

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

LUIZ HENRIQUE DIAS AFONSO

**MODELO LEI DE ESCALA PARA VOTOS BRANCOS E NULOS EM ELEIÇÕES
ESTADUAIS BRASILEIRAS**

CAMPO MOURÃO

2021

LUIZ HENRIQUE DIAS AFONSO

**MODELO LEI DE ESCALA PARA VOTOS BRANCOS E NULOS EM ELEIÇÕES
ESTADUAIS BRASILEIRAS**

Law scale model for white and null votes in Brazilian state elections

Dissertação apresentada como requisito para obtenção do título de Mestre em Inovações Tecnológicas da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientador(a): Magda Cardoso Mantovani.

Coorientador(a): Leandro de Santana Costa.

CAMPO MOURÃO

2021



Esta licença permite remixe, adaptação e criação a partir do trabalho, para fins não comerciais, desde que sejam atribuídos créditos ao(s) autor(es) e que licenciem as novas criações sob termos idênticos. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.



**Ministério da Educação
Universidade Tecnológica Federal do Paraná
Campus Campo Mourão**



LUIZ HENRIQUE DIAS AFONSO

MODELO LEI DE ESCALA PARA VOTOS BRANCOS E NULOS EM ELEIÇÕES ESTADUAIS BRASILEIRAS

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestre Em Inovações Tecnológicas da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Inovações Tecnológicas.

Data de aprovação: 20 de Agosto de 2021

Prof.a Magda Cardoso Mantovani, Doutorado - Universidade Tecnológica Federal do Paraná

Prof Adilandri Mercio Lobeiro, Doutorado - Universidade Tecnológica Federal do Paraná

Prof Denner Serafim Vieira, Doutorado - Universidade Estadual de Maringá (Uem)

Prof Wyrllen Everson De Souza, Doutorado - Universidade Tecnológica Federal do Paraná

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 31/08/2021.

AGRADECIMENTOS

A Deus, por me dar saúde, força e perseverança permitindo assim a conclusão deste trabalho.

A minha esposa e companheira de vida, Érika Janine Maia Afonso, pelo apoio, paciência, conselhos e companheirismo durante esta jornada.

A minha orientadora, Magda Cardoso Mantovani, pelas orientações, ensinamentos, paciência e por ter me acolhido e acreditado em meu potencial.

A banca examinadora, Prof. Denner Serafim Vieira, Prof. Adilandri Mércio e Prof. Wyrllen Everson, pelas contribuições que enriqueceram o trabalho.

Aos meus pais, por terem me dado apoio e condições para chegar até este momento.

Seja quem você for, seja qual for a posição social que você tenha na vida, a mais alta ou a mais baixa social, tenha sempre como meta muita força, muita determinação e sempre faça tudo com muito amor e com muita fé em Deus, que um dia você chega lá. De alguma maneira você chega lá.

(SENNÁ, 2009).

RESUMO

Este estudo foi direcionado a investigar resultados de segundos turnos de eleições para governadores dos estados brasileiros. O objetivo foi compreender, por completo, a opinião dos eleitores num cenário de apenas dois candidatos, considerando que os votos brancos e nulos representam a insatisfação com a opções de candidatos disponíveis, ou seja, estes votos foram considerados uma terceira opção de opinião do eleitor. Inicialmente foram construídas as relações entre o comparecimento e os votos para cada uma das três opções de votos, os dois candidatos e votos brancos e nulos, por cidade, separadamente, todas modeladas por leis de escalas. As relações para os dois candidatos foram bem ajustadas conjuntamente, permitindo obter um modelo de um único parâmetro para esses dados, já as relações entre brancos e nulos foram ajustadas separadamente. As relações para brancos e nulos apresentaram escalas super lineares em 70% das eleições, indicando que maiores eleitorados são mais seletivos. Em seguida, utilizando esses resultados e a conservação da probabilidade sobre o comparecimento entre as três partições de votos, foi obtido um modelo generalizado e robusto de um único parâmetro modelando votos brancos e nulos em função do comparecimento.

Palavras-chave: eleições; sistemas complexos; lei de escala; conservação da probabilidade.

ABSTRACT

This study aimed at investigating the results of second rounds of the elections for governors of Brazilian states. To fully understand the voters' opinions, in a scenario of only two candidates, considering that white and null votes represent dissatisfaction with the options available candidates, these votes were considered the third option of the voters' opinion. Initially, the relations between the attendance and the votes were constructed for each of the three voting options, the two candidates and white and null votes, by city, separately, all modelled by scale laws. The relations for the two candidates were well adjusted together, allowing to obtain a single parameter model for these data, whereas the relations between whites and nulls were adjusted separately. Relationships for whites and nulls showed super linear scales in 70% of the elections, indicating that larger constituencies demand more. Using these results and conserving the probability of the attendance between the three voting partitions, a generalized and robust model of a single parameter was obtained by modelling white and null votes according to the attendance.

Keywords: elections; complex systems; scaling law; probability conservation.

LISTA DE ILUSTRAÇÕES

- Figura 1: Gráfico do número de candidatos $\log_{10}N$ recebendo uma fração de votos $\log_{10}v$, para São Paulo (triângulos) e para o Brasil (círculos). As linhas sólidas representam os ajustes lineares via métodos dos mínimos quadrados para os dados nas regiões de escala e os números inseridos indicam escala..... 19
- Figura 2: Histogramas das frações de votos - Os gráficos acima, em log-log, mostram as distribuições de votos para as eleições 1998 (círculos) e 2002 (triângulos). As linhas sólidas representam os melhores ajustes para as regiões de escala, (a) para deputados estaduais e (b) deputados federais. 20
- Figura 3: Gráfico logarítmico duplo da distribuição de votos para vereadores. Os dados referem-se ao resultado das eleições proporcionais brasileiras realizadas em 2000 considerando as 15 maiores capitais estaduais. A linha sólida é a mais adequada à lei de Zipf. O expoente da escala da lei de potência é $\alpha = 2,63$ 22
- Figura 4: Ajuste via Lei de Zipf generalizada para deputados estaduais - A distribuição de votos para deputados estaduais em log-log. Os círculos são referentes aos dados dos resultados das eleições brasileiras no ano de 1998, considerando os estados com as 15 maiores capitais. A linha contínua é o melhor ajuste para a Lei de Zipf generalizada, Equação (2.1.5): O expoente de escala da lei de potência é $1,03$ 23
- Figura 5: Ajuste via Lei de Zipf generalizada para deputados federais - A distribuição de votos para deputados federais em log-log. Os círculos são referentes aos dados dos resultados das eleições brasileiras no ano de 1998, considerando os estados com as 15 maiores capitais. A linha contínua é o melhor ajuste para a Lei de Zipf generalizada, Equação (2.1.5): O expoente de escala da lei de potência é $1,0$ 24
- Figura 6: (a) Comportamento de escala da distribuição de votos recebidos pelos candidatos. Os dados referem-se às eleições parlamentares italianas em 1972, mas foram obtidos resultados muito semelhantes para a análise de cada eleição da base de dados. O histograma $P(v, Q, N)$ depende apenas da razão $v_0 = N/Q$, então $P(v, Q, N) = P_0(v, v_0)$. (b) A função $P_0(v, v_0)$ mostrado em (a) depende apenas da proporção $v/v_0 = vQ/N$ 26
- Figura 7: Universalidade da função de escala $F(vQ/N)$ para as 5 eleições da base de dados. O ajuste lognormal, realizado na curva polonesa, descreve muito bem os dados. A curva universal é bem reproduzida pelo modelo, onde a dinâmica das opiniões dos eleitores reflete o espalhamento “boca a boca” do eleitorado do partido..... 27
- Figura 8: Histograma em log-log da distribuição dos votos recebidos pelos candidatos em três eleições brasileiras e uma na Finlândia. Além da diferença entre esses países e do intervalo de oito anos no caso do Brasil, todas as curvas seguem a mesma distribuição, uma lei de potência com expoente $\alpha = 1,0$ 29

- Figura 9: Histograma mostrando a distribuição de votos da variável v normalizada de acordo com o número médio de votos v_0 dos candidatos dentro do seu partido. As curvas são ajustadas por diferentes funções. Para o Brasil, a distribuição segue uma queda exponencial para todos os anos, enquanto na Finlândia o melhor ajuste é um log-normal.....29
- Figura 10: Distribuição do desempenho eleitoral dos candidatos em eleições proporcionais com listas abertas, segundo escalonamento do CF. Itália (até 1992), Polónia, Finlândia, Dinamarca e Estónia (após 2002) seguem essencialmente as mesmas regras, o que se reflete no colapso dos dados do painel F. A evolução histórica dos países não parece afetar a forma da distribuição (painéis A a E).....32
- Figura 11: Mesma análise da Figura 10, para Eslovênia, Grécia, Suíça, Brasil e Uruguai. As curvas são razoavelmente estáveis em nível nacional, mas não se comparam bem entre os países e com as curvas universais da Figura 10 (representadas no painel F pela distribuição para as eleições italianas em 1987). É provável que tais discrepâncias se devam às diferentes regras eleitorais desses países em comparação entre si e às examinadas na Figura 10, embora todos adotem listas abertas.32
- Figura 12: Escala CAAMd e listas abertas - Análise para os mesmos países da Figura 10, com escala CAAMd. As curvas são estáveis a nível nacional, mas elas não se assemelham através de todos os países.34
- Figura 13: Escala CAAMd e listas abertas - Análise para os mesmos países da Figura 11, com escala CAAMd. As curvas são estáveis a nível nacional, mas elas não se assemelham através de todos os países34
- Figura 14: Distribuição do desempenho eleitoral dos candidatos em eleições proporcionais com listas semiabertas, usando a escala FC. A nível nacional, as curvas são na sua maioria estáveis, discrepâncias significativas correspondem a alterações nas regras eleitorais, como na Eslováquia (C), República Checa (D) e Estónia (E). A aparente lei de potência da curva holandesa (F) pode ser gerada por um mecanismo mais rico, uma vez que o limite é muito alto (10% em nível nacional) e os eleitores tendem a apoiar os candidatos com base em sua popularidade. A Estônia desde 2002 adotou listas abertas, razão pela qual as distribuições das eleições estonianas após 2002 são ilustradas nas Figs. 10 e 12 (rotulados como Estônia I).....35
- Figura 15: Distribuição do desempenho eleitoral dos candidatos em eleições proporcionais com listas semiabertas, para os mesmos países da Figura 14, agora usando a escala CAAMd35
- Figura 16: Espalhamento dos dados empíricos e do modelo - Gráfico de dispersão do número N de eleitores versus o número n_c de candidatos tomados o logaritmo na base-10 para eleições brasileiras de vereador em 2008. As cruzes vermelhas representam os dados empíricos e os pontos verdes são resultados da simulação obtidos do modelo de Barabási-Albert com $c = 1,12$, $\alpha = 0,196$ e $\beta = 0,32$. Os círculos abertos (preenchidos) são os valores médios, calculados em janelas igualmente espaçadas, para os dados (resultados simulados) e a linha pontilhada e o ajuste linear para os valores médios dos dados empíricos, encontrando $n_c \sim N^{0,36}$. A inserção mostra a mesma comparação para os valores da variância σ_ω^237

- Figura 17: Comportamento médio para vereadores e prefeitos - Valores médios do logaritmo do número de candidatos N_c versus o logaritmo do número de eleitores N , descontado a constante A , para (a) 5 eleições de multi-membros e (b) 6 eleições de um único-membro. As linhas tracejadas representam leis de potências para os valores médios, cujos expoentes médios α são $\alpha = 0,36$ para eleições de multi-membros e $\alpha = 0,18$ para eleições de um único-membro. Os círculos pretos são os dados simulados com o modelo de rede apresentado ao final do deste capítulo, tomando a rede de Barabási-Albert com $\beta = 0,32$ ($\beta = 0,41$) para multi-membros (único-membro); os outros símbolos representam os dados empíricos. ...38
- Figura 18: Valores médios e suas flutuações- O painel superior mostra os valores médios, em log-log, do número N_c de candidatos versus o número N de eleitores descontando a constante A para (a) 5 eleições de multi-membros e (b) 6 eleições de um membro. As linhas tracejadas representam os ajustes lineares para os valores médios com expoente $\alpha = 0,36$ para eleições de multi-membros e $\alpha = 0,18$ para eleições de um único membro. No painel inferior são mostradas as distribuições de probabilidade das flutuações ξ , calculadas pela Equação 2.3.2, sendo que em: (c) para as eleições de (a) e (d) para as eleições de (b). As linhas tracejadas são as funções ajustes gaussianas com média zero e variância um. Os diferentes símbolos representam os dados e os círculos pretos são as previsões do modelo de rede, considerando o modelo Barabási-Albert com $\beta = 0,32$ ($\beta = 0,41$) para a eleição de múltiplos membros (único membro). As outras eleições apresentam o mesmo padrão.....40
- Figura 19: Ilustração do modelo de rede para o processo de candidatura: cada vértice do grafo representa um eleitor e as ligações entre eles são as relações sociais, por exemplo, amizades. Um eleitor pondera sobre sua popularidade (aqui, o número de ligações) na comunidade levando ou não à candidatura. Os nós em vermelho são os mais populares e, conseqüentemente, candidatos potenciais.....41
- Figura 20: Gráfico de dispersão da relação entre $\log_{10} N$ (população de eleitores) e $\log_{10} V$ (número de filiações partidárias) considerando os dados do ano de 2008. Note que a relação é aproximada por uma tendência a lei de potência subjacente à flutuações. Aqui, os pequenos pontos verdes representam os dados empíricos para cada cidade e os círculos são os valores médios desses pontos depois de agrupar os dados em janelas ω igualmente espaçadas em $\log_{10} V$. A linha tracejada é um ajuste linear para esses valores médios e o expoente da lei de potência é mostrado no gráfico. Os resultados simulados da Equação 2.3.5 são representados por pontos cinza, os valores médios nas janelas ω para os dados simulados são representados pelos quadrados cinzentos.42
- Figura 21: Relações médias leis de potência entre $\log_{10} V$ (população de eleitores) e $\log_{10} N$ (número de filiações partidárias). Observe que foram analisados todos os anos em do banco de dados e os valores de α (as inclinações dessas curvas) são quase constantes no tempo. O valor médio de α é mostrado no gráfico e a linha tracejada representa o ajuste linear com este expoente. Os quadrados cinzas representam os resultados simulados da Equação 2.3.7.43

- Figura 22: Flutuações em torno das tendências à lei de potência. (a) Variância σ^2 dos valores de $\log_{10} N$ versus $\log_{10} V$ calculados em janelas ω igualmente espaçadas em $\log_{10} V$. Note que o valor de σ^2 sobre a população de eleitores é quase constante. A linha tracejada horizontal é o valor médio de σ^2 sobre $\log_{10} V$ e ao longo dos anos ($0,027 \pm 0,002$). (b) Distribuição dos resíduos normalizados (ξ) em torno das tendências à lei de potência. A linha tracejada é a distribuição gaussiana com média zero e variância unitária. Em ambos os gráficos, os quadrados em cinza representam os resultados simulados da Equação 2.3.7.44
- Figura 23: A hierarquia das posições políticas. Valores médios dos expoentes da lei de potência α para cada posição política. Os expoentes das relações entre o número de candidatos a vereador e prefeito e a população de eleitores foram obtidos Mantovani (2013). Notamos que a média de α é menor para as posições políticas mais influentes. As barras de erro são desvios padrão do valor médio de α45
- Figura 24: Gráfico de dispersão das variáveis $\log_{10} V$ versus $\log_{10} C$ para as eleições para prefeito de 2000, 2004, 2008, 2012 e 2016, onde V e C representa os números dos eleitores e dos candidatos, respectivamente, em cada cidade. Os pontos verdes claros representam os dados dos homens e os pontos azuis claros os dados das mulheres. Os pontos verdes escuros (para homens) e azuis escuros (para mulheres) representam os valores médios das variáveis $\log_{10} V$ versus os valores médios da variável $\log_{10} C$ em janelas ω igualmente espaçadas em $\log_{10} V$. As retas pretas contínuas representam os ajustes lineares para os valores médios dos candidatos homens e as retas pretas tracejadas representam os ajustes lineares para os valores médios das candidatas mulheres, sendo que R_M^2 e R_F^2 na inserção representam os coeficientes de determinação para os ajustes masculinos e femininos respectivamente. As barras de erro são desvios padrão do valor médio em cada janela ω46
- Figura 25: A evolução dos interceptos A e dos expoentes α das tendências médias das leis de potência para as candidaturas a vereador, apresentadas na Figura 24. Os símbolos verdes referem-se aos homens e os símbolos azuis escuros às mulheres. As barras de erro são os erros padrão para esses ajustes lineares.47
- Figura 26: Valores médios da razão entre o número de candidata feminino C_F e o número de candidato masculino C_M versus $(\log_{10} V)$, em janelas ω igualmente espaçadas de $\log_{10} V$ para eleições de vereador, onde V representa o número total de eleitores. As barras de erro são os desvios padrão do valor médio das razões nas janelas ω48
- Figura 27: Gráfico de dispersão das variáveis $\log_{10} V$ versus $\log_{10} C$ para as eleições para prefeito de 2000, 2004, 2008, 2012 e 2016, onde V e C representam o número de eleitores e o número de candidatos, respectivamente, em cada cidade. Os pontos verdes claros representam os dados dos homens e os pontos azuis claros os dados das mulheres. Os pontos verdes escuros (para homens) e azuis escuros (para mulheres) representam os valores médios das variáveis $\log_{10} V$ e os respectivos valores médios da variável $\log_{10} C$ em janelas ω igualmente espaçadas em $\log_{10} V$. As retas pretas contínuas representam os ajustes lineares para os valores médios

dos candidatos homens e as retas pretas tracejadas representam os ajustes lineares para os valores médios das candidatas mulheres, sendo que R_M^2 e R_F^2 na inserção representam os coeficientes de determinação para os ajustes masculinos e femininos respectivamente. As barras de erro são desvios padrão do valor médio em cada janela ω49

Figura 28: A evolução dos interceptos A e dos expoentes α das tendências médias das leis de potência para as para candidatura a prefeito, apresentadas na Figura 27. Os símbolos verdes referem-se aos homens e os símbolos azuis escuros às mulheres. As barras de erro são os erros padrão para esses ajustes lineares.49

Figura 29: Relação de Escala entre o número de votos e as despesas da campanha. Os círculos roxos claros mostram a relação entre o número de votos v e as despesas declaradas da campanha, m , de cada candidato a deputado estadual e deputado federal (DF) nas eleições de 2014 para os quatro maiores estados do Brasil: São Paulo (a e b), Rio de Janeiro (c e d), Minas Gerais (e e f) e Bahia (g e h). Dentro de cada painel, os dados para candidatos eleitos são destacados em círculos roxos escuros. Para ver as nuances de correlação, traçou-se em uma relação normalizada para deputados estaduais (i) e deputados federais (j) para os oitos maiores estados do Brasil. Os símbolos dispostos na legenda representam a relação normalizada $\langle v \rangle \Delta m / m$, em que primeiro calculou-se o número médio de votos em janelas log-espaçadas ao longo de m . Assumindo-se uma correlação linear, a constante multiplicativa é $\Delta m = M / n$. A normalização fornece uma observação direta da não linearidade na dependência dos votos sobre o dinheiro. Vê-se um comportamento sublinear global, em que os candidatos mais ricos exibem uma fração menor de votos por dinheiro.51

Figura 30: Esboço do Modelo para o preço do voto - Dois candidatos competem por votos em uma rede social com indivíduos indecisos (branco). (A) O candidato azul tem inicialmente um orçamento de 4 unidades monetárias, enquanto a campanha do candidato vermelho tem um orçamento de 2 unidades monetárias. Ambos os candidatos têm um eleitor decidido no início do processo. O eleitor azul na parte superior da rede e o vermelho na parte inferior da rede. As regiões destacadas incluem o eleitor inicial de cada campanha e os conhecidos de cada eleitor inicial. (B) Um eleitor indeciso dentro de cada área operacional é escolhido aleatoriamente, tornando-se então eleitores decididos. Assim, cada campanha de orçamento diminui no valor de $\Delta m = \$$. Como consequência de novos eleitores decididos, as áreas operacionais crescem, e como dois conhecidos passam a eleitores de candidatos diferentes, as áreas operacionais das campanhas agora se sobrepõem. A região representada em amarelo é onde as ambas as campanhas podem atuar. (C) Enquanto a campanha do candidato vermelho escolheu um eleitor indeciso, aumentando sua área operacional, a campanha azul gasta dinheiro em um eleitor decidido. (D) Ao final do processo de campanha, quando todas as campanhas ficam sem recursos financeiros, a campanha com maior orçamento inicial encerra o processo com mais eleitores aderidos, mas sua campanha é menos eficiente, resultando em uma não economia de escala.52

- Figura 31: Modelo para escala não linear. Em (A) e (B) é mostrado a eleição paulista para deputados estaduais e federais em 2014, respectivamente. Cada círculo roxo claro corresponde a um candidato e os círculos roxos escuros são os números médios de votos em janelas com espaçamento logarítmico ao longo de m . Em (C) e (D) foi realizado o mesmo processo de normalização como na Figura 29, mas agora os Δm foram estimados usando a Equação 2.4.3. Cada linha sólida mostra a solução do modelo. Apesar de sua simplicidade, o modelo apresenta todos os regimes não lineares observados nos dados, o que mostra que a ineficiência dos candidatos com mais recursos é principalmente devido à competição. ...55
- Figura 32: Escala urbana nas eleições presidenciais nos Estados Unidos e para o referendo da União Europeia no Reino Unido - Os gráficos em log-log mostram, em (A) os votos para Republicanos (vermelho) e Democratas (azul) em função do comparecimento dos eleitores para as ⁹¹² maiores áreas estatísticas metropolitanas e micropolíticas em 2016. Em (B) os votos para Permanecer (azul) e Sair (vermelho) em função do comparecimento dos eleitores para os distritos eleitorais do Reino Unido. β é a inclinação e R^2 o coeficiente de determinação do ajuste linear.57
- Figura 33: Evolução histórica dos expoentes - Escala dos expoentes para os Republicanos (vermelho) e Democratas (azul) com barras de erro para as 18 eleições presidenciais de 1948 à 2016.57
- Figura 34: Correlação dos parâmetros de escala nas eleições dos Estados Unidos. Em (A), relação dos expoentes de escala dos Republicanos em função dos Democratas para as 18 eleições presidenciais dos EUA de 1948 a 2016 (pontos) e a curva teórica (linha vermelha) decorrente da conservação da probabilidade. Em (B), interceptos das relações $\log_{10} Y_0$ em função dos expoentes β para Republicanos (vermelho) e Democratas (azul) para eleições presidenciais no período 1948 a 2016. A reta em preto representa o ajuste linear dado pela Equação 2.5.2 e na inserção os parâmetros e coeficiente de determinação do ajuste.58
- Figura 35: Gráfico de dispersão da relação entre as variáveis: número de pessoas que compareceram para votar C versus o número de votos para o primeiro colocado, versus número de votos para o segundo colocado, versus número de votos brancos e nulos, todos os montantes de votos representados por V . Dados dos segundos turnos das eleições para governador, em todos os estados onde ocorreu segundo turno, em 2018. Cada painel representa um estado e é identificado pela sigla dos estados.69
- Figura 36: Expoentes e Interceptos dos ajustes mostrados na Figura 35. O painel superior mostra os expoentes relacionados com os 13 estados onde ocorreu o segundo turno para as eleições de governadores no ano de 2018, o painel inferior mostra, da mesma forma, os interceptos correspondentes, Equação 3.2.2.1.73
- Figura 37: Expoentes dos ajustes lineares para as três partições de votos: O painel acima mostra, para cada estado brasileiro onde ocorreu a eleição, os expoentes dos ajustes lineares obtidos para as três relações $\log_{10} C$ versus $\log_{10} V$, em que C é o comparecimento e V os votos: para o primeiro colocado, o segundo colocado e votos brancos e nulos, por cidade, para os 7 anos de eleições da base de dados. Equação 3.2.2.1.74

- Figura 38: Expoentes dos ajustes lineares dos votos brancos e nulos: O painel acima mostra, para cada estado brasileiro onde ocorreu a eleição, os expoentes dos ajustes lineares obtidos para as relações $\log_{10}C$ versus $\log_{10}V$, em que C é o comparecimento e V os votos brancos e nulos, por cidade, para os 7 anos de eleições da base de dados. A reta em preto pontilhada em cada gráfico é a reta horizontal constante em 1, possibilitando visualizar os expoentes acima e abaixo dela, ou seja, as escalas super lineares e sublineares, respectivamente.75
- Figura 39: Gráfico dos expoentes versus interceptos das leis de escala para as três partições de votos. Os pontos no gráfico acima representam a relação entre o intercepto A_k e o expoente α_k para as três partições de votos, votos para o primeiro colocado, para o segundo colocado e brancos e nulos, nas 78 eleições da base de dados. A equação e o coeficiente de determinação para o ajuste dos parâmetros obtido para as partições de votos do primeiro, juntamente, com os dos segundo colocado, reta contínua em preto, é mostrada no inserte, onde o índice p e s representa essas partições, respectivamente. São também mostrados para os ajustes dos parâmetros referentes aos brancos e nulos, reta em preto tracejada, onde o índice bn está relacionado aos votos brancos e nulos. 76
- Figura 40: Flutuações em torno do modelo linear para os dados de 2002. O gráfico acima apresenta as densidades normalizadas da variável ξ as 13 eleições brasileiras de governadores de estados onde ocorreu segundo turno no ano de 2002. A curva tracejada em preto representa a distribuição normal de média $\mu = 0$ e variância $\sigma^2 = 1$79
- Figura 41: Comparação entre os dados reais e os simulados pelo modelo lei de potência. O painel acima mostra as relações $\log_{10}C$ versus $\log_{10}V$ em que C representa o comparecimento e V os brancos e nulos, ambos por cidade. Os círculos em vermelho representam os dados empíricos das eleições brasileiras de governadores de 2002 para cada estado onde ocorreu segundo turno, os círculos em preto correspondem aos dados simulados via modelo lei de potência (3.2.2.9).80
- Gráfico 1: Estados brasileiros x Número de vezes em que ocorreu.....65
- Gráfico 2: Coeficiente de determinação R^2 para todos os ajustes lineares relacionando as variáveis C , comparecimento, e V , votos para o primeiro colocado ou votos para o segundo colocado ou votos brancos e nulos, obtidos para as 78 eleições da base de dados. Eixo (x), representa os estados brasileiros. Eixo(Y) representa o coeficiente de determinação R^272

LISTA DE TABELAS

- Tabela 1 - Estados brasileiros, siglas e região64
- Tabela 2 - Anos e estados que tiveram segundo turno nas eleições para governador. As cédulas marcadas com X representam os estados e anos onde ocorreu segundo turno em eleições para governadores dos estados brasileiros.64
- Tabela 3 – Tabela do total de votos e das porcentagens de votos para o primeiro colocado, o segundo colocado e votos brancos e nulos em cada estado onde ocorreu segundo turno nas eleições para governadores nas 7 eleições da base de dados, 1994, 1998, 2002, 2006, 2010, 2014 e 2018. Aqui as variáveis Ano, Sigla UF, Total_Votos, %Votos_Primeiro, %Votos_Segundo e %Votos_BN, representam, o ano que ocorreu a eleição, a sigla no estado onde ocorreu, o total de votos no ano e estado de acordo com a posição, a porcentagem de votos para o candidato que ficou em primeiro lugar, a porcentagem de votos do candidato que ficou segundo lugar e a porcentagem de votos brancos e nulos, respectivamente.67
- Tabela 4 - Tabela dos coeficientes de determinação R^2 pela Equação 3.2.2.1: para todos os ajustes lineares relacionando as variáveis C , comparecimento, e V , votos para o primeiro colocado ou votos para o segundo colocado ou votos brancos e nulos, obtidos para as 78 eleições da base de dados. Aqui Ano, Sigla da UF, R^2 Primeiro, R^2 Segundo e R^2 Brancos e Nulos, representam, o ano em que ocorreu a eleição, a sigla referente ao estado onde houve o segundo turno, o coeficiente de determinação para o ajuste linear relacionando os votos do primeiro colocado, o coeficiente de determinação para o ajuste linear relacionando os votos do segundo colocado, o coeficiente de determinação para o ajuste linear relacionando os votos brancos e nulos, respectivamente.....69

SUMÁRIO

1	INTRODUÇÃO	13
2	ANÁLISES ESTATÍSTICAS DE PROCESSOS ELEIORAIS VIA SISTEMAS COMPLEXOS	17
2.1	Distribuição de votos em eleições proporcionais	18
2.2	Distribuição de votos em eleições proporcionais levando em conta a relevância do partido	24
2.3	Leis de potência modelando processos de candidaturas	36
2.4	O preço do voto em eleições proporcionais brasileiras	50
2.5	Leis de escala universal em eleições presidenciais dos Estados Unidos	55
3	MODELO LEI DE ESCALA PARA VOTOS BRANCOS E NULOS EM ELEIÇÕES ESTADUAIS BRASILEIRAS	60
3.1	Apresentação dos dados	62
3.2	Análise dos dados	68
3.2.1	Obtenção das leis de escala	68
3.2.2	Modelo analítico para os votos brancos e nulos	76
3.2.3	Relação entre os expoentes	80
3.3	Resultados da análise	82
3.4	Conclusões finais	84
	REFERÊNCIAS	85
	APÊNDICE A - Lei de potência ou escala	87
	APÊNDICE B - Lei de Zipf	89
	APÊNDICE C - Distribuição Gauciana e Distribuição Lognormal	92
	APÊNDICE D – Coeficiente de determinação	95
	APÊNDICE E – Rede de Barabási-Albert	97
	APÊNDICE F – Série de Taylor	99

1 INTRODUÇÃO

Apesar das eleições serem atualmente um tópico efervescente, ainda há muitas questões importantes relacionadas a este processo social complexo que continuam sem resposta. Não há exagero em afirmar que os cientistas estão apenas começando a extrair conhecimento quantitativo e, para descobrir os efeitos de seus resultados no processo eleitoral.

