistema desde o último disparo.

Assim, considerando uma cadeia estocástica  $X_t$  com valor 1 para disparo do neurônio  $i$  no tempo  $t$  e 0 para os demais casos, a probabilidade de um disparo ocorrer para um neurônio  $i$  no tempo  $t$  será dada por:

$$P(X_t(i) = 1 | F_{t-1}) = \phi_i\left(\sum_j W_{j \rightarrow i} g_j(t-s) X_s(j), t - L_i^t\right) \quad (\text{A.o.1})$$

Onde  $g_j$  é uma função filtro,  $s$  é o tempo do último disparo e  $W_{j \rightarrow i}$  significa o peso sináptico das interações de  $j$  com  $i$ .

Como consequência, o modelo GGL é inerentemente estocástico, destacando-se dos seus anteriores do tipo integra-e-dispara. É também, um modelo que condiciona a probabilidade de disparo a toda a evolução do sistema, sendo assim, não-Markoviano.

Neste contexto, Brochini e colaboradores [71] apresentaram uma simplificação para o modelo GGL a partir da aproximação de campo médio, na qual investigaram as características apresentadas pelo sistema na transição de fase e o seu comportamento crítico.

Nessa última abordagem, assume-se uma rede de  $N$  neurônios GGL que mudam de estado paralelamente em intervalos de amostragem com um espaçamento  $\Delta$ . Assim, o potencial de membrana do neurônio  $i$  é modelado por uma variável real  $V_i[t]$  indexada pelo tempo discreto  $t$ , um inteiro que representa o tempo de amostragem  $t\Delta$ .

Cada sinapse transmite sinais de algum neurônio pré-sináptico  $j$  para algum neurônio pós-sináptico  $i$ , e tem um força sináptica  $w_{ij}$ . Se o neurônio  $j$  dispara entre os tempos discretos  $t$  e  $t + 1$ , seu potencial cai para a  $V_R$ . Este evento aumenta por  $w_{ij}$  o potencial de cada neurônio pós-sináptico  $i$  que não dispara nesse intervalo. O potencial de um neurônio que não disparou também pode receber um estímulo externo  $I_i[t]$ , que modela sinais recebidos de fontes fora da rede. Além desses incrementos, o potencial de um neurônio que não disparou decai em cada etapa de tempo em direção à tensão de base  $V_B$  por um fator  $\mu \in [0, 1]$ , que modela o efeito de uma corrente de fuga.

Além disso, introduz-se a variável booleana  $X_i[t] \in \{0, 1\}$  que denota se o neurônio  $i$  disparou entre  $t$  e  $t + 1$ . Assim, os potenciais evoluem como:

$$V_i[t + 1] = \begin{cases} V_R & \text{se } X_i[t] = 1 \\ \mu(V_i[t] - V_B) + V_B + I_i[t] + \sum_{j=1}^N w_{ij}X_j[t] & \text{se } X_i[t] = 0 \end{cases} \quad (\text{A.0.2})$$

Este é um caso especial do modelo GGL, com a função de filtro  $g(t - t_s) = \mu^{t-t_s}$ , onde  $t_s$  é o tempo do último disparo de neurônio  $i$ . Ainda,  $X_i[t + 1] = 1$  com probabilidade  $\Phi(V_i[t])$ , que é chamada de função de disparo. Doutro modo,  $X_i[t + 1] = 0$  se  $X_i[t] = 1$  (período refratário). A função  $\Phi$  é sigmóide, isto é, monotonicamente crescente, com valores limites  $\Phi(-\infty) = 0$  e  $\Phi(+\infty) = 1$ , com apenas uma derivada máxima. Ainda, assume-se que  $\Phi(V)$  é zero até algum potencial de limiar  $V_T$  (possivelmente  $-\infty$ ) e é 1 a partir de algum valor de saturação  $V_S$  (possivelmente  $+\infty$ ). Se  $\Phi$  é a função de Heaviside  $\Theta$ ,  $\Phi(V) = \Theta(V - V_T)$ , temos uma função determinística em tempo discreto. Qualquer outra escolha para  $\Phi(V)$  resulta em um neurônio estocástico. A atividade da rede é medida pela fração (ou densidade)  $\rho[t]$  de neurônios ativos(disparos):

$$\rho[t] = \frac{1}{N} \sum_{j=1}^N X_j[t] \quad (\text{A.0.3})$$

A densidade  $\rho[t]$  pode ser calculada através da densidade de probabilidade  $p[t](V)$  dos potenciais no tempo  $t$ :

$$\rho[t] = \int_{V_T}^{\infty} \Phi(V)p[t](V)dV, \quad (\text{A.0.4})$$

onde  $p[t](V)dV$  é a fração de neurônios com potencial no intervalo  $[V, V + dV]$  no tempo  $t$ .