Dados eleitorais, em geral, representam sistemas em grande escala, abordados no contexto de sistemas complexos utilizando ferramentas estatísticas. Neste sentido, vários estudos têm apresentado modelos de comportamento para essas dinâmicas, sendo que a lei de potência tem-se mostrado muito presente nesses fenômenos.

No intuito de ampliar as discussões referentes a essa temática, no primeiro capítulo desta dissertação apresenta-se alguns estudos que exemplificam a modelagem de sistemas eleitorais. Para tanto, inicia-se com a pesquisa de Costa Filho *et al.* (1999; 2002) na qual foram observadas a parte central da distribuição de votos entre os candidatos e como resultado obteve-se um comportamento de lei de potência em eleições proporcionais brasileiras, revelando que o sucesso dos candidatos é determinado por um processo multiplicativo. Nesta mesma perspectiva, Lyra *et al.* (2003) ao modelar os mesmos dados, propôs um ajuste para toda a extensão das distribuições via uma lei de Zipf generalizada.

Em seguida, apresenta-se o trabalho de Castellano *et al.* (2007) que analisou as distribuições de votos em eleições proporcionais e reescalou os votos levando em consideração a relevância do partido político obtendo um comportamento tipo log-normal para os dados de três países, sugerindo um comportamento universal. No entanto, a pesquisa de Araripe e Costa filho (2009) em que os autores aplicaram o mesmo procedimento para dados eleitorais da Finlândia e do Brasil obteve o mesmo comportamento para a Finlândia e uma distribuição exponencial para os dados eleitorais do Brasil.

Neste sentido, direcionados a compreender melhor as distribuições de votos levando em conta o peso do partido, Chatterjee *et al.* (2012) analisaram dados de 15 países com sistemas de eleições proporcionais utilizando três regras diferentes (listas fechadas, abertas e semi-abertas) e como resultado indicaram que o ajuste via log-normal se confirmou para um grupo de 5 países com listas abertas.

Dando continuidade a essas discussões, ao considerar que os estudos de processos eleitorais não se concentram apenas nos resultados de eleições, às várias fases dos processos eleitorais devem ser investigadas uma vez que todas elas influenciam no resultado. Nessa direção, questões relacionadas as motivações que levam as pessoas a se envolver diretamente no processo eleitoral, como se candidatar a um cargo político ou ser membro de um partido político tem sido investigada por pesquisadores da área.

A pesquisa de Mantovani *et al.* (2011;2013) evidenciou que as relações entre o número de candidatos (prefeito e vereador) para eleições de 5 países, e número de filiados, para partidos brasileiros, versus população de eleitores são descritos por leis de potência sublineares em que os expoentes de escala obedecem a uma hierarquia determinada pelo cargo mais influente em ordem decrescente, ou seja, mostrando que a dificuldade em se candidatar aumenta com o tamanho do eleitorado e de acordo com a posição do cargo.

Ainda nessa fase do processo eleitoral, Mantovani *et al.* (2019), levando em consideração a candidatura para vereadores e prefeitos nas cidades brasileiras, por gênero, observaram que ambos os gêneros apresentaram comportamentos leis de potências, porém para a candidatura a prefeito uma acentuada desvantagem na candidatura feminina foi detectada, diferentemente do comportamento apresentado no processo de candidatura para vereadores em que o comportamento feminino é similar ao masculino observando uma rápida aproximação em função da lei que estabelece uma porcentagem mínima de 30% de candidaturas femininas por partido. No entanto, vale reforçar que a aparente pequena desvantagem no caso citado deve ser cuidadosamente observada devido ao uso das variáveis em logaritmo.

Outro enfoque dos processos eleitorais foi considerado por Melo *et al.* (2018) que analisaram a maneira com que as despesas eleitorais influenciam nos resultados das eleições em que observaram que os principais candidatos gastam mais dinheiro por voto revelando uma sublinearidade e uma deseconomia de escala.

Por fim, como último exemplo de estudos eleitorais foi apresentado o trabalho de Bokányi, Zoltán e Vattay (2018) na qual os autores analisaram resultados de eleições presidenciais nos Estados Unidos ocorridas no período de 1948-2016 e no referendo da União Europeia ocorrido em 2016. Constataram curvas de escalas universais dependente de um único parâmetro, e uma das partições de votos, primeiro ou segundo colocados nas eleições, ou permanece ou sai da União

Europeia, sempre mostram escala superlinear e conduzem o processo, enquanto a outra partição apresenta expoente sublinear que é meramente a consequência da conservação da probabilidade. Uma análise similar para dados de eleições para governadores dos estados brasileiros foi obtida em Souza (2019) que relatou uma maior chance de vencer as eleições quando se investe maiores esforços de campanha em menores eleitorados, devido ao fato de que em mais de 60% das eleições os candidatos segundo colocados apresentam escalas super lineares enquanto o primeiro colocado apresenta escalas sublineares.

Ao analisar as discussões promovidas, levanta-se uma questão: mesmo num cenário “bipartido”, ou seja, de apenas dois candidatos a opinião dos eleitos está mesmo restrita a essas duas possibilidades? Sabemos que entre o montante de votos encontram-se os votos brancos e nulos, e aí pergunta-se novamente: eles representam apenas erros do eleitor? Ou podem representar eleitores insatisfeitos com a configuração de candidatos disponível?

Depois de levantar essas questões foram observadas as quantidades de votos brancos e nulos em algumas eleições brasileiras e constatado uma proporção considerável deles, despertando o interesse em compreender melhor a dinâmica dessa totalidade de votos quando comparado aos votos dos candidatos e ao comparecimento dos eleitores. Direcionados a investigar esse montante de votos, foi proposta nesta pesquisa uma análise estatística de dados de eleições em segundos turnos (para manter o cenário bipartidário) para governadores dos estados brasileiros ocorridas no período 1994 – 2018, totalizando 78 eleições, dados disponibilizados pelo Supremo Tribunal Eleitoral – TSE. Numa primeira fase foram construídas as relações entre o comparecimento e votos para cada uma das três partições, votos para os candidatos primeiro e segundo colocados e votos brancos e nulos, por cidade, as quais apresentaram comportamento leis de escalas para as 78 eleições e para as três partições de votos, contabilizando um total de 234 relações leis de potências. Em seguida, foi observado que 70% das relações obtidas para a partição de brancos e nulos apresentaram escalas super lineares, indicando que a medida que aumenta o eleitorado aumenta a proporção desses votos, sugerindo que em grandes eleitorados tem-se uma maior rejeição dos candidatos, ou seja, são mais seletivos. Posteriormente foram construídas as relações entre interceptos e expoentes para o montante de leis de escala, obtendo um ajuste único e consistente para as partições de votos do primeiro e do segundo colocados, possibilitando um

modelo dependendo de um único parâmetro, modelando essas duas partições de votos em função do comparecimento. No entanto, o ajuste para os parâmetros das escalas de brancos e nulos não é pouco consistente.

Utilizado todos os resultados obtidos anteriormente, as leis de escala entre as três partições, a relação entre os parâmetros do primeiro e do segundo colocado e a conservação de probabilidade sobre o comparecimento e entre as três partições, foi construído um modelo universal e robusto relacionando votos brancos e nulos ao comparecimento. Os resultados obtidos aqui permitem compreender melhor a dinâmica de eleitores insatisfeitos com disponibilidade de candidatos e podem ajudar a planejar melhor as fases dos processos eleitorais, além de prever resultados futuros.

2 ANÁLISES ESTATÍSTICAS DE PROCESSOS ELEITORAIS VIA SISTEMAS COMPLEXOS

Eleições são processos fundamentais em sociedades democráticas, pois representam instrumento de mudanças determinantes para a comunidade. Pesquisadores tem direcionado esforços para compreender os processos eleitorais via sistemas complexos, dados que esses são fenômenos de grande escala que apresentam comportamentos regulares coletivos emergentes. Usando ferramentas estatísticas é possível observar padrões e construir modelos que permitam observar dinâmicas de opinião que ocorrem naturalmente nesses processos. Tais resultados podem ajudar a compreender quais aspectos sociais, econômicos e políticos são relevantes para influenciar, ou mesmo determinar, o resultado de uma eleição, e assim possibilitar corrigir problemas e planejar comportamentos futuros. Neste capítulo serão apresentadas algumas análises de processos eleitorais realizadas no contexto de sistemas complexos, assim como a investigação realizada nesta pesquisa apresentada no próximo capítulo. Como análises iniciais, os trabalhos Costa Filho *et al.* (1999; 2002) observaram um comportamento lei de potência invariante por escala com expoente aproximadamente 1 para a parte central da distribuição de votos em eleições proporcionais brasileiras assumindo que o sucesso dos candidatos é determinado por um processo multiplicativo. Na sequência, para esses mesmos dados, Lyra *et al.* (2003) propõe um ajuste por uma generalização da Lei de Zipf, agora para toda a extensão dos dados. Reescalando a distribuição de votos levando em conta os votos do partido Fortunato e Castellano (2007) ajustou eleições proporcionais de 3 países pelo modelo log-normal, no entanto usando o mesmo procedimento Araripe e Costa Filho (2009) verificaram que eleições brasileiras são ajusta por uma distribuição exponencial. Ainda em uma investigação similar, mas para uma base de dado bem mais ampla, agora para eleições de 15 países, Chatterjee *et al.* (2013) mostram que para um grupo de 5 países com regras eleitorais semelhantes o comportamento é universal e o modelo log-normal se mantém. Processos de candidatura e filiação partidária em eleições brasileiras também apresentam comportamento lei de potência, e com os expoentes seguindo uma hierarquia, em que a posição mais influente assume o maior expoente, como pode ser visto em Mantovani *et al.* (2011; 2013). Quando a análise leva em conta o gênero, Cardoso *et al.* (2019) identifica uma grande desvantagem para as mulheres,

o que pode ser vista através dos expoentes do modelo em eleições para prefeito. Leis de potência também modelam o custo do voto de um candidato em eleições proporcionais brasileiras Araripe *et al.* (2013), e os candidatos mais votados gastam mais dinheiro por voto do que os candidatos com menor sucesso e mais pobres, uma relação que revela uma deseconomia de escala que decorre da competição entre candidatos devido a gastos de campanha ineficientes. Como último exemplo, Bokányi, Zoltán e Vattay (2018) foi o trabalho que direcionou diretamente as investigações desse estudo, contemplado os dados de eleições presidenciais dos US de 1948–2016 os autores mostraram que as curvas de escala modelando as relações entre comparecimento e votos para o primeiro e para o segundo colocado são universais dependendo de um único parâmetro.

2.1 Distribuição de votos em eleições proporcionais

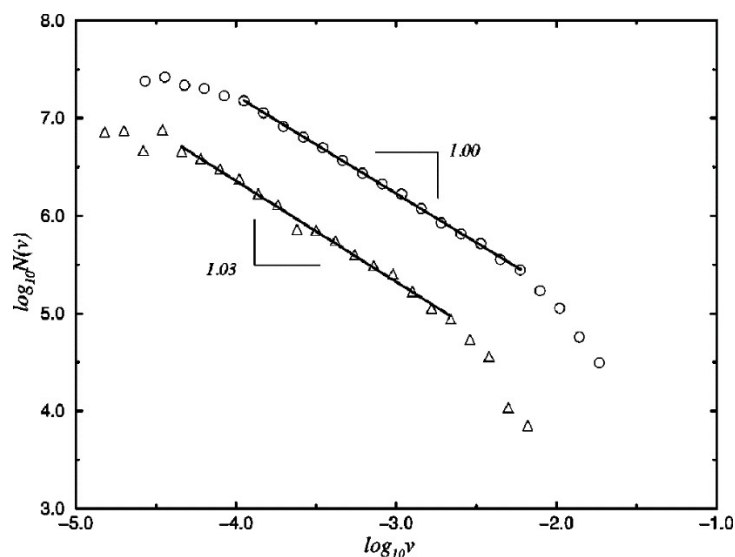
Ao definir o princípio básico do sistema proporcional Mantovani (2013) apresenta que:

[...] todos os eleitores merecem representação e todos os grupos políticos merecem ser representados nas legislaturas na proporção de sua força no eleitorado. Para alcançar essa representação “justa”, cada partido define internamente sua lista de candidatos (no Brasil, candidatos a deputado estadual federal e vereadores). Após a contagem dos votos, as cadeiras a serem ocupadas em dado posto serão atribuídos aos partidos proporcionalmente ao número de votos que o partido recebeu (MANTOVANI, 2013, p. 96).

Iniciando pesquisas de dados eleitorais no contexto de sistemas complexos, usando dados das eleições proporcionais brasileiras de 1998, em Costa Filho *et al.* (1999; 2002), as distribuições de votos dos candidatos para todo o país foram ajustadas por um modelo tipo lei de potência. Considerando os resultados para os cargos de deputados estaduais e tomando N o número de candidatos que receberam a fração de votos v (números de votos do candidato normalizado pelo número total de eleitores). Na Figura 1, é mostrado o histograma do número de candidatos, em logaritmo de base 10, $\log_{10}N$, que receberam uma certa fração de votos v , também tomado o logaritmo na base 10, $\log_{10}v$, para o estado de São

Paulo, onde o número de eleitores e candidatos são os maiores do Brasil. Nessa eleição, São Paulo tinha 1.260 candidatos a deputado estadual e 23 321 034 eleitores, o que representa 21,99% dos eleitores de todo o país. Foram também observados comportamentos semelhantes em outros estados brasileiros com um número de eleitores e candidatos suficientemente grande para possibilitar o mesmo tratamento estatístico. O número total de candidatos a deputado estadual no país foi de 10.535. Como pode ser visto a parte central das distribuições segue modelo $N(v) \propto v^{-\alpha}$, em que $\alpha = 1 \pm 0.02$.

Figura 1: Gráfico do número de candidatos $\log_{10}N$ recebendo uma fração de votos $\log_{10}v$, para São Paulo (triângulos) e para o Brasil (círculos). As linhas sólidas representam os ajustes lineares via métodos dos mínimos quadrados para os dados nas regiões de escala e os números inseridos indicam escala.



Fonte: Costa Filho *et al.* (1999, p.1067).

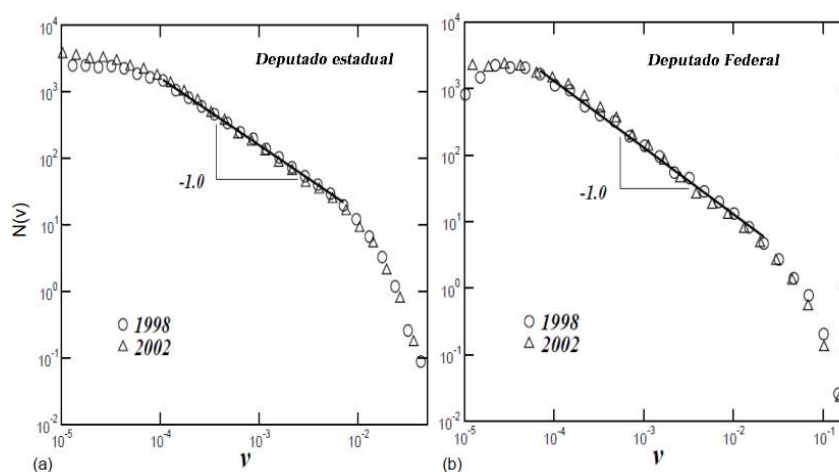
Os autores dessa análise interpretaram o ajuste $1/v$ como resultado descrito por um processo multiplicativo, isto é, o sucesso dos candidatos que é determinado por $N(v)$ pode ser visto como um “grande processo”, dependendo do sucesso da competição de um número n de subprocessos independentes. Em que n é o número de fatores intrinsecamente relacionados aos atributos e / ou habilidades dos candidatos. Assim, associando a cada candidato a probabilidade p_i de realizar o subprocesso i entre os eleitores e então considerando $v = p_1 p_2 \dots p_n$, assumindo que p_i são variáveis aleatórias independentes, para n suficientemente grande, pelo

Teorema do Limite Central a distribuição de v deve ser aproximadamente log-normal. E ainda, se a dispersão da distribuição log-normal for suficientemente grande, pode-se observar um tipo de distribuição $1/v$, ao longo dos valores das variáveis. Com isso os autores concluem que sendo o eleitorado o maior do mundo sujeito ao voto obrigatório juntamente com a divulgação de informações, principalmente por meio da televisão, isso pode ter um efeito significativo no sucesso dos candidatos.

Aplicando os mesmos procedimentos estatísticos de Costa Filho *et al.* (1999; 2002), que utilizou dados das eleições proporcionais ocorridas no Brasil em 1998, obteve o mesmo comportamento lei de potência invariante por escala, para as eleições de 2002 em que o total de 115.253.432 eleitores dos 27 estados do Brasil, escolheram 4.210 candidatos a deputados federais e 11.717 a deputados estaduais, sendo que todas as urnas foram eletrônicas.

Os resultados mostrados na Figura 2, indicam que através das análises das eleições brasileiras de 1998 e 2002 a distribuição dos votos é altamente reprodutível. Apresentando característica de um fenômeno "livre de escala" em que o mecanismo de tomada de decisão governante pode ser modelado em termos de um processo.

Figura 2: Histogramas das frações de votos - Os gráficos acima, em log-log, mostram as distribuições de votos para as eleições 1998 (círculos) e 2002 (triângulos). As linhas sólidas representam os melhores ajustes para as regiões de escala, (a) para deputados estaduais e (b) deputados federais.



Fonte: Costa Filho *et al.* (2002, p.700).

Um candidato ao Congresso Nacional ou ao Legislativo Estadual provavelmente será conhecido do eleitor pela mídia e que quando os instrumentos

de informação estão ausentes ou são ignorados, a decisão de voto é determinada pela comunicação direto com parentes e amigos, como por exemplo, eleições municipais onde existe uma proximidade maior entre eleitores e candidatos criando uma estrutura diferente da encontrada nas eleições estaduais e federais. No entanto os autores realizaram a mesma análise para eleições municipais, mas não obtiveram comportamento claro de escala.

Em Lyra *et al.* (2003) é apresentado um ajuste por uma Lei de Zipf generalizada, anteriormente proposta dentro de uma estrutura mecânica estatística não-extensiva, para os mesmos dados dos trabalhos anteriores, eleições de 1998 e 2002 para deputados federais. Esse trabalho incluiu na análise dados das eleições municipais para vereadores de 2000. Os autores verificaram que o mesmo comportamento de escala para as distribuições de votos para deputados federais e deputados estaduais dentro de uma precisão estatística, mas bastante distinto do obtido para o processo de votação local para vereadores.

Considerando os resultados das 15 maiores cidades para as eleições para prefeito e vereadores, o número médio de votos por candidato é 500, significativamente maior do que para deputados estaduais e federais. Normalizamos, para cada cidade, os votos de cada candidato pelo número total de eleitores e construindo os histogramas para um dado número de candidatos, N , recebendo uma fração de votos, v . Os histogramas para cada cidade apresentaram uma forma padrão dentro de uma previsão estatística. A Figura 3 mostra os dados agrupados em um único histograma em log-log. O ajuste dos pontos mostrado na Figura 3 é construído através da Lei de Zipf expressa por

$$N(v) = \frac{A}{(1 + Cv)^\alpha}, \quad (2.1.1)$$

sendo A uma constante da normalização, C governa a interseção entre o platô inicial e o regime de lei de potência caracterizado pelo expoente α . A lei acima foi deduzida do formalismo termodinâmico não-extensivo proposto por Tsallis (1988), onde a função exponencial (a qual aparece no peso estatístico de Boltzmann-Gibbs) é substituída por uma generalização da exponencial q , com desvio associado ao grau de não extensividade e reflete o grau de correlação do sistema. Tal exponencial que, $exp_q(t)$, deve satisfazer a equação

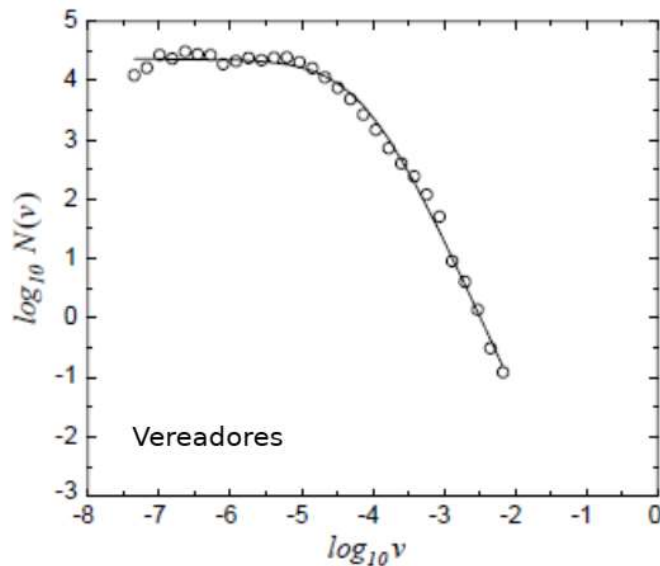
$$\frac{dW(t)}{dt} = \lambda_q W^q. \quad (2.1.2)$$

Tomando $\lambda_q = 1$ e $W(0) = 1$, a solução da equação será $W(t) = \exp_q(t) = [1 + (1 - q)t]^{1/(1-q)}$, para $q = 1$ tem-se a exponencial usual. Fazendo $W = N$ e $t = v$ na equação (2.1.2)

$$N(v) = \frac{N(0)}{[1 + (q - 1)\lambda'_q v]^{1/(q-1)}}, \quad (2.1.3)$$

onde $\lambda'_q = \lambda_q/N(0)^{1-q}$, para $q \rightarrow 1$ recupera-se a exponencial usual. O melhor ajuste de $N(v)$ para eleições de vereadores pela equação (2.2.3) pode ser visto na Figura 3 como uma linha sólida para $\alpha = 1/(q - 1) = 2,63$ ($q = 1,38$).

Figura 3: Gráfico logarítmico duplo da distribuição de votos para vereadores. Os dados referem-se ao resultado das eleições proporcionais brasileiras realizadas em 2000 considerando as 15 maiores capitais estaduais. A linha sólida é a mais adequada à lei de Zipf. O expoente da escala da lei de potência é $\alpha = 2,63$.

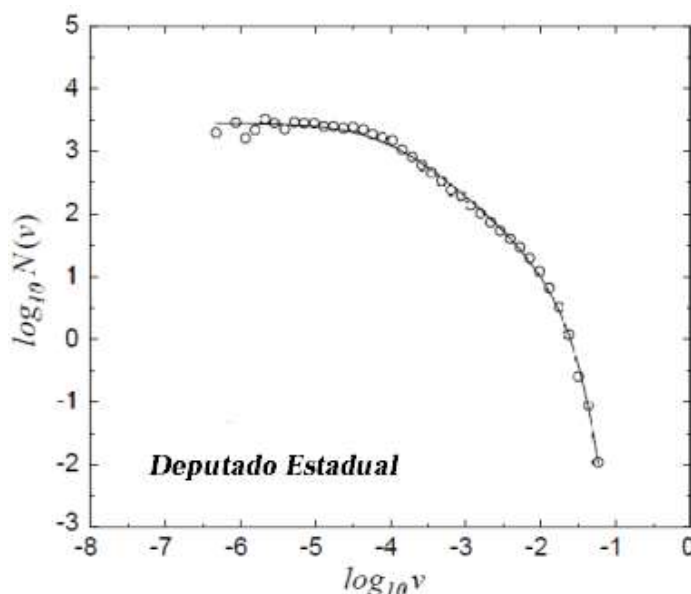


Fonte: Lyra et al. (2003, p.133).

A Figura 4 mostra $N(v)$ calculada para os dados da eleição dos deputados estaduais nos 15 estados correspondentes às maiores capitais consideradas acima. Observe que existe um platô inicial, em seguida um regime de escala de lei de potência com um expoente próximo a $-1,0$. A partir, para grandes frações eleitorais ocorre uma queda mais rápida, o que não ocorre nos dados de eleições para vereadores. A rápida queda observada (exponencial) para grandes frações de votos

pode refletir uma falta de correlação entre grupos de eleitores que optaram por votar no mesmo candidato, mas estão localizados em cidades distintas.

Figura 4: Ajuste via Lei de Zipf generalizada para deputados estaduais - A distribuição de votos para deputados estaduais em log-log. Os círculos são referentes aos dados dos resultados das eleições brasileiras no ano de 1998, considerando os estados com as 15 maiores capitais. A linha contínua é o melhor ajuste para a Lei de Zipf generalizada, Equação (2.1.5): O expoente de escala da lei de potência é 1,03.



Fonte: Lyra et al. (2003, p.134).

Portanto o gráfico da Figura 4 apresenta uma composição entre a exponencial usual ($q = 1$), e a exponencial q , equação (2), ($q \neq 1$). Assim é razoável supor $N(v)$ solução da equação

$$\frac{dN(v)}{dv} = -\lambda N - (\lambda_q - \lambda)N^q, \quad (2.1.4)$$

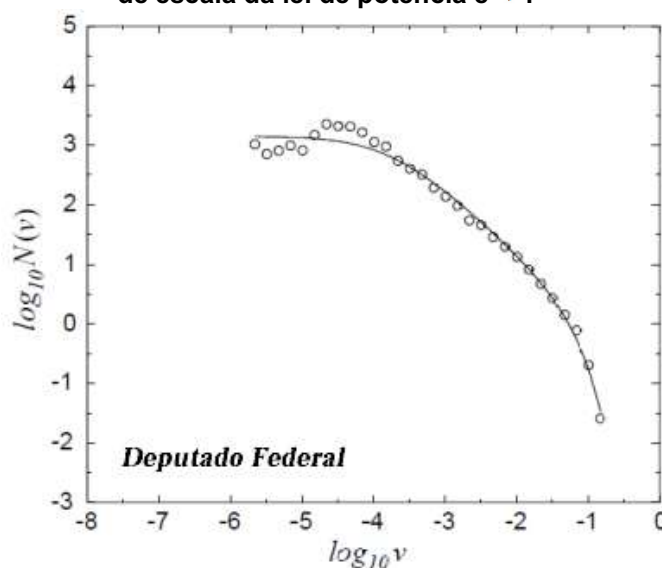
cuja solução é dada por

$$N(v) = \frac{N(0)}{\left[1 - \frac{\lambda'_q}{\lambda} + \left(\frac{\lambda'_q}{\lambda}\right)e^{(q-1)\lambda v}\right]^{1/(q-1)}}, \quad (2.1.5)$$

na qual $\lambda'_q = \lambda - (\lambda - \lambda_q)/N(0)^{(1-q)}$, a Equação 2.1.5 é uma lei de Zipf generalizada. Na Figura 2.1.4 a linha continua representa o melhor ajuste de $N(v)$ para a Equação 2.1.5 correspondendo aos parâmetros $\lambda'_q = 12.227,3$, $\lambda = 135,1$ e $\alpha = \frac{1}{q-1} = 1,03$ ($q = 1,97$). Para dados de deputados federais para as eleições 1998 obtidos nos 15 estados usados anteriormente são mostrados na Figura 5. A linha

contínua mostra o melhor ajuste para os dados com em que os parâmetros são parâmetros $\lambda'_q = 6.774,4$, $\lambda = 33,1$ e $\alpha = \frac{1}{q-1} = 1,07$ ($q = 1,93$).

Figura 5: Ajuste via Lei de Zipf generalizada para deputados federais - A distribuição de votos para deputados federais em log-log. Os círculos são referentes aos dados dos resultados das eleições brasileiras no ano de 1998, considerando os estados com as 15 maiores capitais. A linha contínua é o melhor ajuste para a Lei de Zipf generalizada, Equação (2.1.5): O expoente de escala da lei de potência é 1,0.



Fonte: Lyra *et al.* (2003, p.135).

2.2 Distribuição de votos em eleições proporcionais levando em conta a relevância do partido

Os resultados apresentados na seção anterior foram referentes às distribuições do número de candidatos em função do número de votos em eleições proporcionais, onde foi obtido para a parte central dos dados um ajuste Lei de Potência com expoente -1 Costa Filho *et al.* (1999; 2002) e um ajuste para toda a extensão dos dados por uma Lei de Zipf generalizada, Lyra *et al.* (2003). Os estudos apresentados nessa seção também analisam distribuições de votos recebidos pelos candidatos em eleições proporcionais, porém levando em conta a relevância do partido, ou seja, os votos são reescalados pelo número médio de votos recebidos pelos candidatos do partido do candidato em questão. O artigo Fortunato e Castellano (2007) obtém um comportamento do tipo lognormal para dados de três países sugerindo que com esse procedimento o comportamento obtido deve ser

universal, no entanto em Araripe e Costa Filho (2009) foram analisados, utilizando o mesmo procedimento, dados do Brasil e da Finlândia, obtendo para o Brasil uma distribuição exponencial. Em seguida, também aplicando o mesmo procedimento, foi realizada uma análise de dados para um grupo de 15 países, a qual apresentou o comportamento lognormal obtido em Fortunato e Castellano (2007) para um grupo de 5 desses países para os quais as regras eleitorais similares.

No estudo Fortunato e Castellano (2007), foram considerados 3 países com o sistema eleitoral de eleições proporcionais, Itália (até 1992), Itália (1958, 1972, 1987), Polônia (2005) e Finlândia (2003). Para os procedimentos dessa análise foram feitas as seguintes considerações para o candidato i .

v_i – número de votos do candidato i ;

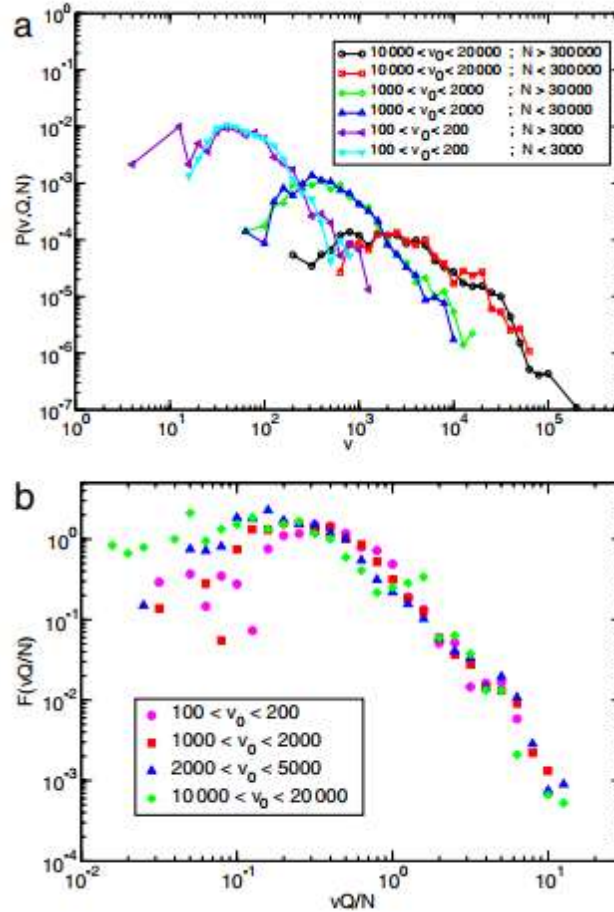
l_i – lista dos candidatos do partido de i ;

Q_i – número de candidatos na lista l_i ;

N_{l_i} – número total de votos recebidos pelos Q_i candidatos.

Considerando-se então a distribuição de probabilidade do número de votos recebidos por candidato $P(v; Q, N)$, que não depende de N e Q separadamente, mas na proporção $v_0 = N/Q$ que representa o número médio de votos coletados por um candidato em sua lista, sendo assim reescrita como uma função de uma única variável reescalada, $F(vQ/N)$. As curvas da Figura 6 (a) correspondem a três valores diferentes de v_0 . Uma vez que v_0 é uma variável contínua, para cada valor de v_0 , foi fixado um limiar para número total N de votos, separando as listas com N maior ou menor que o limiar.

Figura 6: (a) Comportamento de escala da distribuição de votos recebidos pelos candidatos. Os dados referem-se às eleições parlamentares italianas em 1972, mas foram obtidos resultados muito semelhantes para a análise de cada eleição da base de dados. O histograma $P(v, Q, N)$ depende apenas da razão $v_0 = N/Q$, então $P(v, Q, N) = P_0(v, v_0)$. (b) A função $P_0(v, v_0)$ mostrado em (a) depende apenas da proporção $v/v_0 = vQ/N$.



Fonte: Fortunato e Castellano (2007, p.2).

Fixando v_0 os histogramas obtidos colapsam, como pode ser visto na Figura 6 (b), portanto, não há dependência de v_0 . Resultando que $P(v, Q, N)$ é na verdade uma função da única variável $vQ = N$; ou seja,

$$P(v, Q, N) = F(vQ/N) \quad (2.2.1)$$

O gráfico apresentado na Figura 7, indica que cada eleição pode ser caracterizada por uma única função $F(vQ/N)$ é a mesma para diferentes países e anos. Note que sendo v_0 o número médio de votos de um candidato de uma lista l , a taxa $vQ/N = v/v_0$, é um índice do desempenho do candidato em relação aos seus concorrentes na mesma lista. Assim, se $v/v_0 < 1$ o candidato tem desempenho

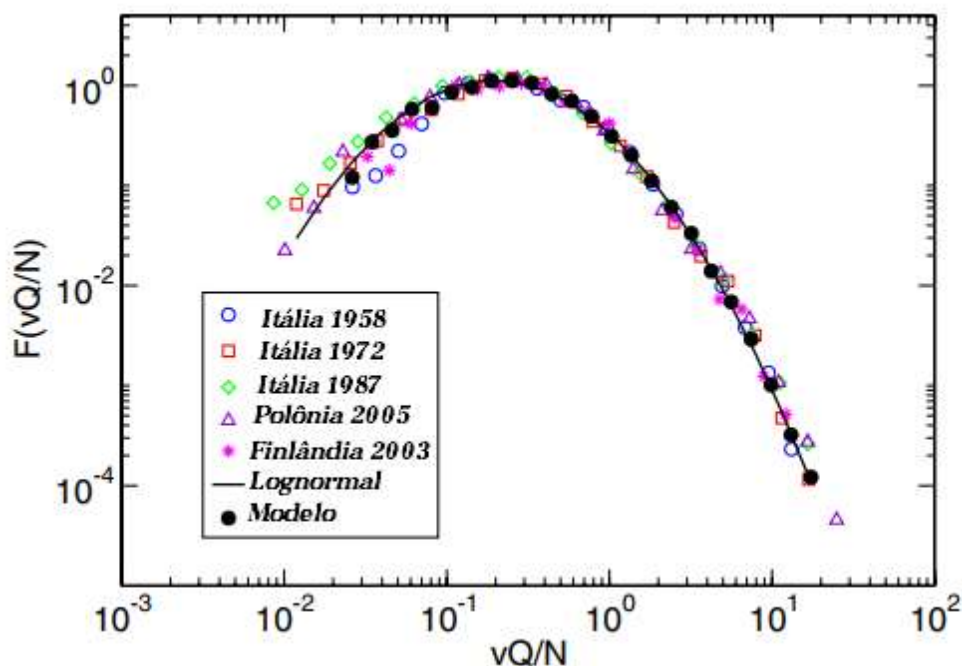
abaixo da média e se $v/v_0 \gg 1$ o candidato teve uma votação muito maior do que a média.

A Equação 2.2.2 mostra que a distribuição reescalada de todas as eleições podem ser ajustadas pela mesma função $F(vQ/N)$ e a curva universal ajustando muito bem as 5 eleições da base de dados é a log-normal dada por,

$$F(vQ/N) = \frac{N}{\sqrt{2\pi}\sigma vQ} e^{-(\log(vQ/N)-\mu)^2/2\sigma^2} \quad (2.2.2)$$

com $\mu = -0,54$ e $\sigma^2 = -2\mu = 1,08$. As distribuições dos dados para as 5 eleições ajustadas pela Equação 2.2.2 são mostradas na Figura 7.

Figura 7: Universalidade da função de escala $F(vQ/N)$ para as 5 eleições da base de dados. O ajuste lognormal, realizado na curva polonesa, descreve muito bem os dados. A curva universal é bem reproduzida pelo modelo, onde a dinâmica das opiniões dos eleitores reflete o espalhamento “boca a boca” do eleitorado do partido.



Fonte: Adaptada de Fortunato e Castellano (2007, p.3).

Note que sendo os autores interpretaram os resultados como uma dinâmica de opinião baseado na boca a boca, onde os eleitores que já escolheram seus candidatos tentam convencer seus contatos a votarem no mesmo candidato, onde apenas os eleitores indecisos podem ser convencidos. Foi implementado um processo representando o eleitorado de um partido como um conjunto de comunidades de eleitores semelhantes a árvores, com candidatos como raízes

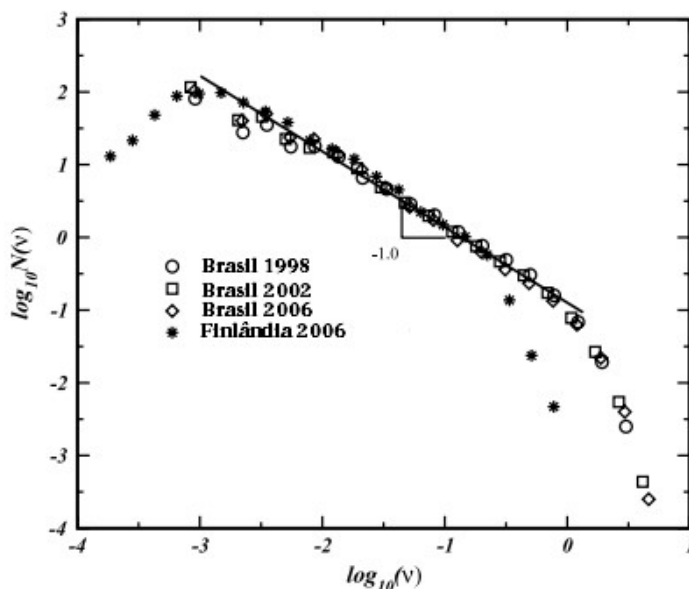
tornando semelhante a uma lei de potência. Concluem que a divulgação boca a boca é um meio crucial para explicar outras instâncias da dinâmica social coletiva, como a propagação de notícias e modismos em uma população e a difusão de novos produtos entre os consumidores em potencial.

Seguindo a análise anterior os autores Araripe e Costa Filho (2009), analisaram dados de eleições proporcionais do Brasil e da Finlândia. Este estudo verificou que dependendo da maneira com que os votos são contados pode-se descobrir distribuições diferentes, porém ao considerar apenas a fração de votos recebida pelo candidato a distribuição apresenta um comportamento que segue a lei potência. No entanto, esse comportamento universal é alterado quando se normaliza a fração de votos pelo número médio de votos do partido do candidato.

Nessa análise foram utilizados dados dos resultados das eleições proporcionais brasileiras realizadas em 1998, 2002 e 2006 e comparada com um país Europeu com o mesmo sistema eleitoral, neste caso a Finlândia nas eleições de 2003.

Os autores consideraram inicialmente a distribuição do número de candidatos recebendo a fração de votos v , sem levar em conta a votação do partido, procedimento mostrado na Figura 8. O gráfico log-log mostra o número de candidatos N , recebendo uma fração de votos v , onde é possível observar um comportamento da lei de potência $N(v) \propto v^{-\alpha}$ com $\alpha = 1 \pm 0,02$. Onde os autores confirmam o comportamento observado nos estudos Costa Filho *et al.* (1999, 2002), em que obtiveram a relação lei de potência $1/v$ ajustando a parte central dos dados.

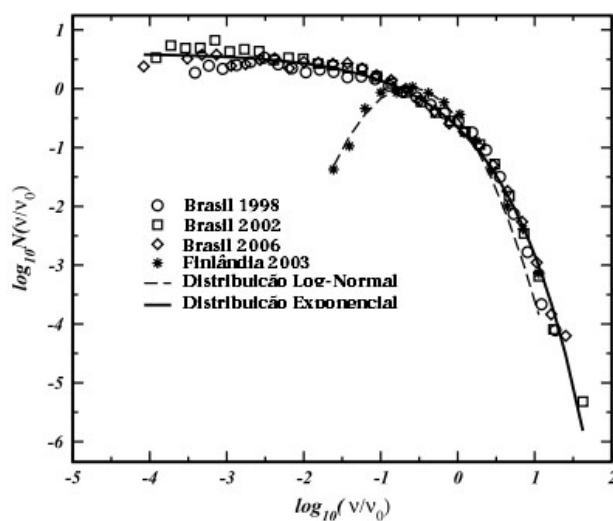
Figura 8: Histograma em log-log da distribuição dos votos recebidos pelos candidatos em três eleições brasileiras e uma na Finlândia. Além da diferença entre esses países e do intervalo de oito anos no caso do Brasil, todas as curvas seguem a mesma distribuição, uma lei de potência com expoente $\alpha = 1,0$.



Fonte: Adaptada Araripe e Costa Filho (2009, p.2).

No procedimento seguinte, assim como no trabalho anterior, Fortunato e Castellano (2007), os votos de cada candidato é normalizado pelo número médio de votos do partido a que pertence. O resultado é mostrado na Figura 9.

Figura 9: Histograma mostrando a distribuição de votos da variável v normalizada de acordo com o número médio de votos v_0 dos candidatos dentro do seu partido. As curvas são ajustadas por diferentes funções. Para o Brasil, a distribuição segue uma queda exponencial para todos os anos, enquanto na Finlândia o melhor ajuste é um log-normal



Fonte: Adaptada de Araripe e Costa Filho (2009, p.3).

Observe que, para as eleições brasileiras, a distribuição é a mesma para as três eleições diferentes e o melhor ajuste encontrado foi:

$$F(v/v_0) = \gamma e^{-\frac{\beta v}{v_0}}, \quad (2.2.3)$$

para $\gamma = 1,86$ e $\beta = 1,262$.

No caso da Finlândia, o ajuste segue uma distribuição lognormal, assim como foi obtido em Fortunato e Castellano (2007), Equação 2.2.2, com parâmetros $\mu = -0,26$ e $\sigma = 0,603$.

Para pontos na vizinhança de $v/v_0 = 1$, onde os candidatos tiveram desempenho próximo a média de seu partido, as duas distribuições apresentam frequências semelhantes. Para valores de $v/v_0 > 1$, onde os candidatos foram bem votados em relação aos outros candidatos de seu partido, a distribuição apresenta uma diferença menor. Já na região $v/v_0 < 1$, existe uma grande discrepância. Esse resultado mostra que mesmo em países com o mesmo sistema de eleições proporcionais com listas abertas o procedimento utilizado pelos autores Fortunato e Castellano (2007), Araripe e Costa Filho (2009), não apresenta um modelo universal.

No intuito de encontrar uma escala que torne a universalidade das distribuições de probabilidade mais gerais possíveis, os autores Chatterjee *et al.* (2012) utilizaram três diferentes escalas para investigar os resultados de eleições parlamentares, com sistemas proporcionais de 15 países: Itália (antes de 1992), Polônia, Finlândia, Dinamarca, Estônia, Suécia, Bélgica, Suíça, Eslovênia, República Checa, Grécia, Eslováquia, Holanda, Uruguai e Brasil.

Países que utilizam sistema de eleições proporcionais são divididos em distritos e cada partido político apresenta uma lista de candidatos para distribuir as cadeiras entre os partidos. Os eleitores indicam sua preferência votando em um partido político e/ou um ou mais candidatos da lista. O número de assentos atribuídos a cada partido é proporcional ao número de votos coletados pelo partido. Os sistemas de lista partidária podem ser classificados em abertos, semi-abertos e fechados. Em listas abertas permite-se que o eleitor escolha partidos e candidatos e a posição de cada candidato depende somente do número de votos que ele recebe. No caso das listas semi-abertas o eleitor vota em um partido ou em um candidato em uma lista de partidos. Os partidos geralmente apresentam uma lista de candidatos de acordo com a preferência do partido. Os candidatos conquistam assentos parlamentares na ordem em que são classificados na lista. No entanto, se um candidato obtiver um número de votos que excede a um dado limiar, então ele

sobe na classificação. Já no sistema de lista fechada, o partido fixa a ordem em que os candidatos serão eleitos. O eleitor vota para um partido e não pode escolher um candidato.

Com o objetivo de verificar o comportamento da distribuição de votos levando em conta a relevância do partido em sistemas com listas abertas (mesmos procedimento de Fortunato e Castellano (2007) agora incluindo mais países), foram feitas as seguintes considerações: v_i número de votos que cada candidato i recebeu e o número de candidatos Q_i da lista do partido l_i dos candidatos do partido de i . Seja N_{l_i} o número total de votos coletados pelos candidatos Q_i da lista l_i . Somando todas as listas partidárias do distrito D_i obtém-se o número de votos N_{D_i} do distrito. O número total de votos na eleição de um dado país é indicado como N_T .

Para construir as distribuições de votos recebidos por um candidato, foram feitas as seguintes normalizações:

- 1) O número de votos v_i de um candidato é dividido pelo número médio de votos $v_0 = N_{l_i}/Q_i$ dos candidatos do partido de i . Mesma escala usada em Fortunato e Castellano (2007), a qual foi chamada nesse trabalho de FC, identificando com os autores, Fortunato e Castellano.

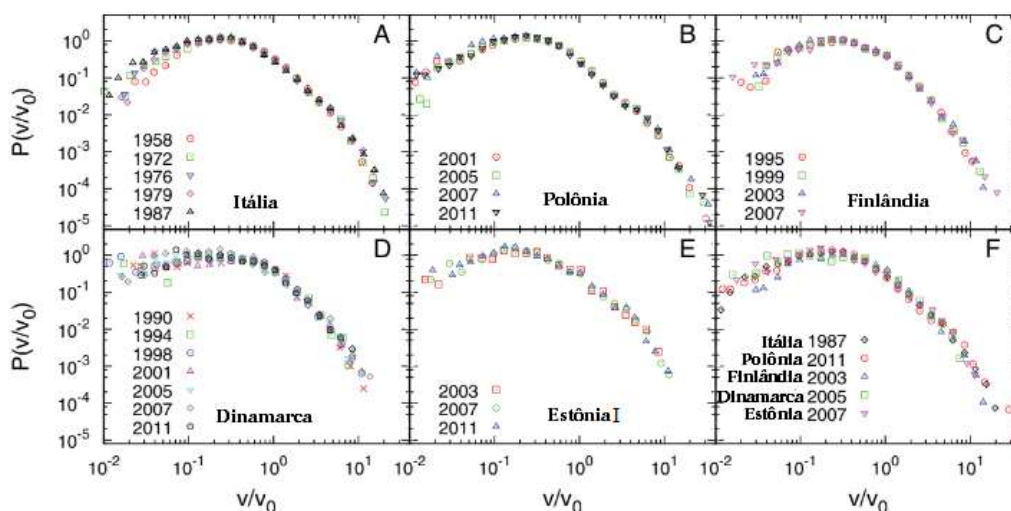
(a) A fração do total de votos no distrito v_i/N_T e (b) a fração do total de votos no país v_i/N_{D_i} . Escalas que serão referidas aqui como CAAMd e CAAMn, respectivamente. Essa identificação é devido a Costa Filho *et al.* (1999).

Usando a escala FC obteve-se o ajuste universal encontrado em Fortunato e Castellano (2007) para as eleições da Finlândia, Polônia e Itália em vários anos, e agora o resultado contempla também os dados eleitorais da Dinamarca e Estônia, como pode ser visto na Figura 10 (A, B, C, D e E). No painel F é mostrado o colapso dessas curvas que podem ser bem ajustadas pela função log-normal.

Outros países que utilizam listas abertas são Eslovênia, Grécia, Suíça, Brasil e Uruguai, onde os resultados utilizando a escala FC são mostrados na Figura 11. Embora exista um histórico persistente da distribuição a nível nacional, as curvas não seguem um padrão comum através dos países distinguindo-se em duas classes de comportamento: Eslovênia, Grécia e Suíça são caracterizadas por um pico em $v/v_0 = 1$ e suas caudas se assemelham bastante. Diferentemente das outras três curvas, Brasil e Uruguai apresentam um padrão monotônico. Comparada com a

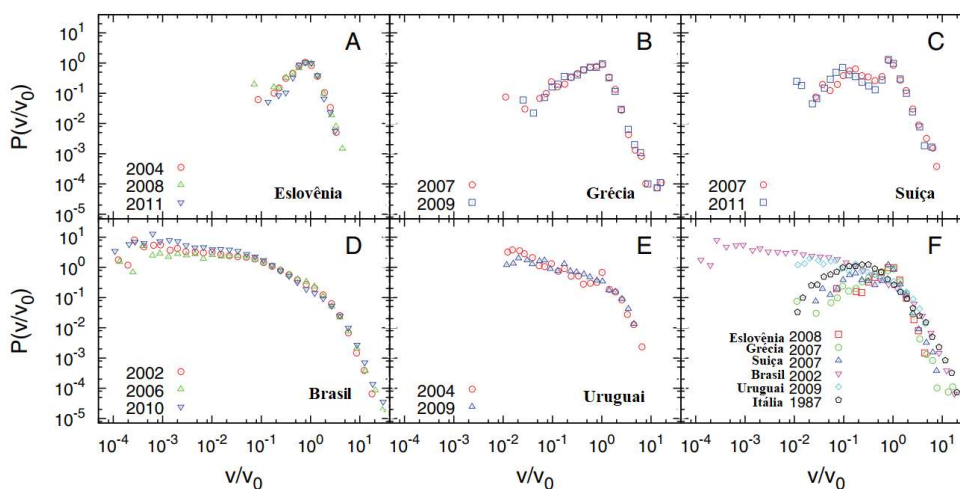
curva universal, a curva brasileira se apresenta muito próxima da Figura 11, sobre a cauda $v/v_0 > 1$. É possível concluir que os sistemas de listas abertas não garantem distribuições idênticas, mas podem ser agrupadas em classes de comportamentos semelhantes.

Figura 10: Distribuição do desempenho eleitoral dos candidatos em eleições proporcionais com listas abertas, segundo escalonamento do CF. Itália (até 1992), Polônia, Finlândia, Dinamarca e Estônia (após 2002) seguem essencialmente as mesmas regras, o que se reflete no colapso dos dados do painel F. A evolução histórica dos países não parece afetar a forma da distribuição (painéis A a E).



Fonte: Adaptada de Chatterjee *et al.* (2013, p.3).

Figura 11: Mesma análise da Figura 10, para Eslovênia, Grécia, Suíça, Brasil e Uruguai. As curvas são razoavelmente estáveis em nível nacional, mas não se comparam bem entre os países e com as curvas universais da Figura 10 (representadas no painel F pela distribuição para as eleições italianas em 1987). É provável que tais discrepâncias se devam às diferentes regras eleitorais desses países em comparação entre si e às examinadas na Figura 10, embora todos adotem listas abertas.

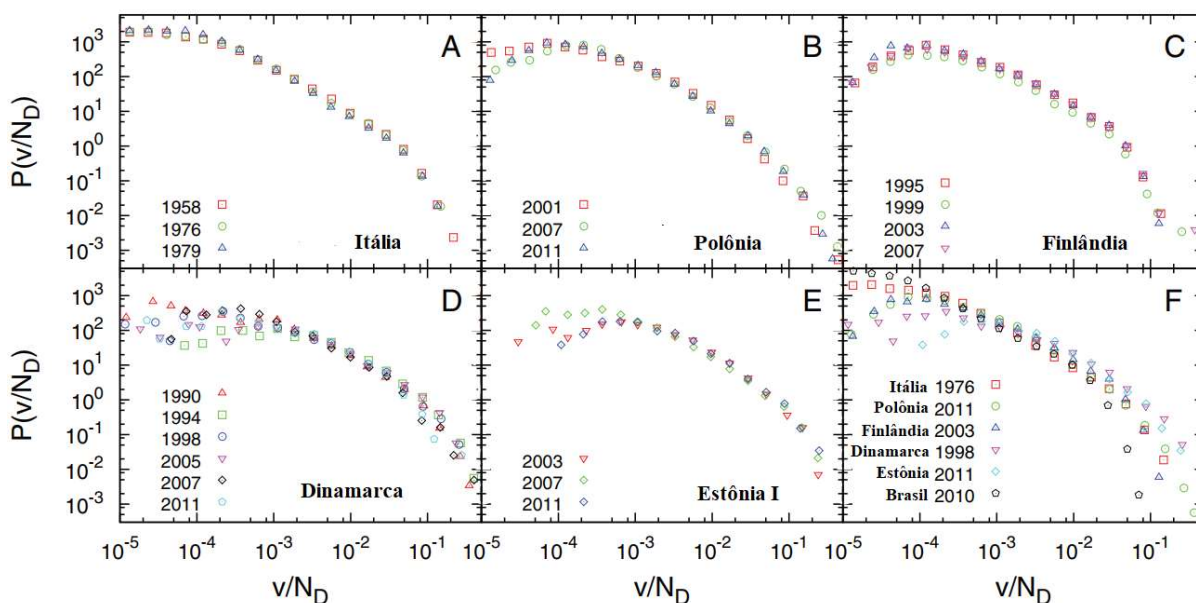


Fonte: Adaptada Chatterjee *et al.* (2013, p.4).

A Figura 12, mostra o procedimento usando a escala CAAMd para a mesma sequência de eleições da Figura 10.

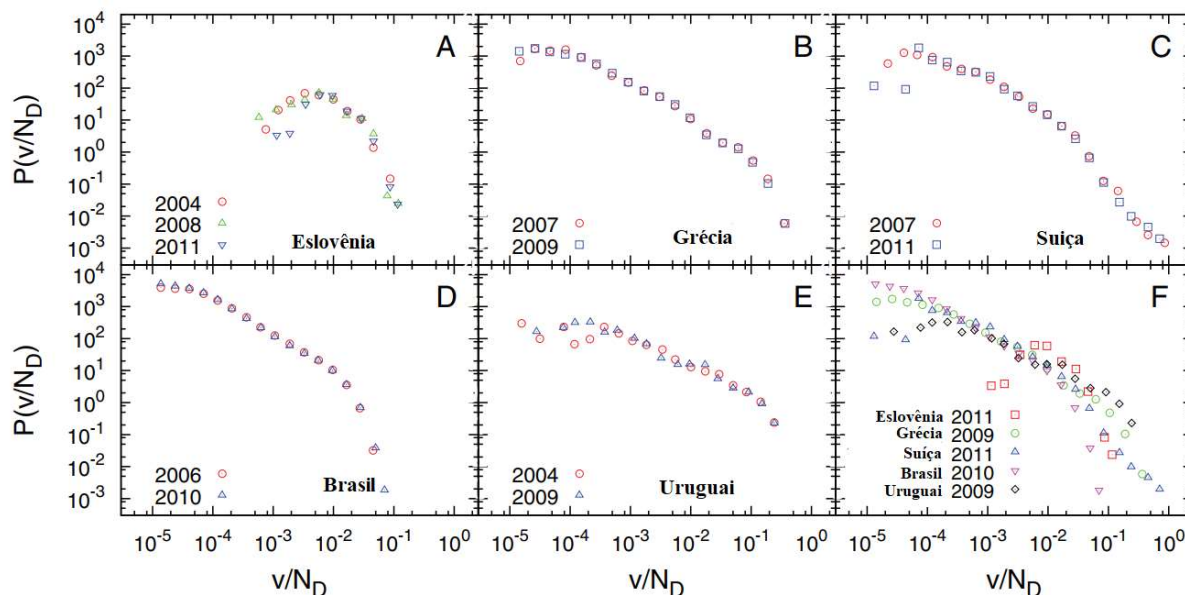
As Figuras 12 e 13 mostram o procedimento usando a escala CAAMd para a mesma sequência de eleições da Figura 10 e 11, respectivamente.

Figura 12: Escala CAAMd e listas abertas - Análise para os mesmos países da Figura 10, com escala CAAMd. As curvas são estáveis a nível nacional, mas elas não se assemelham através de todos os países.



Fonte: Adaptada Chatterjee *et al.* (2013, p.4).

Figura 13: Escala CAAMd e listas abertas - Análise para os mesmos países da Figura 11, com escala CAAMd. As curvas são estáveis a nível nacional, mas elas não se assemelham através de todos os países

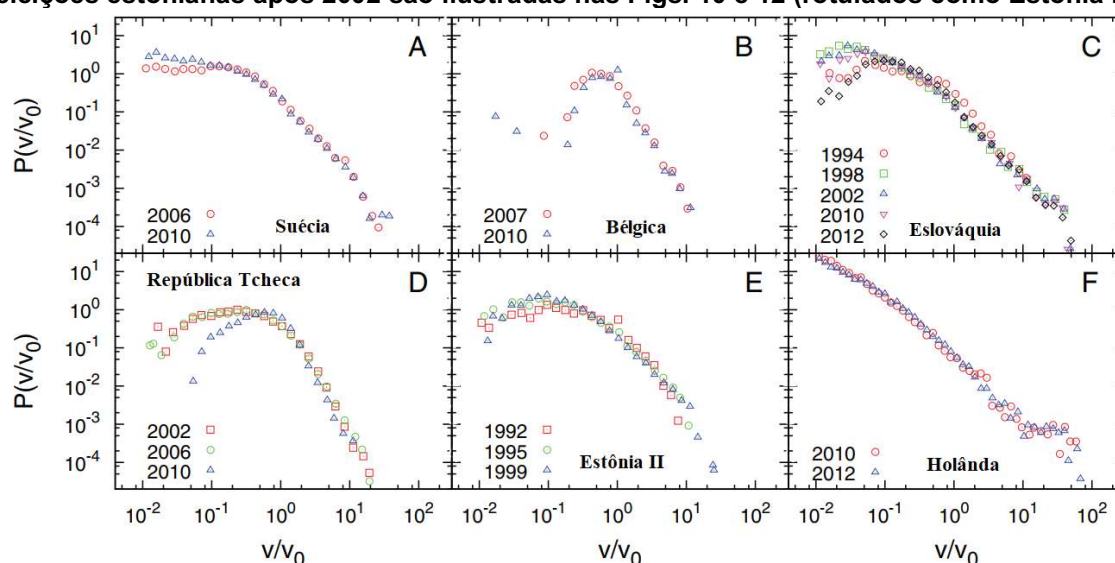


Fonte: Adaptada Chatterjee *et al.* (2013, p.5).