Neurônios que disparam entre  $t$  e  $t + 1$  têm seu potencial igual ao de reset  $V_R$ . Eles contribuem para  $p[t + 1](V)$  como um pulso de Dirac em  $V_R$ , com amplitude (integral)  $\rho[t]$  dada pela equação acima. Nos passos subsequentes, os potenciais de todos os neurônios evoluem de acordo com a equação A.0.2 . Este processo modifica  $p[t](V)$ .

# APPENDIX B – Teste de Kolmogorov-Smirnov para Duas Amostras

O teste de Kolmogorov-Smirnov (KS) para amostras é um método não paramétrico utilizado para determinar se duas amostras de dados provêm da mesma distribuição contínua. Este teste é baseado na distância máxima entre as funções de distribuição empírica das duas amostras.

Assim, suas hipóteses são:

- $H_0$  : As duas amostras vêm de uma mesma distribuição contínua.
- $H_1$  : As duas amostras não vêm da mesma distribuição contínua.

Sejam  $F_{n_1}(x)$  e  $F_{n_2}(x)$  as funções de distribuição acumulada empírica de duas amostras, onde  $n_1$  e  $n_2$  são os tamanhos das amostras. A estatística de teste KS,  $D_{n_1, n_2}$ , é definida como:

$$D_{n_1, n_2} = \sup_x |F_{n_1}(x) - F_{n_2}(x)| \quad (\text{B.0.1})$$

onde  $\sup_x$  denota o supremo da diferença absoluta entre as duas funções de distribuição empírica em todos os pontos  $x$ .

O valor  $p$  do teste é calculado usando a estatística de teste  $D_{n_1, n_2}$  e a distribuição de Kolmogorov. Rejeitamos a hipótese nula  $H_0$  se o valor  $p$  for menor que o nível de significância  $\alpha$  escolhido (por exemplo, 0.10), indicando que as duas amostras provavelmente vêm de distribuições diferentes.



## APPENDIX C – Algoritmo para determinação de $V_C$ e $\omega$

$N \leftarrow$  total number of agents  $W \leftarrow$  list of wealths sorted in descending order  
 $\epsilon_{min} \leftarrow \infty$   $Nu_{opt} \leftarrow 0$   $V_c \leftarrow 0$   $\omega \leftarrow 0$   $W_{bar} \leftarrow \text{average}(W)$

**for**  $Nu = 2$  to  $N$  **do**

$W_u \leftarrow W[1 : Nu];$ $W_{bar\_u} \leftarrow \text{average}(W_u)$ $\epsilon \leftarrow  (Nu/(N - Nu)) \cdot W_{bar\_u} - W_{bar} $ <b>if</b> $\epsilon < \epsilon_{min}$ <b>then</b> $\quad \epsilon_{min} \leftarrow \epsilon$ $Nu_{opt} \leftarrow Nu$ <b>end</b>	// Top $Nu$ wealths
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------

**end**

$V_c \leftarrow W[Nu_{opt}];$	// Wealth of the $Nu_{opt}$ -th agent
$\omega \leftarrow Nu_{opt}/(N - Nu_{opt});$	// Proportion of rich agents

**return**  $V_c, \omega$





## APPENDIX D – L-BFGS-B

O método Limited-memory BFGS é um algoritmo de otimização da família dos métodos quasi-Newton que aproxima o algoritmo Broyden–Fletcher–Goldfarb–Shanno (BFGS) usando uma quantidade limitada de memória do computador. É um algoritmo popular para estimativa de parâmetros em aprendizado de máquina. O problema-alvo do algoritmo é minimizar uma função escalar diferenciável sobre valores não restritos de um vetor real.

A principal característica do "L-BFGS-B" é que ele permite a inclusão de limites (ou restrições) nas variáveis. Isso significa que você pode especificar um intervalo dentro do qual a solução deve ser encontrada. Por exemplo, se você sabe que um determinado parâmetro deve estar entre 0 e 1, você pode impor esses limites ao algoritmo.

O procedimento do algoritmo pode ser resumido nos seguintes passos:

1. Inicialize  $x_0$  (ponto inicial),  $H_0$  (aproximação inicial da inversa da Hessiana) e defina os limites inferiores e superiores para as variáveis.
2. Para cada iteração  $k$ :
  - a) Calcule o gradiente  $g_k = \nabla f(x_k)$ .
  - b) Determine a direção de busca  $d_k$  usando a aproximação da inversa da Hessiana  $H_k$  e o gradiente  $g_k$ .
  - c) Encontre um passo  $\alpha_k$  que satisfaça as condições de Wolfe e mantenha a solução dentro dos limites.
  - d) Atualize a solução:  $x_{k+1} = x_k + \alpha_k d_k$ .
  - e) Atualize a aproximação da inversa da Hessiana  $H_{k+1}$  usando a fórmula de atualização L-BFGS.
  - f) Verifique a convergência. Se satisfeita, pare o algoritmo.
3. Retorne a solução  $x_{k+1}$ .

Essa abordagem permite que o L-BFGS seja eficiente em termos de memória e computacionalmente viável para problemas de otimização de grande escala, especialmente aqueles encontrados em aplicações de aprendizado de máquina.