Observe que as distribuições de v/N_D para o grupo de países com listas abertas não apresenta comportamento universal para nenhum grupo de países como ocorreu com a escala FC.

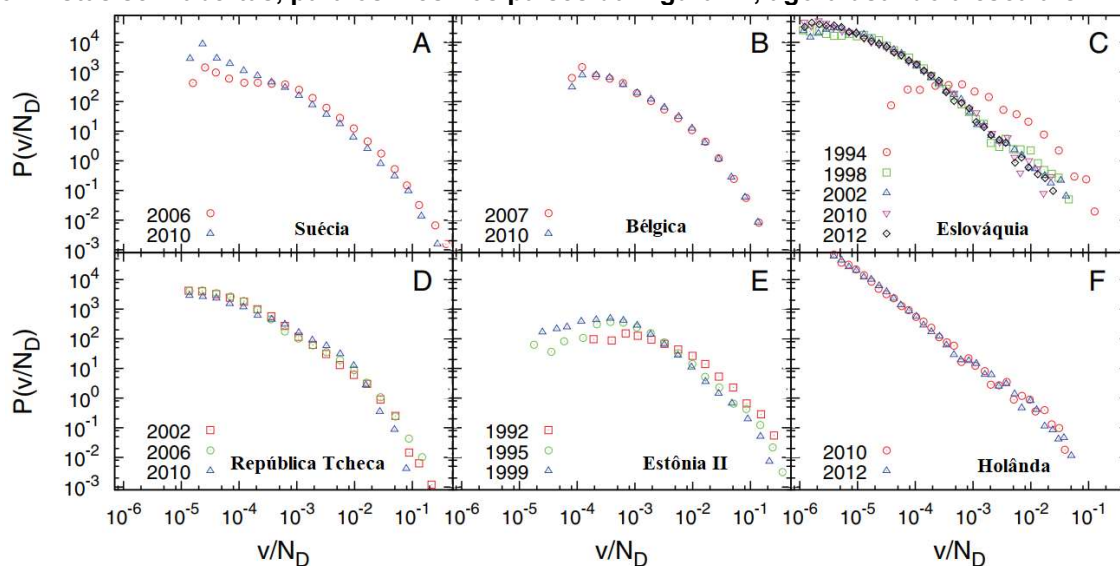
Considerando o restante dos 15 países da base de dados, os quais usam sistemas de listas semi-abertas, foram aplicados os mesmos procedimentos obtendo as distribuições v/v_0 , escala FC, na Figura 14 e em seguida as distribuições de v/N_D , que representa a escala CAAMd, Figura 15.

Figura 14: Distribuição do desempenho eleitoral dos candidatos em eleições proporcionais com listas semiabertas, usando a escala FC. A nível nacional, as curvas são na sua maioria estáveis, discrepâncias significativas correspondem a alterações nas regras eleitorais, como na Eslováquia (C), República Tcheca (D) e Estônia (E). A aparente lei de potência da curva holandesa (F) pode ser gerada por um mecanismo mais rico, uma vez que o limite é muito alto (10% em nível nacional) e os eleitores tendem a apoiar os candidatos com base em sua popularidade. A Estônia desde 2002 adotou listas abertas, razão pela qual as distribuições das eleições estonianas após 2002 são ilustradas nas Figs. 10 e 12 (rotulados como Estônia I).



Fonte: Adaptada de Chatterjee *et al.* (2013, p.5).

Figura 15: Distribuição do desempenho eleitoral dos candidatos em eleições proporcionais com listas semiabertas, para os mesmos países da Figura 14, agora usando a escala CAAMd



Fonte: Adaptada de Chatterjee *et al.* (2013, p.6).

Em resumo, esse estudo concluiu que: dos 15 países e dos 10 com listas abertas 5 deles tem a distribuição do desempenho do candidato, quando comparado ao comportamento médio dentro do partido, ou seja, usando a escala FC, muito similares e ajustada por uma log-normal. O restante dos países, 5 com listas abertas e 5 com listas semi-abertas, usando a escala FC ou a escala CAAMd (desempenho no distrito), não apresentaram padrão entre eles.

2.3 Leis de potência modelando processos de candidaturas

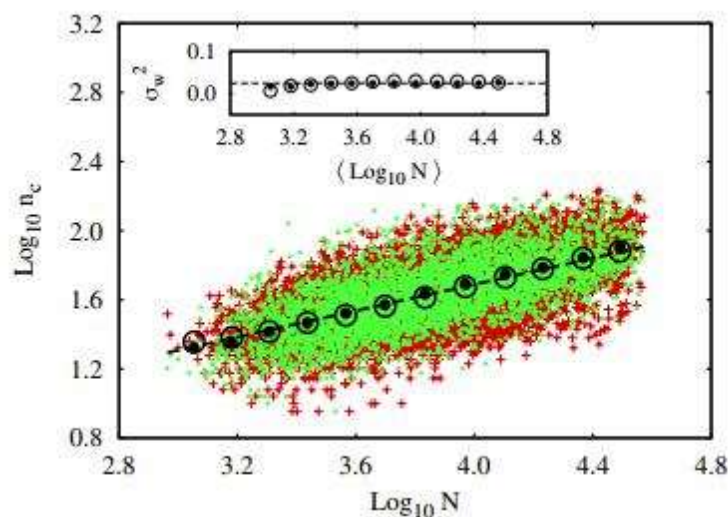
No intuito de investigar a dinâmica de escolha de candidatos em eleições, uma questão ainda em aberto, o artigo Mantovani *et al.* (2011), realizou uma pesquisa que analisou os processos de candidatura para as eleições do Brasil (1996, 2000, 2004, 2008 - eleições para vereador e prefeito), da Itália (2010 - eleições para o conselho municipal), da Inglaterra (2006, 2010 - eleições para o conselho municipal), do Canadá (2006, 2008 - eleições parlamentares) e da Austrália (2004, 2007, 2010 - eleições parlamentares). Das eleições analisadas, cinco são para prefeito, seis para o conselho municipal e cinco correspondem a eleições nacionais.

Para as análises foram considerados apenas municípios que possuem até 47.600 habitantes e que elegem apenas nove vereadores. De modo específico para as eleições italianas, considerou-se apenas eleições para prefeito (Provinciali e Comunali). No caso das eleições inglesas, considerou-se eleições do conselho em todos os 624 distritos nos bairros de Londres, onde a maioria dos distritos tem três assentos. Já para as eleições do Canadá e da Austrália, os dados são das eleições da “Câmara dos Comuns” e da “Câmara dos Representantes”, cada uma com 308 e 160 membros, respectivamente. Ressalta-se que em ambos os casos, os candidatos são eleitos em eleições uninominais disputadas em cada distrito eleitoral.

Inicialmente foi investigado como o número de candidatos n_c cresce com o número de eleitores N , o espalhamento dessa relação para as variáveis em logaritmo na base 10, representado pelas cruzinhas vermelhas, pode ser visto na Figura 16, para os dados das eleições de vereadores no Brasil no ano de 2008.

Neste gráfico é possível observar uma tendência linear resultante dessa relação, ou de forma equivalente, n_c cresce com N seguindo uma lei de potência, ou seja $n_c \sim N^a$. Relacionado ao possível comportamento linear do gráfico, tem-se o coeficiente de correlação de Pearson, $r = 0,63$. Esse valor de r e o aspecto visual à dispersão dos dados indicam que a correspondente lei de potência é subjetivada pelas flutuações.

Figura 16: Espalhamento dos dados empíricos e do modelo - Gráfico de dispersão do número N de eleitores versus o número n_c de candidatos tomados o logaritmo na base-10 para eleições brasileiras de vereador em 2008. As cruzes vermelhas representam os dados empíricos e os pontos verdes são resultados da simulação obtidos do modelo de Barabási-Albert com $c = 1,12$, $\alpha = 0,196$ e $\beta = 0,32$. Os círculos abertos (preenchidos) são os valores médios, calculados em janelas igualmente espaçadas, para os dados (resultados simulados) e a linha pontilhada e o ajuste linear para os valores médios dos dados empíricos, encontrando $n_c \sim N^{0,36}$. A inserção mostra a mesma comparação para os valores da variância σ_ω^2 .



Fonte: Mantovani et al. (2013, p.3).

Para superar as flutuações e evidenciar a lei de potência, os autores construíram janelas ω igualmente espaçadas em $\text{Log}_{10}N$ e então calcularam os valores médios de $\text{Log}_{10}N$ nas janelas e os valores médios correspondentes da variável $\text{Log}_{10}n_c$, obtendo com esse procedimento os pontos médios $(\langle \text{Log}_{10} N \rangle, \langle \text{Log}_{10} n_c \rangle)$ representando a tendência média dos dados, os valores médios são representados pelos círculos em preto mostrados na Figura 16.

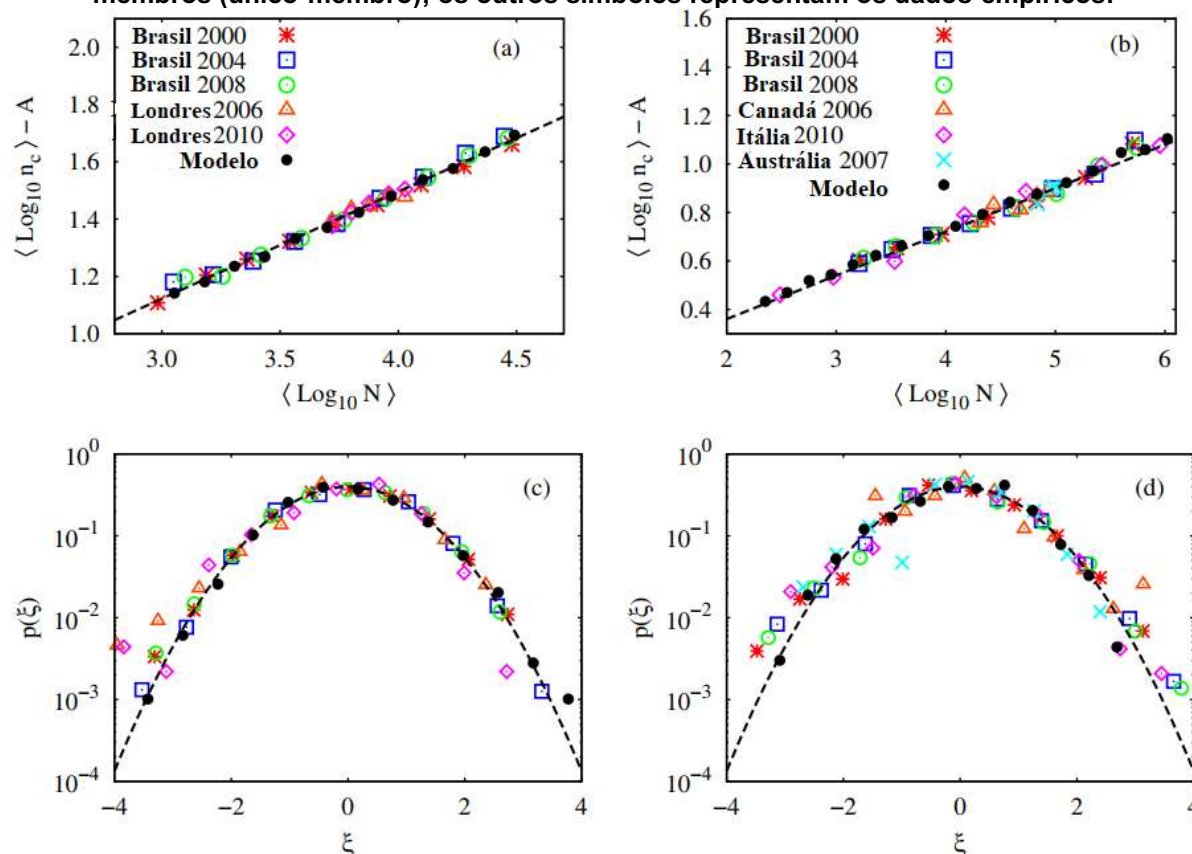
Para os valores médios, a relação de lei de potência torna-se evidente, por meio de:

$$\langle \log_{10} nc \rangle = A + \alpha \langle \log_{10} N \rangle \quad (2.3.1)$$

com $\alpha = 0,36$ e $\langle \dots \rangle$ representando o valor médio em cada janela ω .

Os painéis superiores da Figura 17 (a) mostra o procedimento dos valores médios aplicado a 5 eleições para o conselho municipal a Figura 17 (b) faz o mesmo para 4 eleições para prefeito e 2 parlamentares. Observe que os gráficos dessas relações foram construídos descontando o valor da constante A com o objetivo de colapsar os dados.

Figura 17: Comportamento médio para vereadores e prefeitos - Valores médios do logaritmo do número de candidatos N_c versus o logaritmo do número de eleitores N , descontado a constante A, para (a) 5 eleições de multi-membros e (b) 6 eleições de um único-membro. As linhas tracejadas representam leis de potências para os valores médios, cujos expoentes médios α são $\alpha = 0,36$ para eleições de multi-membros e $\alpha = 0,18$ para eleições de um único-membro. Os círculos pretos são os dados simulados com o modelo de rede apresentado ao final do deste capítulo, tomando a rede de Barabási-Albert com $\beta = 0,32$ ($\beta = 0,41$) para multi-membros (único-membro); os outros símbolos representam os dados empíricos.



Fonte: Mantovani et al. (2013, p.4).

A partir dessas figuras os autores identificaram que a boa qualidade do colapso obtido e os valores de α sugeriram duas classes de universalidades: uma para eleições de um único-membro, caracterizada por $\alpha = 0,18$, e outra para eleições de vereadores (multi-membros) com $\alpha = 0,36$.

Em seguida foi abordado as questões de flutuação em torno do comportamento médio, considerando a variância do logaritmo do número de candidatos. Foram calculadas as variâncias da variável $\log_{10} Nc$, denotada por σ_{ω}^2 , correspondentes as janelas ω . Os resultados obtidos pelos autores indicaram que a variância não depende do número de eleitores N , fato que pode ser observado na Figura 16. Também foi possível investigar a flutuação em torno da lei de potência, em que se considerou:

$$\xi = \frac{\log_{10} Nc - f_w(N)}{\sigma_w} \quad (2.3.2)$$

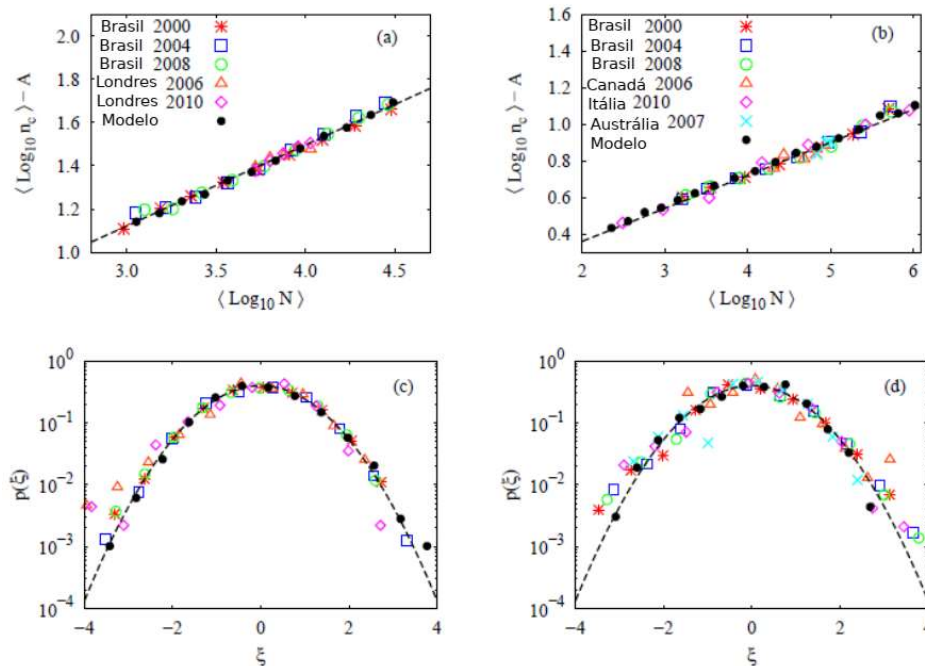
em que:

$$f_w(N) = A + \alpha \langle \log_{10} N \rangle_w \quad (2.3.4)$$

representa a função ajuste para os valores médios considerando cada janela ω .

As Figuras 18 (c) e (d) mostram a função de distribuição de probabilidade de ξ para os ajustes nas Figuras 17 (a) e (b). A partir dessas figuras, observamos que ξ segue muito próximo ao padrão gaussiano.

Figura 18: Valores médios e suas flutuações- O painel superior mostra os valores médios, em log-log, do número N_c de candidatos versus o número N de eleitores descontando a constante A para (a) 5 eleições de multi-membros e (b) 6 eleições de um membro. As linhas tracejadas representam os ajustes lineares para os valores médios com expoente $\alpha = 0,36$ para eleições de multi-membros e $\alpha = 0,18$ para eleições de um único membro. No painel inferior são mostradas as distribuições de probabilidade das flutuações ξ , calculadas pela Equação 2.3.2, sendo que em: (c) para as eleições de (a) e (d) para as eleições de (b). As linhas tracejadas são as funções ajustes gaussianas com média zero e variância um. Os diferentes símbolos representam os dados e os círculos pretos são as previsões do modelo de rede, considerando o modelo Barabási-Albert com $\beta = 0,32$ ($\beta = 0,41$) para a eleição de múltiplos membros (único membro). As outras eleições apresentam o mesmo padrão.



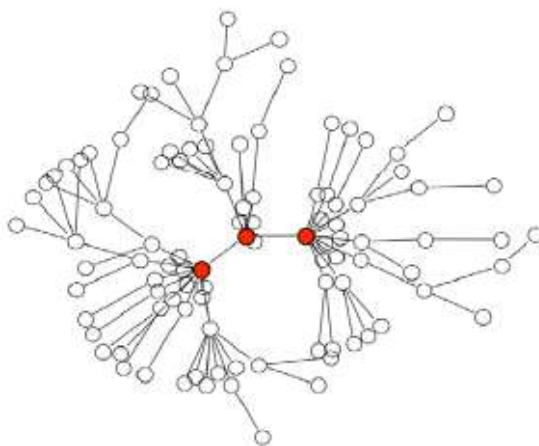
Fonte: Mantovani *et al.* (2011, p.4).

Finalizando as análises desse trabalho foi empregado modelo usando o formalismo de redes complexas Mantovanni *et al.* (2011). Dentro desta estrutura, os eleitores são representados por nós e suas interações ou relacionamentos são expressos por links entre eles. O grau de cada nó (número de links conectados ao nó) mede a “popularidade” do nó, ou seja, quanto mais links (alto grau) que um eleitor tem, maior será sua popularidade, consequentemente maior chance de ser um candidato. Sendo $p(k)$ a distribuição de probabilidade do grau e considerando que um eleitor será um candidato se sua popularidade exceder a um limiar k_c , deve-se ter $n_c \sim N \int_{k_c}^{\infty} p(k) dk$. Para a rede livre de escala utilizada aqui Barabási Albert 1999, $p(k) = k^{-\gamma}$, portanto $n_c \sim N k^{-\gamma+1}$, usando o resultado já obtido nesse trabalho que $n_c \sim N^\alpha$ tem-se, $k_c \sim N^\beta$ em que $\beta = \frac{\alpha-1}{1-\gamma}$. A Figura 19 mostra uma ilustração

dessa abordagem. As simulações via o modelo de rede podem ser vistas nas Figuras 17 e 18.

Como conclusão desse trabalho, um resultado se destaca, para eleições de multi-membros o expoente obtido foi $\alpha = 0,36$ e para eleições de um único membro o expoente encontrado foi $\alpha = 0,18$, mostrando que a dificuldade de se candidatar é maior para eleições de um único membro quando o tamanho do eleitorado aumenta.

Figura 19: Ilustração do modelo de rede para o processo de candidatura: cada vértice do grafo representa um eleitor e as ligações entre eles são as relações sociais, por exemplo, amizades. Um eleitor pondera sobre sua popularidade (aqui, o número de ligações) na comunidade levando ou não à candidatura. Os nós em vermelho são os mais populares e, conseqüentemente, candidatos potenciais.



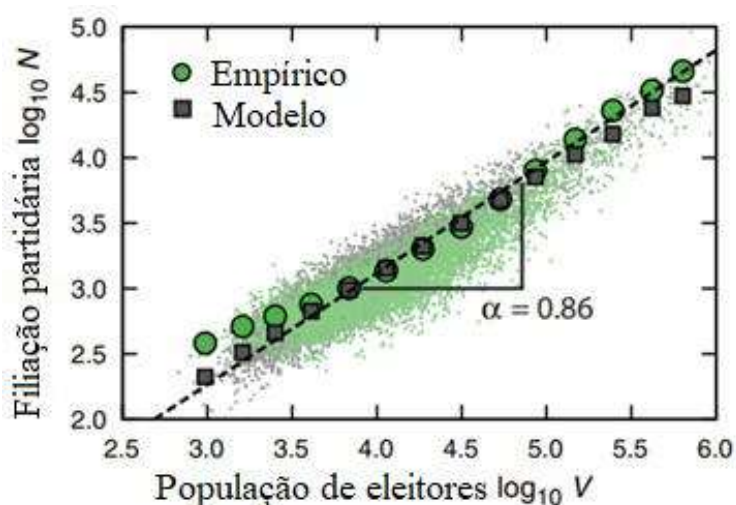
Fonte: Mantovani et al. (2011, p.5).

Dando continuidade as investigações apresentadas anteriormente, ainda com o objetivo de mostrar o engajamento nos processos de candidatura, e agora incluindo filiação partidária, Mantovani *et al.* (2013) apresentaram uma análise estatística de acordo com o número de filiação partidária e o número de candidatos a prefeito e vereador das cidades brasileiras nos anos 2000. Os autores utilizaram os dados eleitorais de todas as cidades brasileiras disponibilizados gratuitamente pelo Supremo Tribunal Eleitoral – TSE e a população de eleitores disponibilizados pelo Instituto Brasileiro de Geografia e Estatística, e descobriram que as formas funcionais dessas relações são bem descritas por leis de potência sublinear (escala alométrica) subjacente a um ruído log-normal multiplicativo.

Durante a captação dos dados, o número de cidades permaneceu quase constante (5.564 em 2006 e 5.565 em 2011) e o número de partidos políticos foi constante. Os autores definiram V como a população de eleitores e N como o

número de candidatos ou o número de membros do partido. Como no trabalho anterior, foi analisado anteriormente a relação $\log_{10} V$ versus $\log_{10} N$. O resultado para os dados de filiação partidária no ano de 2008 são mostrados na Figura 20.

Figura 20: Gráfico de dispersão da relação entre $\log_{10} N$ (população de eleitores) e $\log_{10} V$ (número de filiações partidárias) considerando os dados do ano de 2008. Note que a relação é aproximada por uma tendência a lei de potência subjacente à flutuações. Aqui, os pequenos pontos verdes representam os dados empíricos para cada cidade e os círculos são os valores médios desses pontos depois de agrupar os dados em janelas ω igualmente espaçadas em $\log_{10} V$. A linha tracejada é um ajuste linear para esses valores médios e o expoente da lei de potência é mostrado no gráfico. Os resultados simulados da Equação 2.3.5 são representados por pontos cinza, os valores médios nas janelas ω para os dados simulados são representados pelos quadrados cinzentos.



Fonte: Mantovani (2013, p.2).

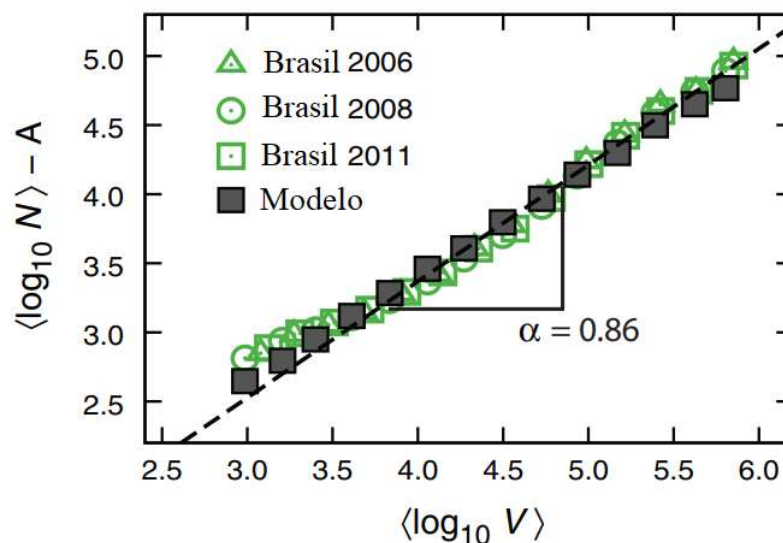
Apesar das flutuações evidenciadas no gráfico foi observado uma tendência a Lei de Potência. Para superar essas flutuações Mantovani *et al.* (2013) colocaram em ω janelas igualmente espaçadas em $\log_{10} V$ e calculando os valores médios da variável $\log_{10} V$ em ω e dos respectivos valores da variável $\log_{10} N$. Os círculos preenchidos em verde na Figura 2.3.5 representam estes valores médios para os quais foi capturada a lei de potência média, ou seja, o ajuste linear, dado por:

$$\langle \log_{10} N \rangle = A + \alpha \log_{10} V \quad (2.3.5)$$

Além disso, a lei de potência também foi avaliada em todo os anos do banco de dados conforme mostra a Figura 21. Foram obtidos os expoentes $\alpha = 0,87 \pm 0,02$, $\alpha = 0,86 \pm 0,02$, e $\alpha = 0,85 \pm 0,02$ para os anos de 2006, 2008 e 2011, respectivamente. Com os resultados obtidos é consiste

considerar que os valores de α são constantes ao longo do tempo levando em consideração os erros padrão de α .

Figura 21: Relações médias leis de potência entre $\log_{10} V$ (população de eleitores) e $\log_{10} N$ (número de filiações partidárias). Observe que foram analisados todos os anos em do banco de dados e os valores de α (as inclinações dessas curvas) são quase constantes no tempo. O valor médio de α é mostrado no gráfico e a linha tracejada representa o ajuste linear com este expoente. Os quadrados cinzas representam os resultados simulados da Equação 2.3.7.



Fonte: Mantovani (2013, p.2).

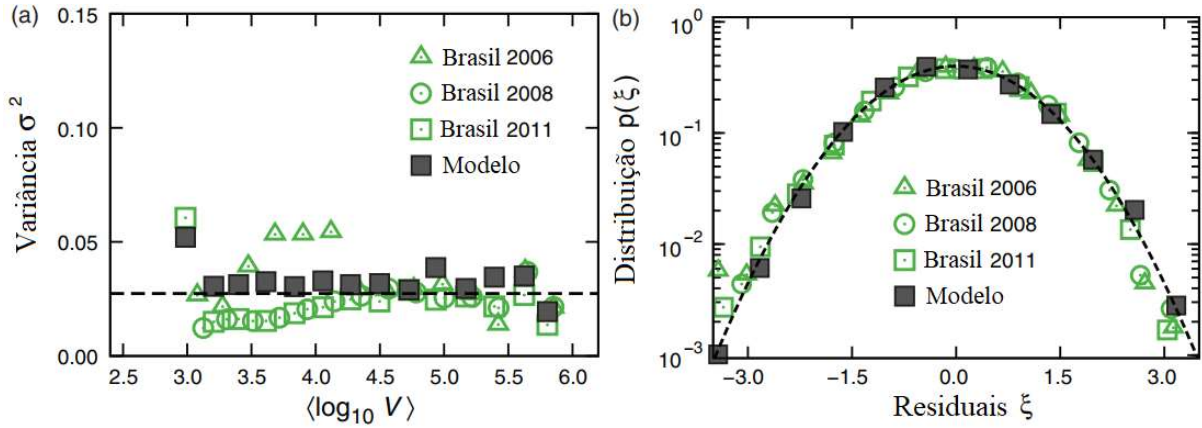
As variâncias σ^2 da variável $\log_{10} N$ nas janelas ω em função dos valores médios $\langle \log_{10} V \rangle$ para cada eleição da base de dados é mostrado na Figura 22 (a). Como pode ser visto são aproximadamente constantes mostrando a consistência do modelo obtido via valores médios.

Considerando as distribuições das flutuações normalizadas em torno das tendências a lei de potência

$$\xi(V) = \frac{\log_{10} N(V) - \log_{10} N}{\sigma} \quad (2.3.6)$$

onde σ é o desvio padrão de $\log_{10} N$ dentro das janelas ω . Pode ser observado na Figura 22 (b) as distribuições de ξ ajustadas pela distribuição gaussiana com média zero e variância unitária.

Figura 22: Flutuações em torno das tendências à lei de potência. (a) Variância σ^2 dos valores de $\log_{10} N$ versus $\log_{10} V$ calculados em janelas ω igualmente espaçadas em $\log_{10} V$. Note que o valor de σ^2 sobre a população de eleitores é quase constante. A linha tracejada horizontal é o valor médio de σ^2 sobre $\log_{10} V$ e ao longo dos anos ($0,027 \pm 0,002$). (b) Distribuição dos resíduos normalizados (ξ) em torno das tendências à lei de potência. A linha tracejada é a distribuição gaussiana com média zero e variância unitária. Em ambos os gráficos, os quadrados em cinza representam os resultados simulados da Equação 2.3.7.



Fonte: Mantovani (2013, p.3).

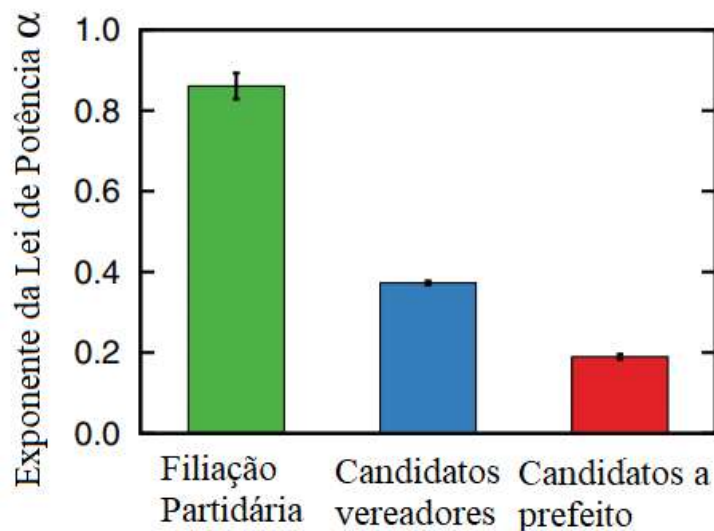
Os resultados foram modelados pela seguinte equação estocástica:

$$N = \tilde{A} \eta(V) V^\alpha \quad (2.3.7)$$

Em que \tilde{A} é uma constante, $\eta(V)$ é um ruído estocástico e α é o expoente da lei de potência da relação entre N e V . Comparando a Equação 2.3.7 com a 2.3.6, tem-se $\log_{10} \tilde{A} = A$, $\log_{10} \eta(V) = \sigma \xi(V)$ e $\langle \xi(V) \rangle = 0$.

Dentro deste comportamento universal, verificou-se que os expoentes da lei de potência α permitem distinguir as diferentes posições políticas. Comparando o valor médio do expoente da lei de potência relatado aqui para o número de membros do partido com aqueles anteriormente relatado para vereadores e candidatos a prefeito, observou-se que há uma hierarquia clara para os valores de α , como mostrado na Figura 23.

Figura 23: A hierarquia das posições políticas. Valores médios dos expoentes da lei de potência α para cada posição política. Os expoentes das relações entre o número de candidatos a vereador e prefeito e a população de eleitores foram obtidos Mantovani (2013). Notamos que a média de α é menor para as posições políticas mais influentes. As barras de erro são desvios padrão do valor médio de α .

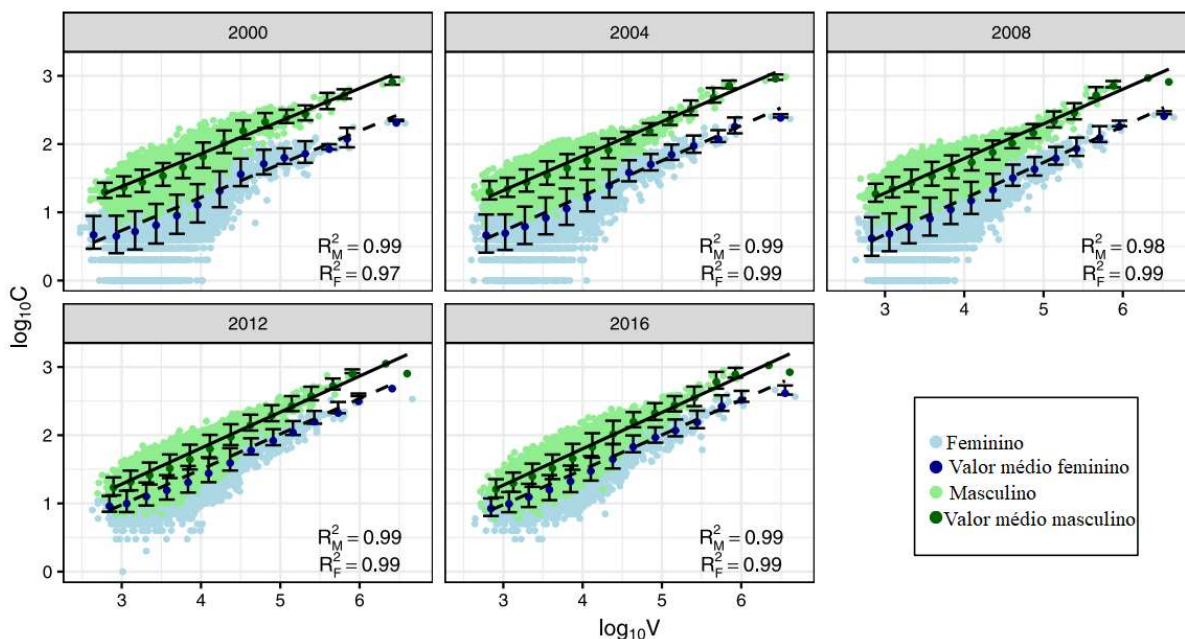


Fonte: Mantovani (2013, p.3).

O trabalho de Cardoso *et al.* (2019), assim como os trabalhos Mantovani *et al.* (2011) e Mantovani *et al.* (2013) propõem investigar processos de candidatura, porém agora levando em conta a candidatura por gênero, as variáveis tomadas aqui foram: número de candidatos femininos e masculinos e número de eleitores masculinos e femininos por cidade brasileira nas eleições para prefeito e vereador nos anos de 2000, 2004, 2008, 2012 e 2016.

Para contabilizar o efeito do tamanho do eleitorado nos processos de candidatura de mulheres e homens a vereador, foram tomados: o número de candidatos C versus o número de eleitores V , em logaritmo na base 10, para cada cidade brasileira investigada, como mostra a Figura 24. Note que ocorre a sobreposição ao longo dos anos.

Figura 24: Gráfico de dispersão das variáveis $\log_{10}V$ versus $\log_{10}C$ para as eleições para prefeito de 2000, 2004, 2008, 2012 e 2016, onde V e C representa os números dos eleitores e dos candidatos, respectivamente, em cada cidade. Os pontos verdes claros representam os dados dos homens e os pontos azuis claros os dados das mulheres. Os pontos verdes escuros (para homens) e azuis escuros (para mulheres) representam os valores médios das variáveis $\log_{10}V$ versus os valores médios da variável $\log_{10}C$ em janelas ω igualmente espaçadas em $\log_{10}V$. As retas pretas contínuas representam os ajustes lineares para os valores médios dos candidatos homens e a retas pretas tracejadas representam os ajustes lineares para os valores médios das candidatas mulheres, sendo que R_M^2 e R_F^2 na inserção representam os coeficientes de determinação para os ajustes masculinos e femininos respectivamente. As barras de erro são desvios padrão do valor médio em cada janela ω .



Fonte: Cardoso (2019, p.3).

Os autores chamam a atenção para a sobreposição que tem que ser observada com cautela, tendo em vista que a escala dos gráficos é logarítmica. Mesmo com algumas oscilações, também é possível visualizar uma tendência alométrica, ou seja, um aumento de C com V seguindo uma lei de potência que foi expressa na equação 14:

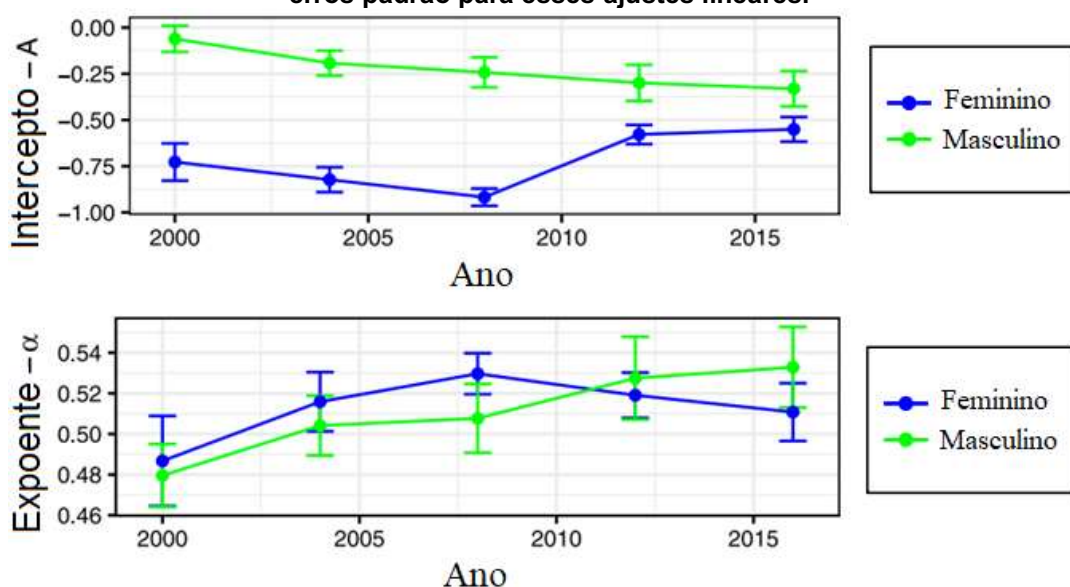
$$C \sim V^\alpha \quad (2.3.8)$$

Assim como nos dois trabalhos anteriores, o procedimento para interpretar esses resultados, foi construir janelas logaritmicamente espaçadas ω da variável V , calcular os valores médios da variável $\log_{10}V$ nas janelas ω dos correspondentes valores da variável $\log_{10}C$, o procedimento juntamente com os ajustes lineares para os valores médios é apresentado na Figura 24. Com isso é possível confirmar a tendência da lei de potência, obtendo a Equação 2.3.9:

$$\langle \log_{10}C \rangle = A + \alpha \langle \log_{10}V \rangle, \quad (2.3.9)$$

em que o símbolo $\langle \dots \rangle$ representa o valor médio em ω . Foram calculadas as variâncias em cada janela e constatado que a pouca flutuação, o que torna consistente o modelo via valores médios. Com isso a dinâmica de candidatura feminina e masculina pode ser interpretada via expoente do modelo lei de potência, expoente e intercepto da lei de potência para candidatos femininos e masculinos para cada eleição são mostrados na Figura 25.

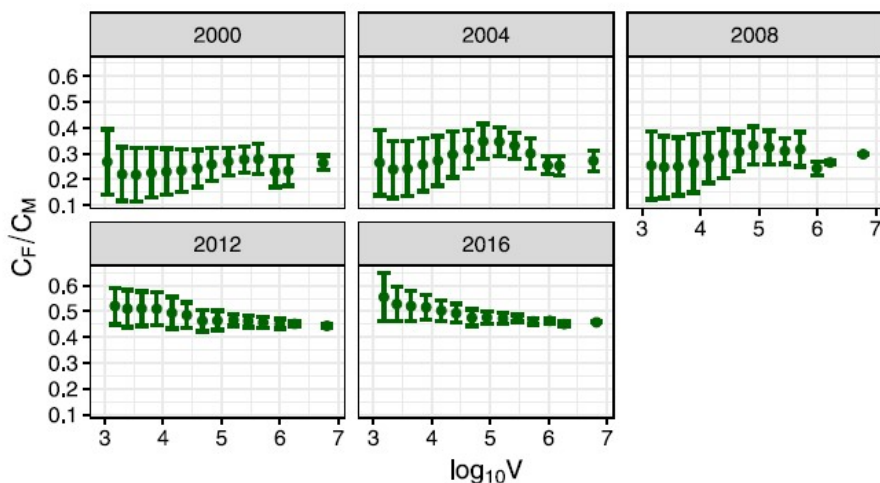
Figura 25: A evolução dos interceptos A e dos expoentes α das tendências médias das leis de potência para as candidaturas a vereador, apresentadas na Figura 24. Os símbolos verdes referem-se aos homens e os símbolos azuis escuros às mulheres. As barras de erro são os erros padrão para esses ajustes lineares.



Fonte: Cardoso (2019, p.4).

A proximidade entre os expoentes para candidatos masculinos e femininos sugerem que a diferença entre os gêneros na candidatura a vereador não aumenta com o tamanho do eleitorado. Um outro procedimento utilizando os valores médios das razões C_F/C_M nas janelas ω é mostrado na Figura 26. O fato da razão se manter aproximadamente constante em função de $\langle \log_{10} V \rangle$, confirma o resultado.

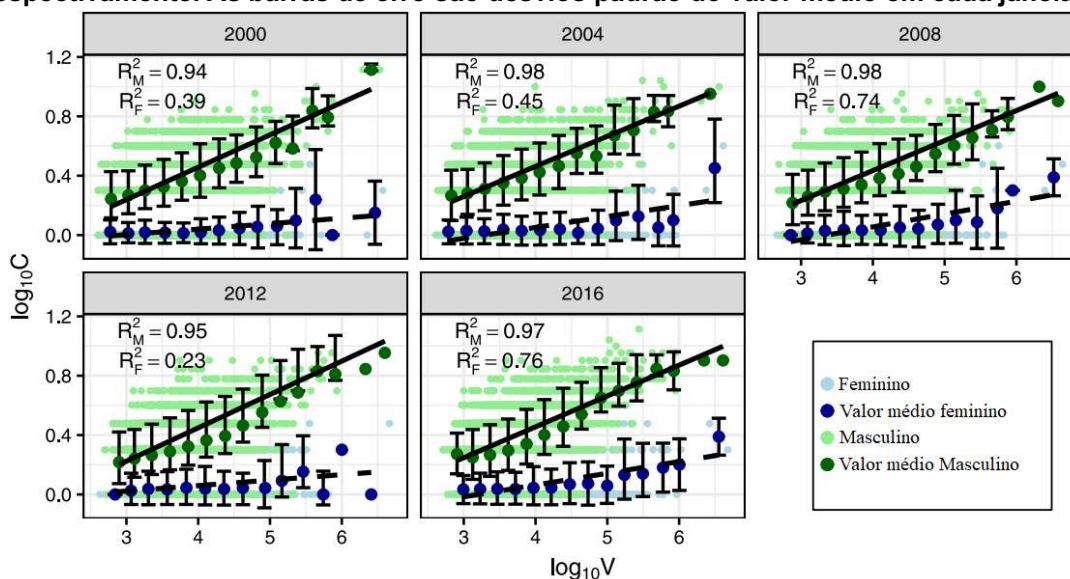
Figura 26: Valores médios da razão entre o número de candidata feminino C_F e o número de candidato masculino C_M versus $\langle \log_{10} V \rangle$, em janelas ω igualmente espaçadas de $\log_{10} V$ para eleições de vereador, onde V representa o número total de eleitores. As barras de erro são os desvios padrão do valor médio das razões nas janelas ω .



Fonte: Cardoso (2019, p.4).

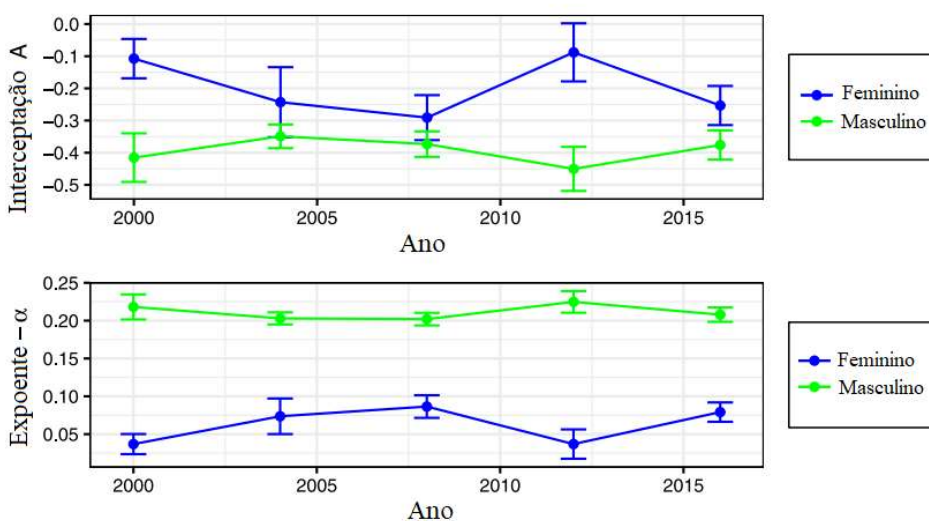
Os mesmos procedimentos utilizados para analisar as candidaturas para vereador foram aplicados para investigar o processo de candidatura para prefeito. O procedimento completo é mostrado na Figura 27. Como pode ser observado, ambas as dinâmicas de gênero na candidatura a prefeito também apresentam comportamento lei de potência. Em contraste com os dados para candidatura para vereador, os dados das candidaturas para prefeito, apresentam uma diferença nos expoentes ao longo dos anos. Também foi observado que o número médio de candidatas a prefeita é muito inferior ao número médio de candidatos homens para prefeito. A discrepância entre os gêneros aumenta consideravelmente com o tamanho do eleitorado. A dinâmica dessa discrepância pode ser mais bem observada na Figura 28, a qual disponibiliza os interceptos e expoentes das leis de potência.

Figura 27: Gráfico de dispersão das variáveis $\log_{10} V$ versus $\log_{10} C$ para as eleições para prefeito de 2000, 2004, 2008, 2012 e 2016, onde V e C representam o número de eleitores e o número de candidatos, respectivamente, em cada cidade. Os pontos verdes claros representam os dados dos homens e os pontos azuis claros os dados das mulheres. Os pontos verdes escuros (para homens) e azuis escuros (para mulheres) representam os valores médios das variáveis $\log_{10} V$ e os respectivos valores médios da variável $\log_{10} C$ em janelas ω igualmente espaçadas em $\log_{10} V$. As retas pretas contínuas representam os ajustes lineares para os valores médios dos candidatos homens e a retas pretas tracejadas representam os ajustes lineares para os valores médios das candidatas mulheres, sendo que R_M^2 e R_F^2 na inserção representam os coeficientes de determinação para os ajustes masculinos e femininos respectivamente. As barras de erro são desvios padrão do valor médio em cada janela ω .



Fonte: Adaptada de Cardoso (2019, p.5).

Figura 28: A evolução dos interceptos A e dos expoentes α das tendências médias das leis de potência para as para candidatura a prefeito, apresentadas na Figura 27. Os símbolos verdes referem-se aos homens e os símbolos azuis escuros às mulheres. As barras de erro são os erros padrão para esses ajustes lineares.



Fonte: Adaptada de Cardoso (2019, p.5).

Para o caso de eleições para vereador é possível notar que os expoentes masculinos e femininos são muito próximos, indicando que a dificuldade relativa

média para ser candidato a vereador, medida pela razão entre o número de candidatos do sexo masculino e o número de candidatas do sexo feminino, não depende muito do número de eleitores de cada cidade.

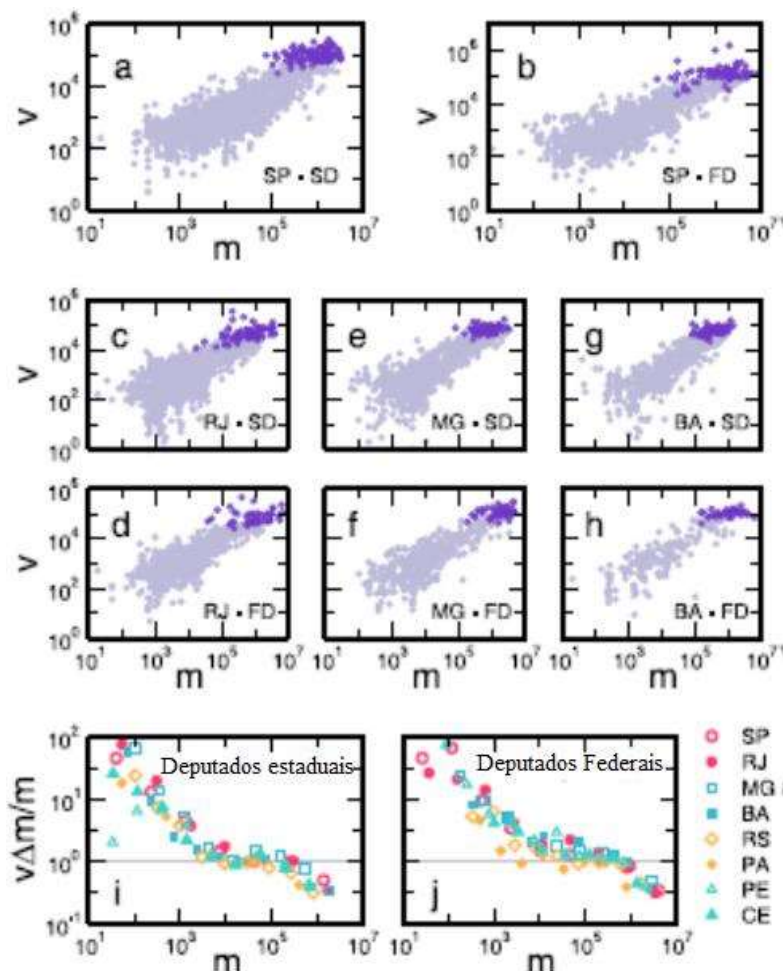
Conforme pode ser visto na Figura 28 existe uma tendência dos dados de gênero se aproximarem com o passar dos anos, revelando que o comportamento feminino é cada vez mais semelhante ao masculino para o cargo de vereador. Desse modo, Cardoso *et al.* (2019), concluíram que é notável que desde 2012 essa aproximação vem ocorrendo de forma mais rápida, em função da alteração da lei que estabelece normas para as eleições brasileiras que estabeleceram um percentual mínimo de 30% de candidaturas femininas por partido. No entanto, a discrepância na candidatura de prefeito tem marcante desvantagem para as mulheres e aumenta fortemente com o tamanho do eleitorado.

2.4 O preço do voto em eleições proporcionais brasileiras

Como os investimentos de campanha são transformados em voto era uma questão em aberto quando Melo *et al.* (2017), analisou o modelo de que forma as despesas das campanhas eleitorais influenciam no resultado da eleição. Os dados analisados são das eleições proporcionais no Brasil realizadas em 2014. O conjunto de dados é composto pelo resultado das eleições para a câmara baixa nacional e o congresso estadual e as despesas da campanha dos candidatos de todos os 26 estados brasileiros. A Figura 29 disponibiliza o número de votos v versus despesas de campanha declaradas m de cada candidato para os 4 principais estados, do Brasil, em termos de população, ou seja, São Paulo, Rio de Janeiro, Minas Gerais e Bahia. Para obter a principal relação entre v e m , tomou-se a média do número de votos v em janelas log-espaciais ao longo de m . Para representar graficamente os resultados de diferentes estados na mesma figura, foi realizada uma transformação de escala em $\langle v \rangle$, supondo uma relação linear simples $\langle v \rangle = c \times m$, em que c é uma constante característica de uma dada eleição. Definindo o preço médio de um voto como $\Delta m = \sum_i m_i / \sum_i v_i$ e supondo que é aproximadamente uniforme entre os candidatos, deve-se ter que $c = 1/\Delta m$. Onde v_i é o número de votos dos candidatos

i. Se a relação entre votos e dinheiro for linear, então o gráfico de $\langle v \rangle \times (\Delta m/m)$ deve ser uma função constante de m com valor próximo a 1,0.

Figura 29: Relação de Escala entre o número de votos e as despesas da campanha. Os círculos roxos claros mostram a relação entre o número de votos v e as despesas declaradas da campanha, m , de cada candidato a deputado estadual e deputado federal (DF) nas eleições de 2014 para os quatro maiores estados do Brasil: São Paulo (a e b), Rio de Janeiro (c e d), Minas Gerais (e e f) e Bahia (g e h). Dentro de cada painel, os dados para candidatos eleitos são destacados em círculos roxos escuros. Para ver as nuances de correlação, traçou-se em uma relação normalizada para deputados estaduais (i) e deputados federais (j) para os oito maiores estados do Brasil. Os símbolos dispostos na legenda representam a relação normalizada $\langle v \rangle \Delta m/m$, em que primeiro calculou-se o número médio de votos em janelas log-espaciações ao longo de m . Assumindo-se uma correlação linear, a constante multiplicativa é $\Delta m = M/n$. A normalização fornece uma observação direta da não linearidade na dependência dos votos sobre o dinheiro. Vê-se um comportamento sublinear global, em que os candidatos mais ricos exibem uma fração menor de votos por dinheiro.

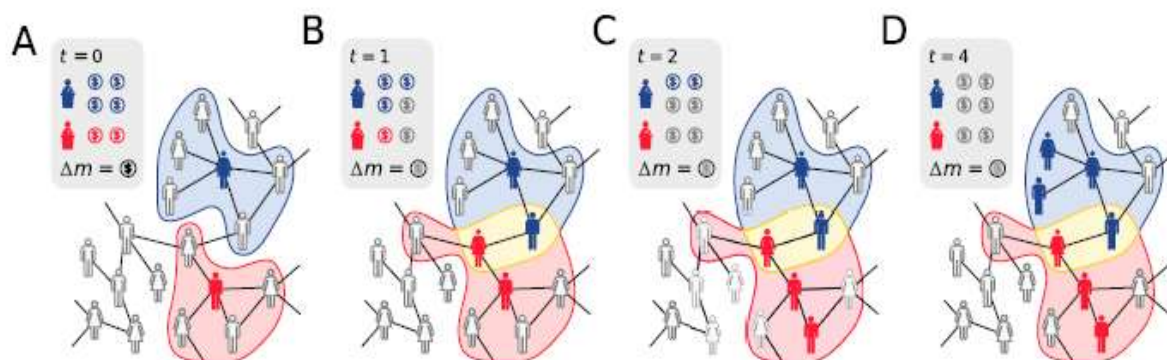


Fonte: Adaptado de Melo *et al.* (2018, p.4).

A Figura 29 i e j, mostra o gráfico de $v \Delta m/m \times m$ para deputado estadual e federal, respectivamente, para os oito estados mais populosos do Brasil no ano de 2014. O resultado mostra uma dependência não trivial entre votos e despesa de

campanha. Para pequenos valores de m , nota-se um rápido decréscimo de $v\Delta m/m$. Já para valores intermediários no intervalo $R\$10.000,00 < m < R\$100.000,00$, observa-se uma aparente dependência linear de v em relação a m . Por fim, para $m > R\$100.000,00$, observa-se um desvio perceptível da linearidade, ou seja, os candidatos mais ricos precisam de uma quantia consideravelmente grande de dinheiro para obter um único voto em comparação com os candidatos de menor sucesso na mesma faixa de despesas. Este trabalho também propôs modelar o preço de um voto, supondo um processo eleitoral composto por dois grupos distintos de indivíduos, candidatos e eleitores. Todos os candidatos s podem competir pelo voto de todos os n eleitores, sendo que cada candidato i possui uma quantia limitada de dinheiro m_i para usar em suas campanhas. Aqui assumimos que os eleitores, uma vez escolhido o seu voto, não podem mais mudar de ideia. Um esboço do modelo pode ser visto na Figura 30.

Figura 30: Esboço do Modelo para o preço do voto - Dois candidatos competem por votos em uma rede social com indivíduos indecisos (branco). (A) O candidato azul tem inicialmente um orçamento de 4 unidades monetárias, enquanto a campanha do candidato vermelho tem um orçamento de 2 unidades monetárias. Ambos os candidatos têm um eleitor decidido no início do processo. O eleitor azul na parte superior da rede e o vermelho na parte inferior da rede. As regiões destacadas incluem o eleitor inicial de cada campanha e os conhecidos de cada eleitor inicial. (B) Um eleitor indeciso dentro de cada área operacional é escolhido aleatoriamente, tornando-se então eleitores decididos. Assim, cada campanha de orçamento diminui no valor de $\Delta m = \$$. Como consequência de novos eleitores decididos, as áreas operacionais crescem, e como dois conhecidos passam a eleitores de candidatos diferentes, as áreas operacionais das campanhas agora se sobrepõem. A região representada em amarelo é onde as ambas as campanhas podem atuar. (C) Enquanto a campanha do candidato vermelho escolheu um eleitor indeciso, aumentando sua área operacional, a campanha azul gasta dinheiro em um eleitor decidido. (D) Ao final do processo de campanha, quando todas as campanhas ficam sem recursos financeiros, a campanha com maior orçamento inicial encerra o processo com mais eleitores aderidos, mas sua campanha é menos eficiente, resultando em uma não economia de escala.



Fonte: Adaptado de Melo *et al.* (2018, p.6).

A evolução do número de votos de um candidato i pode ser escrita por

$$\frac{dv_i}{dt} = \left(1 - \frac{S(t)}{n}\right) [m_i(t) > 0], \quad (2.4.1)$$

onde $S(t) = \sum_i v_i$ é o número de eleitores decididos no tempo t , e o colchete de Iverson, $[m_i(t) > 0]$, significa que é 1 se a condição dentro dos colchetes for satisfeita e 0 caso contrário. O lado direito da Equação 2.4.1 representa a probabilidade do candidato i escolher um eleitor indeciso no tempo t . Considerando que o montante de dinheiro gasto na campanha decresce linearmente, $dm_i/dt = -\Delta m_i$, e que a taxa é constante e a mesma para todos os candidatos, $\Delta m_i = \Delta m$.

Primeiramente, considerou-se o caso sem competição, em que $S \ll n$. Assumindo que $n\Delta m \gg m_i$ para todo i , para que o candidato com a maior quantia de fundos não tenha dinheiro suficiente para alcançar toda a rede. Com isso as campanhas dos candidatos não se sobreporão. Assim, $S(t)$ pode ser substituído por v_i na Equação 2.4.1, o que leva a um sistema desacoplado de equações diferenciais, com solução dada por

$$v_i = n(n - v_{0,i})e^{-m_i/n\Delta m}, \quad (2.4.2)$$

Onde $v_{0,i}$ é o número inicial de votos do candidato i . Sendo $n\Delta m \gg m_i$, e assumindo que $n - v_{0,i} \approx n$, expandindo a exponencial e tomando sua aproximação de primeira ordem, pode-se escrever o número de votos como $v_i \approx v_{0,i} + m_i/\Delta m$. Os primeiros dois regimes apresentado na Figura 29 pode ser aproximado por esse modelo simples, mas não em sua totalidade, pois no cenário sem competição, o comportamento linear permanece para m .

Considerado agora a competição entre candidatos como uma possível causa da transição do regime linear para o sublinear. Integrando Equação 2.4.1, tem-se

$$v_i = v_{0,i} + \frac{m_i}{\Delta m} - \frac{1}{n\Delta m} \int_0^{m_i} S(m') dm', \quad (2.4.3)$$

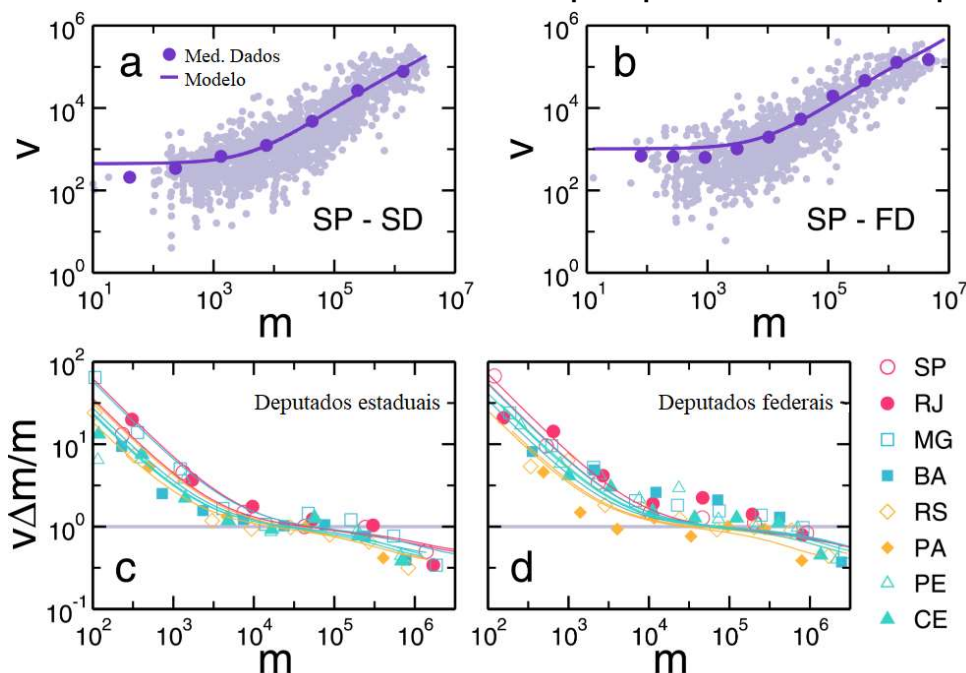
em que a integração do colchete de Iverson ao longo do tempo dá o tempo total que o candidato i tem para realizar sua campanha, $m_i/\Delta m$, e substituindo a variável de integração no último termo por $dm'/dt' = -\Delta m$. Resolvendo a Equação 2.4.1 para obter $S(m')$ e substituindo na Equação 2.4.3, encontra-se um conjunto de equações não lineares que devem ser resolvidas, para cada candidato, seguindo uma ordem

crescente de valores de m_i . Em que, o número de votos do candidato i depende da distribuição completa $P(m)$ no termo integral na Equação 2.4.3. Para obter uma solução para o modelo, foi usado o dinheiro m_i de cada candidato i , obtido a partir dos dados, o número total de eleitores n , um número inicial de votos v_0 e um valor estimado para Δm . Por simplificação foi assumido que todos os candidatos iniciam a campanha com o mesmo número de votos $v_0 = v_{0,i}$, calculado como o número médio de votos dos candidatos montante menor do que R\$1.000,00, isso pela baixa correlação entre v e m quando $m < 1.000$. O parâmetro Δm é obtido como função da taxa de comparecimento $T = S_f/n$ em que $S_f = \lim_{t \rightarrow \infty} S(t)$ é o número total de votos no estado estacionário. Assim a taxa final dos votos pode ser escrita como

$$T = 1 - e^{-\frac{M}{n\Delta m}}, \quad (2.4.4)$$

onde $M = \sum_i m_i$ é o montante total de dinheiro utilizado na campanha. Desta forma, Δm é estimado pela Equação 2.4.4 de forma que o número total de votos se ajuste aos dados da votação. A Figura 31 A e B mostra os resultados para a eleição no estado de São Paulo para deputados estaduais e federais em 2014.

Figura 31: Modelo para escala não linear. Em (A) e (B) é mostrado a eleição paulista para deputados estaduais e federais em 2014, respectivamente. Cada círculo roxo claro corresponde a um candidato e os círculos roxos escuros são os números médios de votos em janelas com espaçamento logarítmico ao longo de m . Em (C) e (D) foi realizado o mesmo processo de normalização como na Figura 29, mas agora os Δm foram estimados usando a Equação 2.4.3. Cada linha sólida mostra a solução do modelo. Apesar de sua simplicidade, o modelo apresenta todos os regimes não lineares observados nos dados, o que mostra que a ineficiência dos candidatos com mais recursos é principalmente devido à competição.



Fonte: Adaptado de Melo *et al.* (2018, p.8).

Observe que para $m < R\$ 1.000,00$, o modelo exibe um comportamento constante, capturando a natureza não correlacionada dos dados. E para $m > R\$ 1.000$, há uma correlação evidente entre votos e dinheiro, o que pode ser mais bem observado na Figura 31 c e d que mostra a razão normalizada $\langle v \rangle \Delta m / m$ para os oito estados brasileiros mais populosos. Para valores pequenos e grandes de m , o modelo desvia claramente de um comportamento linear. E mostra também uma clara sublinearidade para $m > R\$ 100.000,00$. A deseconomia de escala observada é uma consequência direta da competição entre os candidatos.

2.5 Leis de escala universal em eleições presidenciais dos Estados Unidos

O estudo abordado em Bokányi, Zoltán e Vattay (2018), analisa resultados eleitorais de grandes cidades e áreas rurais para eleições presidenciais de 1948 à 2016 nos Estados Unidos (EUA) e no referendo da União Europeia (UE) de 2016 no

Reino Unido. Encontrando curvas de escala universais e dependendo de um único parâmetro

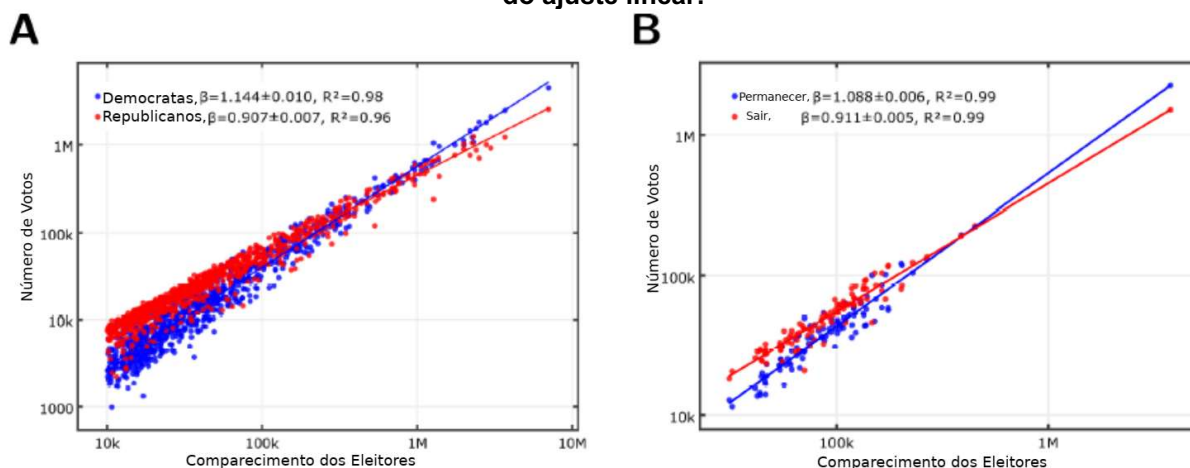
$$Y = Y_0 \cdot N^\beta. \quad (2.5.1)$$

Aqui N representa a participação do eleitor e Y os votos para os partidos (Democratas ou Republicanos) ou uma das opiniões no referendo (Permanecer ou Sair), em que uma das partes sempre mostra a escala superlinear ($\beta > 1$) e direciona o processo, enquanto o expoente sublinear ($\beta < 1$) da outra parte é meramente a consequência da conservação de probabilidade.

A Figura 32 A mostra a relação entre a taxa de comparecimento e os votos para os partidos democratas e republicanos pra 912 maiores áreas metropolitanas e micropolitanas representando cerca de 82% da população total de eleitores para as eleições presidenciais de 2016 nos EUA. A Figura 32 B mostra os votos em função da participação eleitoral nas opiniões de Permanecer e Deixar no referendo da UE de 2016 para os distritos eleitorais urbanos do Reino Unido (UK). Observe que os votos para Democratas e Permanecer na UE apresentam escala superlinear com expoentes $\beta = 1,14$ e $\beta = 1,09$, enquanto votos para Republicanos e Deixar a UE seguem escala sublinear com $\beta = 0,92$ e $\beta = 0,91$, respectivamente, com altos coeficientes de determinação $R^2 \approx 0,9$.

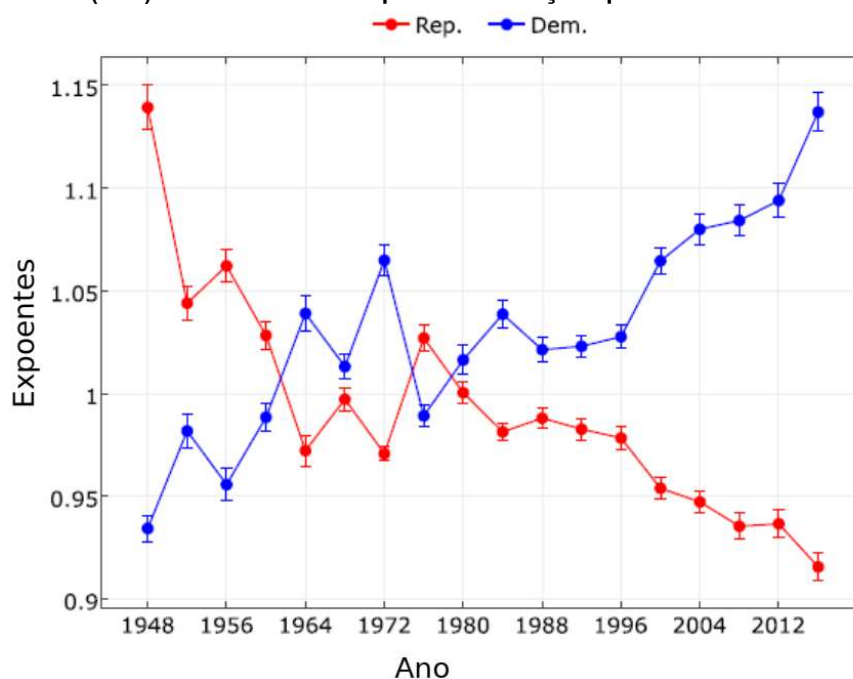
Na Figura 33, é mostrado os expoentes β_D (da relação de escala para os Democratas) e β_R (da relação de escala para os Republicanos) para as 18 eleições presidenciais no período de 1948 à 2016. Os democratas apresentam uma tendência histórica crescente, enquanto os republicanos têm uma tendência histórica decrescente.

Figura 32: Escala urbana nas eleições presidenciais nos Estados Unidos e para o referendo da União Europeia no Reino Unido - Os gráficos em log-log mostram, em (A) os votos para Republicanos (vermelho) e Democratas (azul) em função do comparecimento dos eleitores para as 912 maiores áreas estatísticas metropolitanas e micropolíticas em 2016. Em (B) os votos para Permanecer (azul) e Sair (vermelho) em função do comparecimento dos eleitores para os distritos eleitorais do Reino Unido. β é a inclinação e R^2 o coeficiente de determinação do ajuste linear.



Fonte: Adaptado de Bokányi, Zoltán e Vattay (2018, p.3).

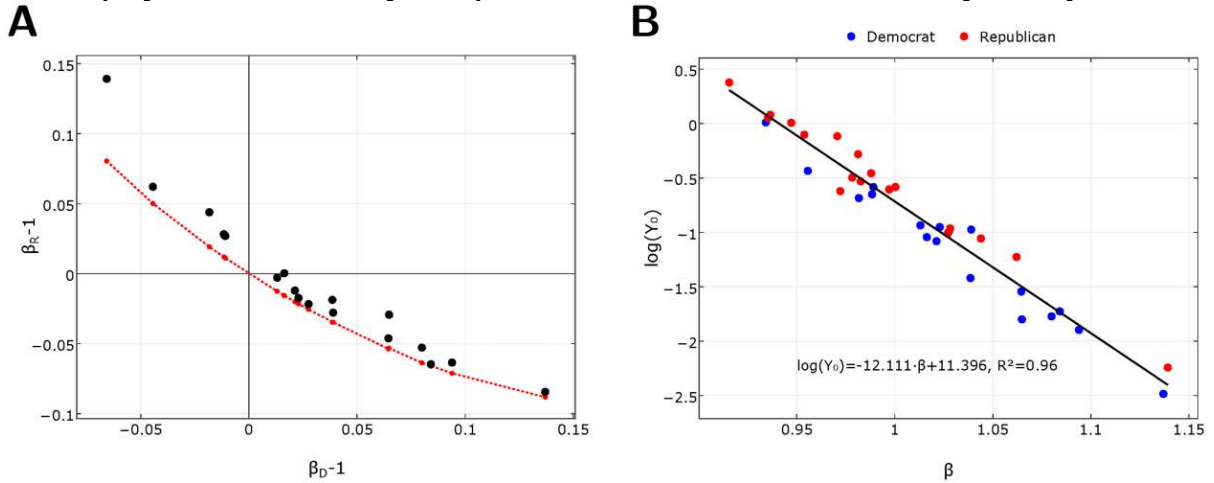
Figura 33: Evolução histórica dos expoentes - Escala dos expoentes para os Republicanos (vermelho) e Democratas (azul) com barras de erro para as 18 eleições presidenciais de 1948 à 2016.



Fonte: Adaptado de Bokányi, Zoltán e Vattay (2018, p.4).

A relação entre o expoente para republicanos e o expoente para os democratas é mostrada na Figura 34 A. O gráfico dos expoentes β versus interceptos $\log_{10} Y_0$ para cada eleição e para cada um dos dois partidos dos Estados Unidos, Republicanos e Democratas, podem ser vistas na Figura 34 B.

Figura 34: Correlação dos parâmetros de escala nas eleições dos Estados Unidos. Em (A), relação dos expoentes de escala dos Republicanos em função dos Democratas para as 18 eleições presidenciais dos EUA de 1948 a 2016 (pontos) e a curva teórica (linha vermelha) decorrente da conservação da probabilidade. Em (B), interceptos das relações $\log_{10} Y_0$ em função dos expoentes β para Republicanos (vermelho) e Democratas (azul) para eleições presidenciais no período 1948 a 2016. A reta em preto representa o ajuste linear dado pela Equação 2.5.2 e na inserção os parâmetros e coeficiente de determinação do ajuste.



Fonte: Adaptado de Bokányi, Zoltán e Vattay (2018, p.5).

Observe que esses parâmetros apresentam uma relação linear robusta com $R^2 = 0,96$ dada por

$$\log_{10} Y_0 = -\alpha\beta + \delta, \quad (2.5.2)$$

para todas as eleições e para os dois partidos. Sendo $\alpha = 12,111$ e $\delta = 11,396$. Isso significa que a Equação 20 tem a forma universal

$$Y = e^{\delta - \alpha N^*} \left(\frac{N}{N^*}\right)^\beta \approx \frac{1}{2} N^* \left(\frac{N}{N^*}\right)^\beta, \quad (2.5.3)$$

onde N é o comparecimento em uma cidade, β é o expoente do partido e $\log N^* = \alpha$. A propriedade utilizada acima é que em média no comparecimento $N = N^*$ os partidos compartilham os votos igualmente ($Y_D = Y_R = N/2$), independentemente de seus expoentes β_D e β_R ou do ano da eleição e não são afetados por mudanças históricas na população. Observe que a competição é expressa matematicamente pela conservação de probabilidade $Y_D/N + Y_R/N = 1$. Usando a Equação 2.5.3 e calculando a média para todas as cidades

$$\frac{1}{2} \langle (N/N^*)^{\beta_D - 1} \rangle + \frac{1}{2} \langle (N/N^*)^{\beta_R - 1} \rangle = 1. \quad (2.5.4)$$

Esta equação garante que um dos expoentes deve ser superlinear enquanto o outro sublinear.

3 MODELO LEI DE ESCALA PARA VOTOS BRANCOS E NULOS EM ELEIÇÕES ESTADUAIS BRASILEIRAS

Compreender melhor os efeitos da aglomeração urbana sobre os fenômenos sociais podem possibilitar o planejamento de projetos futuros e soluções para os problemas já existentes. Em vista dos sistemas complexos, tem surgido muitos estudos quantitativos de fenômenos sociais com objetivo de buscar padrões no comportamento coletivo. Entre eles os processos eleitorais são de fundamental importância dentro de uma sociedade, visto que, o futuro de seus componentes depende de tais resultados, além disso os outros sistemas sociais, em grande parte, são influenciados pelos resultados desses processos. Nesse contexto, no capítulo anterior foram apresentados alguns exemplos de trabalhos com análise de sistemas eleitorais nas diferentes fases do processo. Os trabalhos iniciais Costa e Filho *et al.* (1999, 2002), Lyra *et al.* (2003), encontraram o padrão de comportamento lei de potência para a distribuição da fração de votos em eleições proporcionais entre candidatos em eleições brasileiras, indicando que se trata de um processo multiplicativo. Ainda direcionados a analisar distribuição de votos, porém agora reescalando os votos levando em conta a relevância dos partidos as investigações em Fortunato e Castellano (2007), Araripe e Costa Filho (2009), Chatterjee *et al.* (2012), obtiveram um comportamento log-normal para um grupo de 5 países, no entanto para eleições brasileiras o procedimento obteve um ajuste exponencial.

Em uma fase anterior do processo eleitoral, a filiação partidária e a candidatura para eleição de vereador e prefeito, o comportamento obtido em Mantovani *et al.* (2011), Mantovani *et al.* (2013) identificaram uma relação lei de escala entre o número de candidatos e o tamanho da população, verificando uma hierarquia através dos expoentes, onde a posição mais influente está associada ao menor expoente. Ainda analisando candidatura, Cardoso *et al.* (2019), realiza o procedimento dos artigos anteriores para cada gênero separadamente mostrando, através dos expoentes das leis de potência, uma marcante desvantagem para as mulheres na candidatura para prefeito.

O estudo apresentado em Farhat (1996), analisa votos e gastos de campanha verificando que candidatos mais bem sucedidos pagam mais por voto do que os candidatos com menor sucesso. Em uma análise de resultado de eleições presidenciais dos Estados Unidos entre 1948-2016 e do referendo sobre a

permanência do Reino Unido na União Europeia em 2016 Bokányi, Zoltán e Vattay (2018) identificando um comportamento de escala universal com um único parâmetro, em que, sempre um dos partidos apresenta escala superlinear enquanto outro apresenta escala sublinear, consequência da conservação da probabilidade.

Uma investigação análoga foi desenvolvida em Souza (2019), para segundos turnos de eleições para governadores dos estados brasileiros, além do modelo lei de escala dependendo de um único parâmetro extraído da conservação da probabilidade entre os votos dos dois candidatos, foi possível detectar um resultado robusto entre os expoentes para cada um deles, em mais de 60% das eleições o expoente do segundo colocado é maior do o expoente do primeiro colocado, a interpretação qualitativa desse resultado é que vencer a eleição para governador dos estados é mais provável quando se investe mais esforços de campanha em menores eleitorados.

Os trabalhos usando dados da política na literatura da área de sistemas complexos, em sua grande maioria, tem focado em compreender as dinâmicas de opinião do eleitor através de análises de votos dados aos candidatos, como pode ser observado nos estudos relatados no Capítulo 2, no entanto ao observar com mais cuidado os dois últimos trabalhos citados, notamos que para utilizar com rigor a conservação da probabilidade na partição dos votos precisamos levar em conta os votos brancos e nulos.

No Brasil nosso sistema de votação é compulsório, isto é, todos os cidadãos entre 18 e 70 anos devem obrigatoriamente serem registrados como eleitores, e em cada eleição realizada no país, seja a nível municipal, estadual ou federal, é obrigatório o voto sob pena de multa. No entanto, mesmo diante disso uma proporção considerável, em média, aproximadamente 10%, dos votos são brancos ou nulos, o que nos leva a questionar se esses votos também não representam a opinião de uma parte do eleitorado.

Seguindo esse ponto de vista propomos nesse estudo uma investigação direcionada a comparar e modelar os votos branco e nulos dando a eles a devida importância. Para proceder nossas análises utilizamos dados de eleições para governadores dos estados brasileiros em segundos turnos, onde os votos são particionados em três partes, votos para o primeiro colocado, votos para o segundo colocado e votos inválidos (brancos e nulos), construímos o gráfico da relação entre o tamanho do eleitorado e o montante de votos, para cada uma das partes, nas

últimas 7 eleições estaduais brasileiras, para cada estado onde houve segundo turno, totalizando 78 eleições. Obtivemos leis de escala bem ajustadas para as três partições de votos em todas as eleições. Seguimos construindo a relação entre expoentes e interceptos das partes em todas as eleições, verificamos a relação entre os parâmetros para o primeiro colocado e o segundo colocado apresenta um único e robusto ajuste linear, no entanto os parâmetros para as leis de escala de votos brancos e nulos não pode fazer parte desse ajuste e também não são ajustados separadamente. Usando o ajuste obtido para o primeiro e o segundo colocados, juntamente com a conservação da probabilidade para as três partições de votos, construímos um modelo de um único parâmetro para todas as 78 eleições, o que sugere um comportamento padrão temporal para esse montante de votos modelando a opinião dos eleitores que aceitam as duas opções de candidatos.

3.1 Apresentação dos dados

Este estudo analisou dados eleitorais que possibilitaram compreender melhor a dinâmica de opinião em segundos turnos de eleições para governadores dos estados brasileiros. No Brasil, eleições gerais para os cargos executivos (prefeitos, governadores e presidente), acontecem a cada 4 anos.

O sistema eleitoral brasileiro determina que, deve haver uma segunda eleição, a qual chamamos segundo turno, para os três poderes executivos, sendo que, para prefeitos somente em cidades com mais de 200.000 habitantes, caso o primeiro colocado não tenha atingido na primeira votação a maioria absoluta dos votos válidos ($50\% + 1\%$), em que votos válidos se referem aos votos para os dois candidatos excluindo os votos brancos e nulos, instituído pela constituição de 1988.

No sistema eleitoral brasileiro, o voto em branco é considerado “Aquele em que o eleitor não expressa preferência por nenhum dos candidatos” Diniz (1998), antes da Constituição da República Federativa do Brasil de 1988, Brasil (1988), votos em branco eram considerados válidos e contados para o primeiro colocado, considerado como uma manifestação de insatisfação, mas de concordância com a opinião da maioria. Após a Constituição de 1988, em que constava “o candidato que obtiver a maioria dos votos válidos, excluindo brancos e nulos” é eleito, votos brancos deixam de serem votos válidos.

O sistema eleitoral brasileiro é composto praticamente em sua totalidade de urnas eletrônicas, as quais mostram para o eleitor o voto em branco como uma opção, disponibilizando um botão branco, com a escrita branco, de modo que é pouco provável que o eleitor vote em branco por engano. Já o voto nulo ocorre quando um eleitor insere um número que não corresponde a uma opção de candidato, e no caso de urnas tradicionais com cédulas de papel, quando faz algum tipo de marcação que impossibilita identificar o candidato ou o partido escolhido Farhat (1996). Portanto, desde 1988, votos brancos e nulos são calculados apenas como estatísticas para as eleições. No entanto, eles podem representar, em especial os votos em branco, a opinião de eleitores.

Consideramos nesse trabalho, juntamente, votos brancos e nulos como uma manifestação de insatisfação do eleitor, e dando a devida importância a opinião desses eleitores propomos aqui analisar e modelar esse montante de votos em função da população de eleitores.

O banco de dados utilizado nesse estudo foi composto por resultados de eleições para governadores dos estados brasileiros em segundos turnos, que ocorreram nos anos e para os números de estados: 1994 - 11, 1998 - 11, 2002 - 13, 2006 - 10, 2010 - 12, 2014 - 13 e 2018 - 13, contemplando um total de 78 eleições, disponibilizados no sítio¹ do Tribunal Superior Eleitoral – TSE. Foram escolhidos os segundos turnos por apresentarem um cenário em que os votos são divididos em três partes, votos para o primeiro e segundo colocados e os votos brancos e nulos (contados conjuntamente), o que possibilitou relacioná-las pela conservação de probabilidades entre as três partes, permitindo a abordagem do modelo para os votos brancos e nulos. Além disso optamos por eleições de governadores por englobar uma quantidade considerável de eleições, diferente do que ocorre para eleições de presidente, e no caso de eleições para prefeito seria uma possibilidade para similar futuramente.

Para facilitar a identificação dos estados na descrição das análises, a Tabela 1 e o Gráfico 1 mostra os nomes, as siglas e a região onde estão localizados os 27 estados do Brasil. A Tabela 2 apresenta os estados e os anos onde houve segundo turno para eleição de governador. É possível observar um fato interessante, no Rio

¹ Disponível em: <https://www.tse.jus.br/>. Acesso em 20 set. 2020.

Grande do Sul - RS ocorreram segundos turnos em todos os anos de eleição, sugerindo que a população tende a ter opinião fragmentada, enquanto nos estados do Espírito Santo - ES, Mato Grosso - MT e Tocantins - TO não ocorreu segundo turno em nenhum dos anos o que permite interpretar que a maioria da população nesses estados tende a compartilhar da mesma opinião.

Tabela 1 - Estados brasileiros, siglas e região

Estado do Brasil	Sigla	Região de Localização	Estado do Brasil	Sigla	Região de Localização
Acre	AC	Norte	Pará	PA	Norte
Alagoas	AL	Nordeste	Paraíba	PB	Nordeste
Amapá	AP	Norte	Paraná	PR	Sul
Amazonas	AM	Norte	Pernambuco	PE	Nordeste
Bahia	BA	Nordeste	Piauí	PI	Nordeste
Ceará	CE	Nordeste	Rio de Janeiro	RJ	Sudeste
Distrito Federal	DF	Centro-Oeste	Rio Grande do Norte	RN	Nordeste
Espírito Santo	ES	Sudeste	Rio Grande do Sul	RS	Sul
Goiás	GO	Centro-Oeste	Rondônia	RO	Norte
Maranhão	MA	Nordeste	Roraima	RR	Norte
Mato Grosso	MT	Centro-Oeste	Santa Catarina	SC	Sul
Mato Grosso do Sul	MS	Centro-Oeste	São Paulo	SP	Sudeste
Minas Gerais	MG	Sudeste	Sergipe	SE	Nordeste
			Tocantins	TO	Norte

Fonte: Autoria própria (2021).

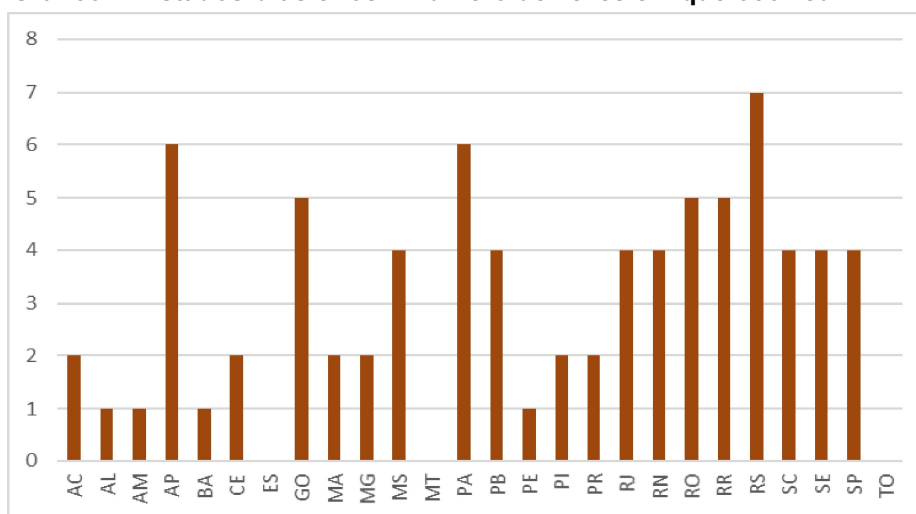
Tabela 2 - Anos e estados que tiveram segundo turno nas eleições para governador. As cédulas marcadas com X representam os estados e anos onde ocorreu segundo turno em eleições para governadores dos estados brasileiros.

UF	1994	1998	2002	2006	2010	2014	2018	TOTAL
AC	X					X		2
AL					X			1
AM							X	1
AP	X	X	X		X	X	X	6
BA	X							1
CE			X			X		2
ES								0
GO	X	X		X	X	X		5
MA	X			X				2
MG		X					X	2

Continuação da Tabela 2									
MS		X	X				X	X	4
MT									0
PA		X	X	X	X	X	X	X	6
PB			X	X	X	X			4
PE					X				1
PI	X					X			2
PR			X	X					2
RJ		X		X			X	X	4
RN			X	X			X	X	4
RO		X	X			X	X	X	5
RR	X	X	X				X	X	5
RS	X	X	X	X	X	X	X	X	7
SC	X		X	X				X	4
SE	X	X	X					X	4
SP	X	X	X					X	4
TO									0
TOTAL	11	11	13	10	8	12	13	78	

Fonte: Autoria própria (2021).

Gráfico 1: Estados brasileiros x Número de vezes em que ocorreu



Fonte: Autoria própria (2021).

De forma mais clara, as variáveis filtradas para serem analisadas nesse trabalho foram, número total de votos, votos para o primeiro colocado, votos para o segundo colocado e votos brancos e nulos em cada cidade onde ocorreu segundo turno da eleição para governador nas 7 eleições da base de dados, como descrito acima. Com o objetivo de apresentar uma ideia generalizada desses dados, a Tabela 3 apresenta os dados totais referentes ao total de votos, porcentagens de votos para o primeiro e o segundo colocados e de votos brancos e nulos em cada ano de eleição e cada estado. Uma observação interessante a ser feita, é que em 2018 a proporção de votos brancos e nulos foi consideravelmente mais alta que nos outros anos. De fato, tomando a média entre todos os estados onde houve segundo turno em cada ano temos: 1994-7,11; 1998 – 4,51; 2002 – 6,31; 2006 – 9,61; 2010 – 7,69; 2014 – 8,61; 2018 -12,08.

Tabela 3 – Tabela do total de votos e das porcentagens de votos para o primeiro colocado, o segundo colocado e votos brancos e nulos em cada estado onde ocorreu segundo turno nas eleições para governadores nas 7 eleições da base de dados, 1994, 1998, 2002, 2006, 2010, 2014 e 2018. Aqui as variáveis Ano, Sigla UF, Total_Votos, %Votos_Primeiro, % Votos_Segundo e %Votos_BN, representam, o ano que ocorreu a eleição, a sigla no estado onde ocorreu, o total de votos no ano e estado de acordo com a posição, a porcentagem de votos para o candidato que ficou em primeiro lugar, a porcentagem de votos do candidato que ficou segundo lugar e a porcentagem de votos brancos e nulos, respectivamente.

Ano	Sigla UF	Total_Votos	% Votos_Primeiro	% Votos_Segundo	% Votos_BN
1994	AC	184.699	49,809	43,008	7,182
1994	AP	131.534	53,147	43,728	3,125
1994	BA	4.368.192	51,18	36,103	12,717
1994	GO	1.959.722	51,701	39,938	8,361
1994	MA	1.580.238	47,708	46,565	5,727
1994	PI	1.182.923	52,07	41,223	6,707
1994	RR	85.890	55,044	39,104	5,853
1994	RS	5.405.386	49,575	45,384	5,041
1994	SC	2669948	48,242	46,726	5,032
1994	SE	760633	48,878	45,704	5,419
1994	SP	17748536	48,804	38,152	13,044
1998	AP	178666	52,433	45,404	2,163
1998	GO	2247614	51,521	45,174	3,305
1998	MG	8990598	53,485	39,346	7,168
1998	MS	935280	58,596	37,044	4,36
1998	PA	1917481	51,182	43,799	5,019
1998	RJ	7702924	55,295	40,077	4,628
1998	RO	523049	51,357	44,466	4,177
1998	RR	127783	53,128	44,882	1,989
1998	RS	5748205	49,49	47,97	2,541
1998	SE	817124	50,788	42,586	6,626
1998	SP	19174067	51,112	41,205	7,683
2002	AP	237820	52,909	44,052	3,039
2002	CE	3699233	47,732	47,65	4,618
2002	MS	1119145	51,963	44,725	3,311
2002	PA	2588852	49,871	46,555	3,575
2002	PB	1894314	46,979	44,508	8,513
2002	PR	5428354	49,404	40,176	10,42
2002	RN	1520599	53,962	34,435	11,604
2002	RO	641443	54,265	37,604	8,131
2002	RR	170766	50,968	44,315	4,717
2002	RS	6278208	50,154	45,069	4,777
2002	SC	3238941	46,696	46,056	7,248
2002	SE	945260	50,935	41,678	7,387
2002	SP	21493289	55,872	39,412	4,716

Fonte: Autoria própria (2021).

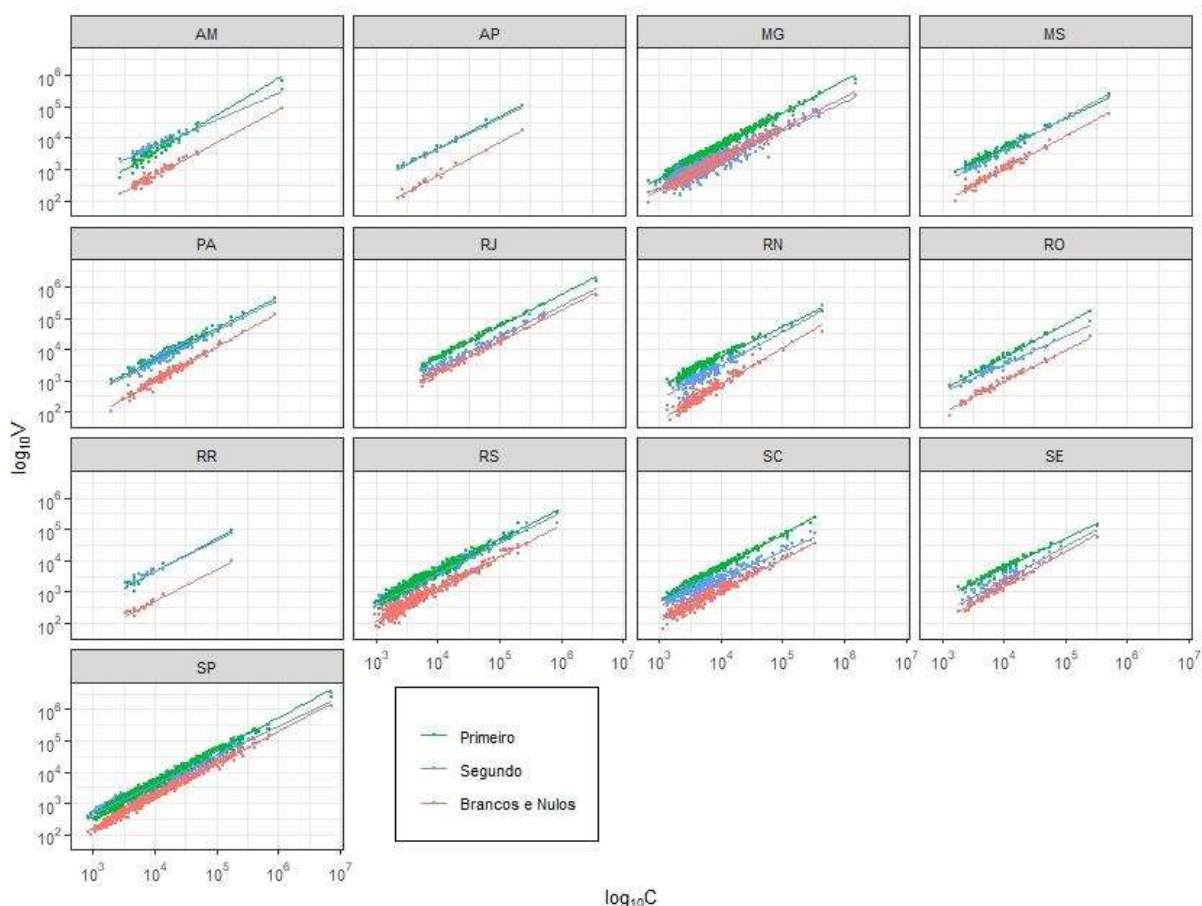
3.2 Análise dos dados

3.2.1 Obtenção das leis de escala

Este estudo deu início às investigações por construir os gráficos para a relação entre o número total de eleitores que compareceram para votar (denotado por C) versus número de votos para o primeiro colocado, votos para o segundo colocado e votos brancos e nulos (denotados por V). Os gráficos estão representados por cidade, para as duas variáveis em logaritmo na base 10, em cada estado onde ocorreu segundo turno para todas os 7 anos de eleições da base de dados (eleições de 1994, 1998, 2002, 2006, 2010, 2014 e 2018), o que representa um total de 78 eleições.

O procedimento mostrou relações aproximadamente leis de potências para todas as 78 eleições e para as três relações ($C \times V$) em cada eleição, totalizando 234 relações leis de potência. A Figura 35 apresenta um painel com as 13 eleições ocorridas em 2018. Escolhemos mostrar o procedimento par apenas um ano de eleição, dado que os resultados para todos os anos foram muito semelhantes, o ano de 2018 por ser um entre os dois anos que tiveram o maior número de estados com segundo turno, ou seja, com maior número de eleições.

Figura 35: Gráfico de dispersão da relação entre as variáveis: número de pessoas que compareceram para votar C versus o número de votos para o primeiro colocado, versus número de votos para o segundo colocado, versus número de votos brancos e nulos, todos os montantes de votos representados por V . Dados dos segundos turnos das eleições para governador, em todos os estados onde ocorreu segundo turno, em 2018. Cada painel representa um estado e é identificado pela sigla dos estados.



Fonte: Autoria própria (2021).

Tabela 4 - Tabela dos coeficientes de determinação R^2 pela Equação 3.2.2.1: para todos os ajustes lineares relacionando as variáveis C , comparecimento, e V , votos para o primeiro colocado ou votos para o segundo colocado ou votos brancos e nulos, obtidos para as 78 eleições da base de dados. Aqui Ano, Sigla da UF, R^2 Primeiro, R^2 Segundo e R^2 Brancos e Nulos, representam, o ano em que ocorreu a eleição, a sigla referente ao estado onde houve o segundo turno, o coeficiente de determinação para o ajuste linear relacionando os votos do primeiro colocado, o coeficiente de determinação para o ajuste linear relacionando os votos do segundo colocado, o coeficiente de determinação para o ajuste linear relacionando os votos brancos e nulos, respectivamente.

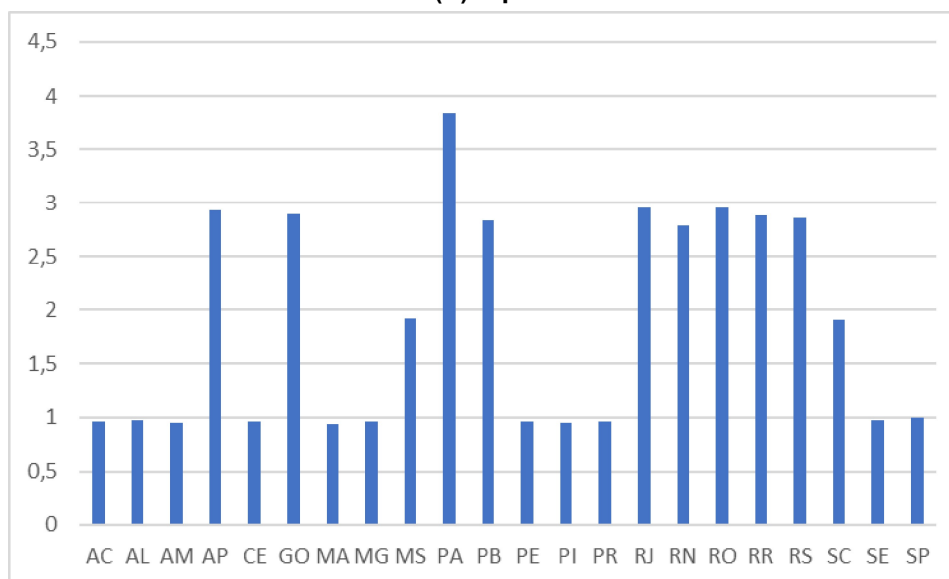
Ano	Sigla da UF	R² Primeiro	R² Segundo	R² Brancos e Nulos
1994	AC	0,96	0,95	0,95
1994	AP	0,99	0,98	0,94
1994	BA	0,94	0,87	0,77
1994	GO	0,97	0,93	0,92
1994	MA	0,93	0,91	0,9
1994	PI	0,91	0,9	0,85
1994	RR	1	0,99	0,98
1994	RS	0,98	0,96	0,93
1994	SC	0,97	0,97	0,92
1994	SE	0,96	0,94	0,88
1994	SP	0,98	0,98	0,96
1998	AP	0,99	0,96	0,96
1998	GO	0,97	0,96	0,92
1998	RR	0,98	0,96	0,93
1998	RS	0,95	0,97	0,93
1998	SE	0,97	0,95	0,82
1998	SP	0,98	0,98	0,96
1998	MG	0,93	0,86	0,89
1998	MS	0,97	0,95	0,92
1998	PA	0,93	0,85	0,9
1998	RJ	0,98	0,93	0,95
1998	RO	0,96	0,95	0,94
2002	AP	0,99	0,99	0,96
2002	RR	0,99	0,98	0,97
2002	RS	0,98	0,96	0,94
2002	SC	0,96	0,97	0,95
2002	SE	0,97	0,92	0,94
2002	SP	0,99	0,98	0,97
2002	MS	0,97	0,96	0,95
2002	PA	0,94	0,93	0,93
2002	RO	0,96	0,89	0,97
2002	CE	0,94	0,92	0,95
2002	PB	0,87	0,88	0,92
2002	PR	0,95	0,94	0,95
2002	RN	0,91	0,8	0,92

Continuação da Tabela 4

Ano	Sigla da UF	R^2	Primeiro	R^2	Segundo	R^2	Branco e Nulos
2006	GO		0,98		0,96		0,96
2006	MA		0,87		0,89		0,93
2006	PA		0,96		0,95		0,95
2006	PB		0,92		0,89		0,93
2006	PE		0,96		0,81		0,96
2006	PR		0,95		0,93		0,96
2006	RJ		0,99		0,96		0,99
2006	RN		0,9		0,88		0,92
2006	RS		0,98		0,96		0,95
2006	SC		0,98		0,97		0,97
2010	AP		0,99		0,98		0,98
2010	AL		0,93		0,89		0,97
2010	GO		0,98		0,95		0,96
2010	PA		0,96		0,93		0,95
2010	PB		0,95		0,93		0,96
2010	PI		0,96		0,87		0,95
2010	RO		0,97		0,96		0,99
2010	RR		0,99		0,99		0,94
2014	AC		0,98		0,97		0,96
2014	AP		0,99		0,96		0,97
2014	CE		0,95		0,92		0,96
2014	GO		0,98		0,94		0,97
2014	MS		0,97		0,97		0,95
2014	PA		0,91		0,92		0,96
2014	PB		0,92		0,92		0,95
2014	RJ		0,98		0,95		0,99
2014	RN		0,94		0,92		0,94
2014	RO		0,97		0,97		0,98
2014	RR		0,98		0,97		0,97
2014	RS		0,98		0,93		0,96
2018	AP		0,99		0,99		0,98
2018	AM		0,93		0,94		0,95
2018	MG		0,98		0,88		0,96
2018	MS		0,96		0,96		0,97
2018	PA		0,97		0,94		0,97
2018	RJ		0,99		0,97		0,98
2018	RN		0,96		0,86		0,93
2018	RO		0,99		0,96		0,98
2018	RR		0,93		0,94		0,97
2018	RS		0,97		0,92		0,95
2018	SC		0,97		0,91		0,94
2018	SE		0,97		0,9		0,97
2018	SP		0,98		0,97		0,99

Fonte: Autoria própria (2021).

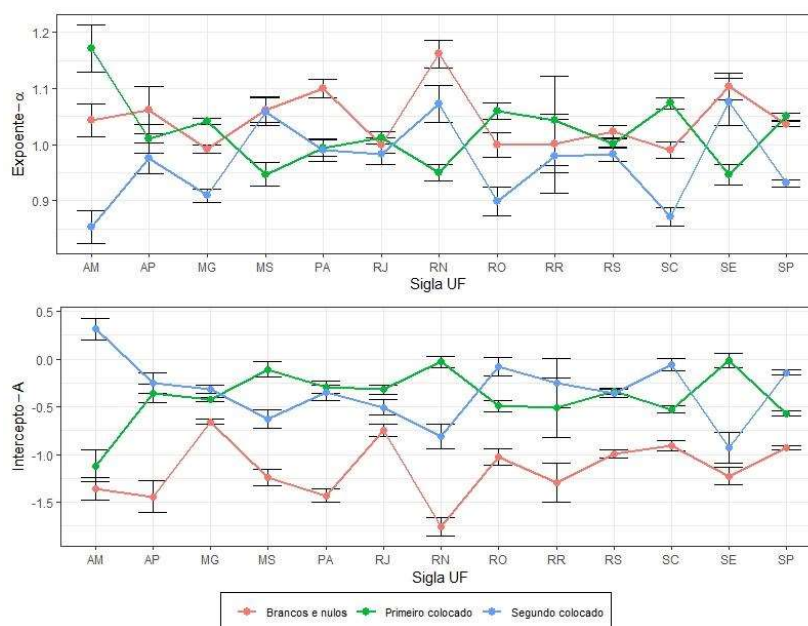
Gráfico 2 - Coeficiente de determinação R^2 para todos os ajustes lineares relacionando as variáveis C , comparecimento, e V , votos para o primeiro colocado ou votos para o segundo colocado ou votos brancos e nulos, obtidos para as 78 eleições da base de dados. Eixo (x), representa os estados brasileiros. Eixo(Y) representa o coeficiente de determinação R^2 .



Fonte: Autoria própria (2021).

No entanto, a Tabela 4 e o Gráfico 2 mostra os coeficientes de determinação R^2 para todos os 234 ajustes. Como pode ser visto todos são maiores ou igual a 0,77, sendo que, para mais de 90% dos casos $R^2 > 0,9$, o que confirma a boa qualidade dos ajustes lineares. Além disso o p-valor obtidos para todos os ajustes foram muito próximos a 0, menores do que $3,7 \times 10^{-6}$. Os expoentes e interceptos dos ajustes mostrados na Tabela 4 são disponibilizados na Figura 36. Note que, os pontos em vermelho representando o expoente para os votos brancos e nulos, painel superior, das 13 eleições em 9 delas o expoente de ajuste de brancos e nulos é maior do que 1, ou seja, a escala é super linear, mostrando que esse montante de votos aumenta em maior proporção nos maiores eleitorados. α representa os expoentes relacionados com os 13 estados. Os interceptos correspondentes podem ser vistos no painel inferior, representados por A.

Figura 36: Expoentes e Interceptos dos ajustes mostrados na Figura 35. O painel superior mostra os expoentes relacionados com os 13 estados onde ocorreu o segundo turno para as eleições de governadores no ano de 2018, o painel inferior mostra, da mesma forma, os interceptos correspondentes, Equação 3.2.2.1.



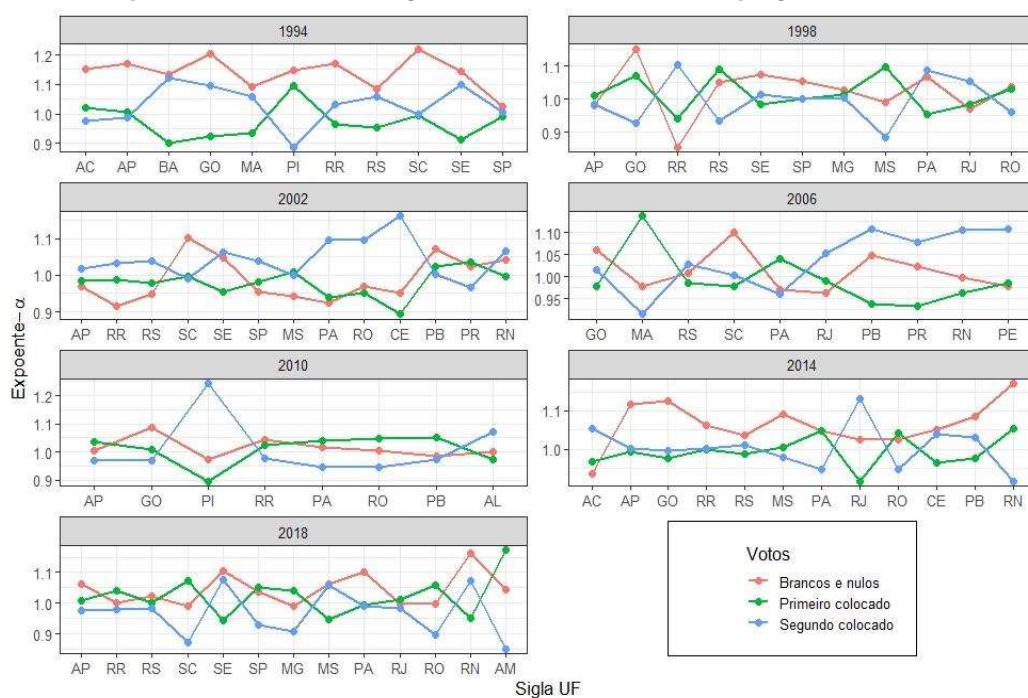
Fonte: Autoria própria (2021).

Esse resultado sugere uma questão, será que: à medida que o tamanho do eleitorado aumenta, tende a aumentar as proporções de votos brancos e nulos? A resposta é: provavelmente sim, veja o total de expoentes em cada ano disponibilizados na Figura 36. Para as 78 eleições tem-se escalas super lineares em 44% para as relações dos votos do primeiro colocado, em 56% dos votos do segundo colocado e em 70% dos votos brancos. Mostrando que o segundo colocado tem na maioria das eleições sua proporção de votos aumentada com o aumento do tamanho do eleitorado, enquanto isso ocorre na minoria delas para o primeiro colocado. Isso é um indicativo de que os candidatos vencedores tendem a concentrar seus esforços de campanha em eleitorados menores. No estudo Souza (2019), foi verificado, usando conservação de probabilidade e desconsiderando brancos e nulos, que quando a escala do primeiro colocado é super linear a escala do segundo é sublinear e vice-versa, sendo que a soma dos dois expoentes é sempre muito próxima a 2.

Esse resultado continua válido mesmo levando em conta brancos e nulos, no entanto aqui foi obtida uma relação de proporções entre os expoentes das três partições de votos cujo resultado é aproximadamente 3 para todas as eleições, resultado que será verificado analiticamente, na seção 3.2.3, usando conservação

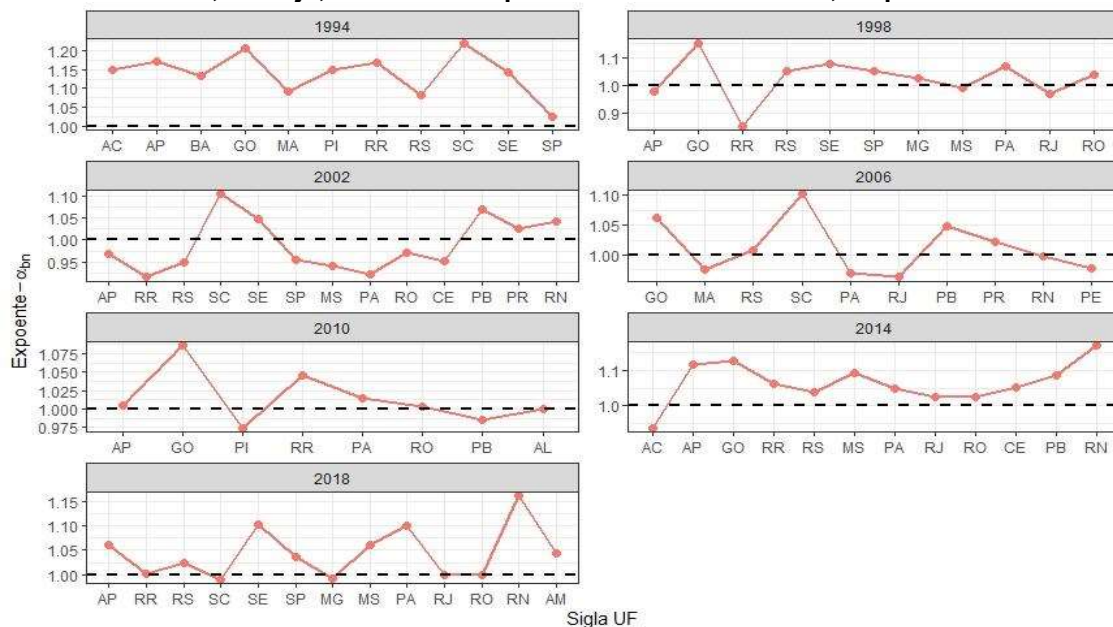
de probabilidade. Nesse procedimento se destaca, em especial, o fato dos expoentes para as relações de brancos e nulos serem em grande parte das eleições. 70% delas como mencionado acima, maiores do que 1, a Figura 38 mostra os expoentes somente para os ajustes brancos e nulos, separadamente, comparados a reta vertical constante igual a 1. Pode ser observado com clareza essa proporção de relações super lineares. Esse resultado mostra que, à medida que aumenta o eleitorado tende a aumentar a proporção desse montante de votos, o que leva a interpretar que eleitores em grandes eleitorados são mais seletivos ou mais exigentes na hora de escolher seus representantes.

Figura 37: Expoentes dos ajustes lineares para as três partições de votos: O painel acima mostra, para cada estado brasileiro onde ocorreu a eleição, os expoentes dos ajustes lineares obtidos para as três relações $\log_{10} C$ versus $\log_{10} V$, em que C é o comparecimento e V os votos: para o primeiro colocado, o segundo colocado e votos brancos e nulos, por cidade, para os 7 anos de eleições da base de dados. Equação 3.2.2.1.



Fonte: Autoria própria (2021).

Figura 38: Expoentes dos ajustes lineares dos votos brancos e nulos: O painel acima mostra, para cada estado brasileiro onde ocorreu a eleição, os expoentes dos ajustes lineares obtidos para as relações $\log_{10} C$ versus $\log_{10} V$, em que C é o comparecimento e V os votos brancos e nulos, por cidade, para os 7 anos de eleições da base de dados. A reta em preto pontilhada em cada gráfico é a reta horizontal constante em 1, possibilitando visualizar os expoentes acima e abaixo dela, ou seja, as escalas super lineares e sublineares, respectivamente.

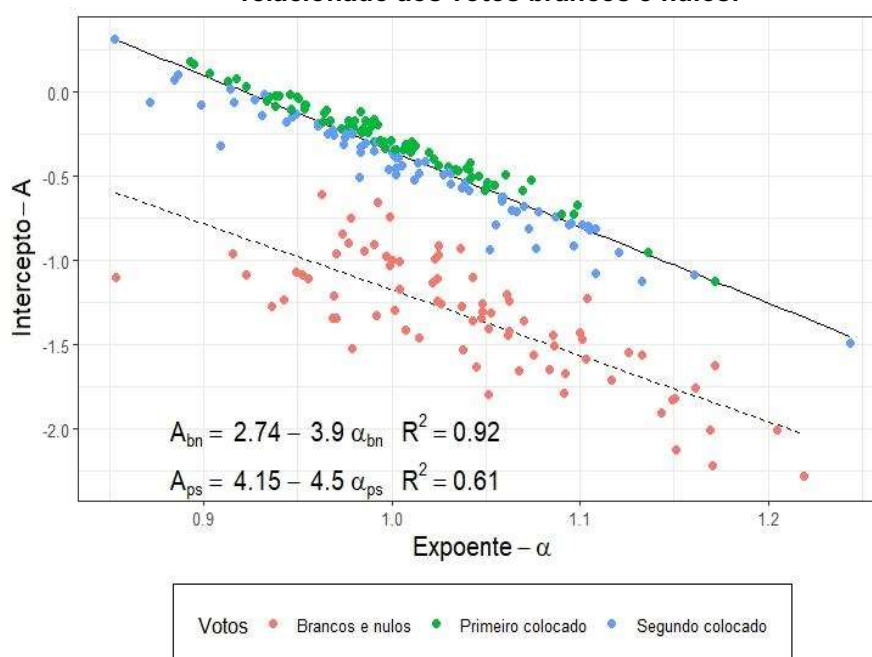


Fonte: Autoria própria (2021).

Sejam α_p , α_s e α_{bn} os expoentes relativos às escalas dos votos para o primeiro colocado, o segundo colocado e votos brancos e nulos, respectivamente. A Figura 39 mostra o gráfico da relação entre os parâmetros, expoentes e interceptos, dos ajustes para cada uma das três partições de votos, para as 78 eleições da base de dados, procedimento análogo ao usado nos trabalhos Bokányi, Zoltán e Vattay (2018).

Como pode ser observado, os parâmetros das relações de votos do primeiro e do segundo colocados é obtido um único e robusto ajuste, a equação da reta e o coeficiente de determinação são mostrados na inserção do gráfico, veja que $R^2 = 0,92$. No entanto, os modelos das relações de escala de votos brancos e nulos, visivelmente, não podem ser ajustados na primeira relação, juntamente com os parâmetros do primeiro e do segundo colocados, e nem mesmo separadamente como pode também ser visto na inserção do gráfico que mostra a equação e o coeficiente de determinação $R^2 = 0,61$.

Figura 39: Gráfico dos expoentes versus interceptos das leis de escala para as três partições de votos. Os pontos no gráfico acima representam a relação entre o intercepto A_k e o expoente α_k para as três partições de votos, votos para o primeiro colocado, para o segundo colocado e brancos e nulos, nas 78 eleições da base de dados. A equação e o coeficiente de determinação para o ajuste dos parâmetros obtido para as partições de votos do primeiro, juntamente, com os do segundo colocado, reta contínua em preto, é mostrada no inserte, onde o índice p e s representa essas partições, respectivamente. São também mostrados para os ajustes dos parâmetros referentes aos brancos e nulos, reta em preto tracejada, onde o índice bn está relacionado aos votos brancos e nulos.



Fonte: Autoria própria (2021).

3.2.2 Modelo analítico para os votos brancos e nulos

Dos procedimentos realizados até aqui foram obtidas relações de escalas relacionando votos, V , e comparecimento, C , para as três partições de votos e para todas as 78 eleições da base de dados, ou seja, 234 relações leis de potência no

total. Inicialmente foram observadas através dos ajustes lineares para os gráficos $\log_{10}C$ versus $\log_{10}V$, isto é, obteve-se as relações

$$\log_{10}V = A_{j,k} + \alpha_{j,k}\log_{10}C, \quad (3.2.2.1)$$

em que, $A_{j,k}$ e $\alpha_{j,k}$ são os parâmetros dos ajustes para $j = p, s, bn$, variando entre as três partições de votos (para o primeiro colocado, segundo colocado e brancos e nulos, respectivamente), e k variando entre as 78 eleições que compõem a base de dados (os 7 anos de eleições, 1994, 1998, 2002, 2006, 2010, 2014, 2018 para cada estado onde houve segundo turno), obtendo-se um total de 234 ajustes. Tomando a exponencial de base 10 na Equação 3.2.2.1, segue-se as relações leis de potência para os dados

$$V \approx a_{j,k}C^{\alpha_{j,k}}, \quad (3.2.2.2)$$

para $a_k = 10^{A_k}$ e k percorrendo os 234 ajustes. Da relação entre os parâmetros para as duas partições de votos, para o primeiro e o segundo colocados, mostradas na Figura 39, tem-se a equação

$$A_{i,k} = 4,15 - 4,5 \alpha_{i,k}, \quad (3.2.2.3)$$

sendo $i = p, s$ e k variando sobre as 78 eleições da base de dados, ou seja, a equação acima relaciona os parâmetros de 156 ajustes, não percorrendo os ajustes para votos brancos e nulos. Os parâmetros para esse montante de votos são razoavelmente ajustados, separadamente, pela equação $A_{bn} = 2,74 - 3,9 \alpha_{bn}$, assim A_{bn} e α_{bn} representam os coeficientes dos ajustes obtidos para as relações entre o comparecimento e os votos brancos e nulos nas 78 eleições.

Substituindo a Equação 3.2.2.3 em 3.2.2.1 e aplicando a exponencial de base 10, tem-se a equação

$$V \approx 0,45 C^* \left(\frac{C}{C^*} \right)^{\alpha_{i,k}}, \quad (3.2.2.4)$$

considerando $C^* = 10^{4,5}$, para a relação entre o comparecimento C e os votos V , representando o primeiro e o segundo colocados para os seus respectivos expoentes relacionados a cada eleição da base de dados ($i = p, s$ e $k = 1, \dots, 78$).

Para obtenção do modelo, é conveniente separar a notação para as três partições de votos, por isso foi tomado V_p , V_s e V_{bn} os números de votos para o primeiro colocado, para o segundo colocado e brancos e nulos, por cidade,

respectivamente. A conservação da probabilidade para relação entre os votos é dada por

$$\frac{V_p}{C} + \frac{V_s}{C} + \frac{V_{bn}}{C} = 1. \quad (3.2.2.5)$$

Uma equação aproximada a partir da Equação (3.2.2.5) quando substituimos os votos pelas respectivas leis de escala descritas pela Equação aproximada (3.2.2.4)

$$0,45 C^* \left(\frac{C}{C^*}\right)^{\alpha_{p,k}} + 0,45 C^* \left(\frac{C}{C^*}\right)^{\alpha_{s,k}} + V_{bn} \approx C. \quad (3.2.2.6)$$

Para uma cidade com $C = C^*$ a relação acima determina que $V_{bn} \approx 0,1 C^*$, por outro lado, pelo modelo lei de potência da Equação 3.2.2.2, $V_{bn} \approx a_{bn,k} C^{*\alpha_{bn,k}}$, comparando os dois resultados a constante do modelo deve ser da forma $a_{bn} \approx 0,1 C^{*(1-\alpha_{bn,k})}$, com esse resultado o modelo lei de escala para votos brancos e nulos pode ser escrito como:

$$V \approx 0,1 C^* \left(\frac{C}{C^*}\right)^{\alpha_{bn,k}}. \quad (3.2.2.7)$$

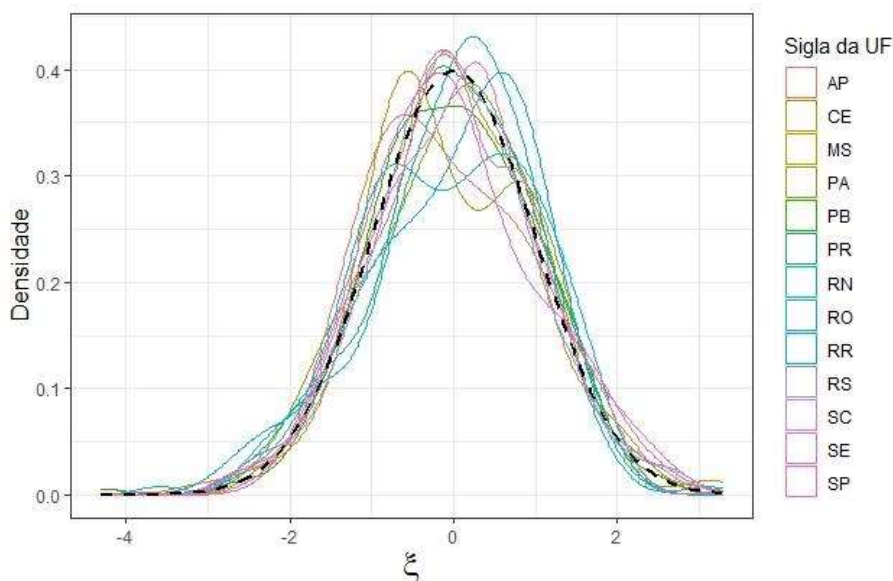
O modelo acima representa o principal resultado obtido nesse trabalho.

Com o objetivo de construir um modelo completo foram analisadas as flutuações em torno dos ajustes lineares descritos pela Equação 3.2.2.1. Seja ξ a função que calcula a flutuação em torno do modelo linear dada por

$$\xi = \log_{10} V - A_{j,k} + \alpha_{j,k} \log_{10} C. \quad (3.2.2.8)$$

A Figura 40 mostra a densidade de ξ , normalizada, para eleições de 2002, ou seja, as flutuações dos ajustes mostrados na Figura 35.

Figura 40: Flutuações em torno do modelo linear para os dados de 2002. O gráfico acima apresenta as densidades normalizadas da varável ξ as 13 eleições brasileiras de governadores de estados onde ocorreu segundo turno no ano de 2002. A curva tracejada em preto representa a distribuição normal de média $\mu = 0$ e variância $\sigma^2 = 1$.



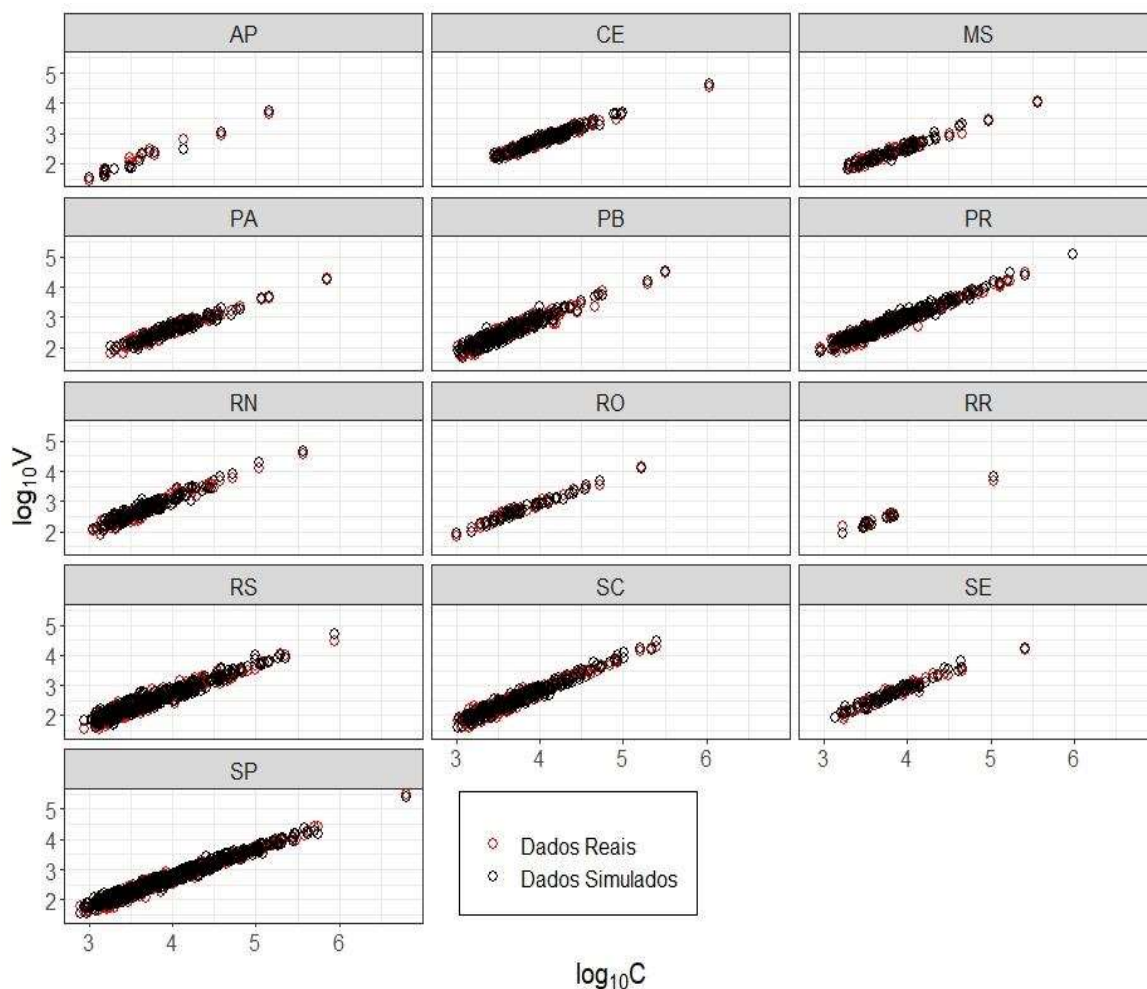
Fonte: Autoria própria (2021).

Foram construídas as densidades normalizadas dos 234 ajustes obtidos nessa análise com resultados, para todos os 7 anos de eleições, muito similares aos mostrados na Figura 40 para os dados de 2002. Onde podemos observar que as flutuações se apresentam consideravelmente próximas a distribuição normal, dessa forma, com o intuito de obter um modelo robusto, generalizado e completo para o número de votos brancos e nulos, inserimos um ruído gaussiano ζ na relação (3.2.2.7)

$$V \approx 0,1 \left(\frac{C}{C^*} \right)^{\alpha_{bn,k}} \zeta(C). \quad (3.2.2.9)$$

A simulação do modelo foi realizada para todas as eleições estudadas aqui, no entanto a Figura 41 mostra a simulação via modelo (3.2.2.9) em comparação com os dados empíricos para as eleições de 2002, como vem sendo apresentado desde o primeiro procedimento.

Figura 41: Comparação entre os dados reais e os simulados pelo modelo lei de potência. O painel acima mostra as relações $\log_{10}C$ versus $\log_{10}V$ em que C representa o comparecimento e V os brancos e nulos, ambos por cidade. Os círculos em vermelho representam os dados empíricos das eleições brasileiras de governadores de 2002 para cada estado onde ocorreu segundo turno, os círculos em preto correspondem aos dados simulados via modelo lei de potência (3.2.2.9).



Fonte: Autoria própria (2021).

3.2.3 Relação entre os expoentes

Foi observado que em todas as eleições $\alpha_p + \alpha_s + \alpha_{bn} \approx 3$. De fato, considerando os modelos lei de potências para as três partições de votos obtidas nessa análise, a conservação de probabilidade pode ser escrita na forma

$$1,35 \left(\frac{C}{C^*}\right)^{\alpha_{p,k}-1} + 1,35 \left(\frac{C}{C^*}\right)^{\alpha_{s,k}-1} + 0,3 \left(\frac{C}{C^*}\right)^{\alpha_{bn,k}-1} \approx 3. \quad (3.2.3.1)$$

Sendo cada expoente próximos de 1, cada termo do lado esquerdo pode ser expandido em série de Taylor na vizinhança de 1 e a equação aproximada (3.2.3.1) assume a forma

$$\begin{aligned}
 & 1 + 1,35 (\alpha_{p,k} - 1) \log_{10} \frac{C}{C^*} + 1,35 \frac{1}{2} (\alpha_{p,k} - 1)^2 \left(\log_{10} \frac{C}{C^*} \right)^2 + \dots \\
 & 1 + 1,35 (\alpha_{s,k} - 1) \log_{10} \frac{C}{C^*} + 1,35 \frac{1}{2} (\alpha_{s,k} - 1)^2 \left(\log_{10} \frac{C}{C^*} \right)^2 + \dots \\
 & 1 + 0,30 (\alpha_{bn,k} - 1) \log_{10} \frac{C}{C^*} + 0,30 \frac{1}{2} (\alpha_{bn,k} - 1)^2 \left(\log_{10} \frac{C}{C^*} \right)^2 + \dots \approx 3. \quad (3.2.3.2)
 \end{aligned}$$

Tomando na equação aproximada (3.2.3.2) a média sobre o comparecimento por cidade

$$\begin{aligned}
 & 1 + 1,35 (\alpha_{p,k} - 1) \langle \log_{10} \frac{C}{C^*} \rangle + 1,35 \frac{1}{2} (\alpha_{p,k} - 1)^2 \langle \left(\log_{10} \frac{C}{C^*} \right)^2 \rangle + \dots \\
 & 1 + 1,35 (\alpha_{s,k} - 1) \langle \log_{10} \frac{C}{C^*} \rangle + 1,35 \frac{1}{2} (\alpha_{s,k} - 1)^2 \langle \left(\log_{10} \frac{C}{C^*} \right)^2 \rangle + \dots \\
 & 1 + 0,30 (\alpha_{bn,k} - 1) \langle \log_{10} \frac{C}{C^*} \rangle + 0,30 \frac{1}{2} (\alpha_{bn,k} - 1)^2 \langle \left(\log_{10} \frac{C}{C^*} \right)^2 \rangle + \dots \approx 3. \quad (3.2.3.3)
 \end{aligned}$$

Para uma aproximação em primeira ordem e levando em conta que $(\alpha_{p,k} - 1)^2$ é consideravelmente pequeno, pode se usar as aproximações

$$(\alpha_{p,k} - 1)^2 \approx (\alpha_{s,k} - 1)^2 \approx (\alpha_{bn,k} - 1)^2. \quad (3.2.3.4)$$

Substituindo as aproximações (3.2.3.4) em (3.2.3.3) a equação aproximada pode ser simplificada como segue

$$\begin{aligned}
 & 1,35 (\alpha_{p,k} - 1) + 1,35 (\alpha_{s,k} - 1) + 0,30 (\alpha_{bn,k} - 1) \\
 & + \frac{3}{2} (\alpha_{p,k} - 1)^2 \langle \left(\log_{10} \frac{C}{C^*} \right)^2 \rangle / \langle \log_{10} \frac{C}{C^*} \rangle \approx 0. \quad (3.2.3.5)
 \end{aligned}$$

E nesse caso a aproximação em primeira ordem resulta na relação aproximada para os expoentes

$$1,35\alpha_{p,k} + 1,35\alpha_{s,k} + 0,30\alpha_{bn,k} = 3. \quad (3.2.3.6)$$

3.3 Resultados da análise

Foram investigadas 78 eleições brasileiras para governadores dos estados em segundo turno, como disponibilizado na Tabela 2, sendo que nesse cenário, os votos necessariamente são particionados entre: votos para os candidatos primeiro colocado, segundo colocado e brancos e nulos. Como primeiro procedimento foram construídos os gráficos em log-log para as relações do comparecimento versus os votos em cada partição por cidade, o qual apresentou leis de escala para as 3 partições de votos nas 78 eleições, ou seja, foram obtidas 234 leis de escala.

Expoentes e interceptos obtidos para as três partições foram dispostos relacionados a cada estado onde ocorreu a eleição, Figuras 37 e 38, possibilitando a comparação entre eles, sendo que foi possível observar que: em 44% das relações dos votos do primeiro colocado, em 56% dos votos do segundo colocado e em 70% dos votos brancos tem-se escalas super lineares. Verificou-se aqui que quando a escala do primeiro colocado é super linear a escala do segundo é sublinear e vice-versa, assim como foi verificado em Souza (2019), estudo análogo que desconsiderou os votos brancos e nulos.

A Figura 39 mostra a relação entre os parâmetros, interceptos e inclinações (expoentes), dos ajustes lineares obtidos da relação $\log_{10}C$ versus $\log_{10}V$, C comparecimento e V votos do primeiro colocado, do segundo colocado ou brancos e nulos. É visível que a relação entre os parâmetros do primeiro e do segundo colocados apresentam um ajuste consistente e único, como pode ser confirmado do pelo coeficiente de determinação $R^2 = 0,92$, enquanto, que para os parâmetros de brancos e nulos foi obtido um ajuste, separadamente, visivelmente não poderiam ser ajustados com os parâmetros do primeiro e segundo colocados, no entanto, mesmo assim o resultado não é satisfatório, visto que o coeficiente de determinação é $R^2 = 0,61$.

Os ajustes lineares para os dados em log-log, mostrados na Figura 35, fornecem um modelo lei de escala universal para as três partições de votos e para todas as 78 eleições estudadas aqui, totalizando 234 leis de potências, veja relação (3.2.2.3). A robustez desse resultado pode ser confirmada na Tabela 3 que mostra o coeficiente de determinação para cada um dos ajustes, onde pode ser observado

que $R^2 > 0,77$ para todos, sendo $R^2 > 0,90$ para 90% deles. Os poucos ajustes, com $0,77 < R^2 < 0,90$ ocorrem para os dados de estados como Amazonas – AM, Roraima – RR, que possuem um número pequeno de cidades, ou seja, poucos dados a serem ajustados. Substituindo o ajuste dos parâmetros para o primeiro e o segundo colocado, dado em (3.2.2.3), no modelo universal lei de escala (3.2.2.2), obteve-se um modelo dependendo de um único parâmetro (3.2.2.4) relacionando comparecimento e votos para os dois candidatos, por cidade. Observando que para um eleitorado fixo, $C^* = 10^{4,5}$, e juntando a equação de conservação de probabilidade para as três partições de votos, obrigatória, dada por (3.2.2.5), foi possível de terminar um parâmetro, a constante multiplicativa, do modelo lei de escala para brancos e nulos. Obtendo assim o principal resultado desse estudo, o modelo lei de escala universal dependendo de um único parâmetro, (3.2.2.7), relacionando comparecimento a votos brancos e nulos, por cidade.

Para completar o modelo lei de escala para votos brancos e nulos, foram analisadas as flutuações em torno do modelo linear (3.2.2.1), mostradas na Figura 40 para os dados de 2002, respectivas aos ajustes da Figura 3.2.1.1, e calculadas pela equação (3.2.2.7). É mostrada juntamente com a curva para cada estado a densidade de probabilidade normal padrão, o que possibilita observar uma proximidade considerável entre elas. Devido a esse resultado, foi inserido no modelo (3.2.2.7) um ruído gaussiano, obtendo o modelo (3.2.2.9), os dados simulados, juntamente com os dados empíricos, para a eleições de 2002, podem ser vistos na Figura 41.

Finalizando os procedimentos das análises, foi novamente utilizada a conservação de probabilidade entre as três partições de votos, substituídas as respectivas leis de escala, expandido os três termos em série de Taylor para os expoentes em torno de 1, dado que observamos durante os procedimentos, que os expoente para as três partições são sempre próximos a 1, e tomado a aproximação em primeira ordem, foi obtida uma equação proporcional aproximada, (3.2.3.6), relacionando os expoentes das três partições de votos.

3.4 Conclusões finais

Esse estudo obteve leis de escala relacionando comparecimento e votos para as três partições de votos em segundo turnos de eleições para governadores dos estados brasileiros, mostrando que existe um comportamento padrão para a dinâmica eleitoral nesse contexto.

A as proporções 44%, 56% e 70% de escalas super lineares, para as escalas do primeiro colocado, segundo colocado e brancos e nulos, respectivamente, permitem fazer observações importantes. Note o segundo colocado venceu a eleição nos maiores eleitorados, em 12% a mais do que o primeiro colocado, resultado que indica que os candidatos vencedores, na maior parte das eleições, 56%, dedicaram mais esforços de campanha aos eleitorados menores. A grande proporção de escalas super lineares para brancos e nulos indicam que quanto maior o eleitorado o eleitor tende a ser mais rigoroso na sua escolha, supondo que brancos e nulos representam a rejeição de eleitores insatisfeitos com ambas as opções de candidatos.

O cenário escolhido, segundo turnos de eleições para governadores, que fornece uma relação de conservação entre o montante de votos para as três partes, possibilitou a obtenção de um modelo robusto dependendo de um único parâmetro, resultado central desse estudo, para a relação entre comparecimento e votos brancos e nulos. Resultado que possibilita analisar o “descontentamento” do eleitorado em função do número total de votos. Com isto, a proposta desse trabalho, que foi compreender a dinâmica dos votos brancos e nulos, supondo que esses representam eleitores insatisfeitos com a configuração de candidatos disponível no segundo turno, foi alcançada com êxito.

O estudo apresentando nesse trabalho permitiu ampliar a compreensão de uma questão pouco abordada nas investigações do ponto de vista dos sistemas complexos, a dinâmica das opiniões de eleitores insatisfeitos com as opções de candidatos. Em geral a maiorias dos estudos focam na opinião dos eleitores apenas considerando votos para os candidatos, até mesmo o sistema eleitoral contabiliza votos brancos e nulos apenas como estatísticas. No entanto, os resultados obtidos podem contribuir na interpretação do resultado de uma eleição e ajudar no planejamento de uma campanha e até mesmo processo de candidatura.

REFERÊNCIAS

- ARARIPE, L. E.; COSTA FILHO, R. N. Role of parties in the vote distribution of proportional elections. **Physica A**, v. 388, p. 4167- 4170, jun. 2009.
- ARARIPE, L.E.; COSTA FILHO, R. N.; HERRMANN, H. J.; ANDRADE, J. S. Plurality voting: the statistical laws of democracy in Brazil. **International Journal Modern Physics C**, v. 17, n. 12, p. 1809-1813, 2006.
- BARABÁSI, A-L.; ALBERT, R. Emergence of Scaling in Random Networks. **Science**, v. 286, p. 509 - 512, oct. 1999.
- BETTENCOURT, L. M. A.; LOBO, J.; HELBING, D.; KUHNERT, C.; WEST, G.B. Growth, innovation, scaling, and the pace of life in cities. **Proceedings of the National Academy of Sciences of the United States of America**, v. 104, n. 17, p. 7301 – 7306, apr. 2007.
- BOKÁNYI, E.; SZÁLLÁSI, Z.; VATTAY, G. Universal scaling laws in metro area election results. **PLoS One**, v. 13, p. 1 – 11, feb. 2018.
- BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil de 1988**. [Brasília]: Presidência de República, 2016. Disponível em: http://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm. Acesso em: 9 set. 2020.
- CARDOSO, M.; MENDES, R.S.; SOUZA, J.T.Z.; RIBEIRO, H.V. Gender difference in candidature processes for Brazilian elections. **Physica A**, v. 537, p. 1-7, sep. 2019.
- CHATTERJEE, A., MITROVIC, M; FORTUNATO, S. Universality in voting behavior: an empirical analysis, **Scientific Reports**, v. 3, n. 1049, p. 1 – 8, jan. 2013.
- COSTA FILHO, R.N.; ALMEIDA, M. P.; ANDRADE, JR. J. S.; MOREIRA, J.E. Scaling behavior in a proportional voting process. **Physical Review E**, v. 60, n. 1, p. 1067 – 1068, jul. 1999.
- COSTA FILHO, R.N.; ALMEIDA, M. P.; ANDRADE, JR. J. S; MOREIRA, J.E. Brazilian elections: voting for a scaling democracy. **Physica A**, v. 322, p. 698-700, nov. 2002.
- TSALLIS, C. Possible generalization of Boltzmann-Gibbs statistics. **Journal of Statistical Physics**, v. 52, p. 479 – 487, nov. 1988.
- DINIZ, M. H. **Dicionário jurídico**. São Paulo: Saraiva, v. 4, p. 760, 1998.
- FARHAT, S. **Dicionário parlamentar e político**. São Paulo: Melhoramentos; Fundação Peirópolis, 1996.

FORTUNATO, S.; CASTELLANO, C. Scaling and Universality in Proportional Elections. **Physical Review Letters**, v. 99, n. 138701, sep. 2007.

HONG, B. H.; LEE, K. E.; LEE, J. W. Power Law in Firms Bankruptcy. **Physics Letter A**, v. 361, p. 1 – 7, jan. 2007.

LYRA, M. L.; COSTA, U. M. S.; COSTA FILHO, R. N.; JUNIOR ANDRADE, J.S.. Generalized Zipf's law in proportional voting processes. **Europhysics Letters**, v. 62, p.131-137, 2003.

MANTOVANI, M. C.; RIBEIRO, H. V.; MORO, M. V.; PICOLI JUNIOR, S.; MENDES, R.S. Scaling laws and universality in the choice of election candidates. **A Letters Journal Exploring The Frontiers of Physics**, v. 96, p. 1 – 5, sep. 2011.

MANTOVANI, M.C.; RIBEIRO, H.V.; LENZI, E.K.; MENDES, R.S. Engagement in the electoral processes: scaling laws and the role of the political position. **Physical Review E**, v. 88, p. 1, aug. 2013.

MELO, H. P. M.; MOREIRA, A. A.; MAKSE, H. A.; ANDRADE, J. S. Statistical signs of social influence on suicides. **Scientific Reports**, v. 4, n. 6239, p. 1 – 5, sep. 2014.

MELO, H. P. M.; REIS, S. D. S.; MOREIRA, A. A.; MAKSE, H. A.; ANDRADE JUNIOR, J. S. The price of a vote: diseconomy in proportional elections. **PLoS One**, v. 13, p. 1 – 13, aug. 2018.

POOSALA, V. Zipf's Law. **ResearchGate**, p. 1, 1999.

SENNA, A. Mensagem do nº1 Ayrton Senna - seja quem você for. Youtube, 1 set. 2009. Disponível em: <https://www.youtube.com/watch?v=RqIPoz4Cf8U>. Acesso em: 20 jan. 2021.

SOUZA, J. T. G. **Leis de potência e eleições**. 2019. Dissertação (Mestrado em Inovações Tecnológicas) – Programa de Pós-Graduação em Inovações Tecnológicas, Universidade Tecnológica Federal do Paraná, Campo Mourão, 2019.

APÊNDICE A - Lei de potência ou escala²

² Adaptado de Souza (2019)

LEI DE POTÊNCIA OU ESCALA

São conhecidas como lei de potência ou comportamento de escala manifestações fractais onde um grande conjunto de dados e fenômenos interagem entre si e pode ser descrito pelo modelo matemático:

$$Y = Y_0 N^\beta \quad (\text{A.1})$$

Y_0 = Constante de Proporcionalidade e β = Constante

Tomando o logaritmo da Equação A.1:

$$\log(Y) = \log(Y_0) + \beta \log(N) \quad (\text{A.2})$$

APÊNDICE B - Lei de Zipf

LEI DE ZIPF

A lei de Zipf pode ser encontrada em muitos fenômenos sociais, como define Hong et al. (2007), pode descrever o tamanho de cidades, tamanho de terremotos, crateras lunares, erupções solares, a intensidade das guerras, a frequência do uso de palavras em um determinado idioma humano, as vendas de livros, o número de espécies em taxa biológica, dentre outras coisas. A lei de Zipf apresenta algumas versões, aqui apresenta-se a mais simples e generalizada Poosala (1999).

$$x_1 > x_2 > x_3 > \dots > x_n \quad (\text{B.1})$$

Sendo:

n = posto de x_n

x_n = Tamanho do valor de dados no conjunto ordenado

A lei de Zipf é uma relação entre a classificação de um valor de dados e seu valor real

valor que foi empiricamente observado como o seguinte Poosala (1999)

$$nx_n = \text{Constante} \quad (\text{B.2})$$

Seja x o tamanho de um elemento e $f(x)$ a frequência relativa da ocorrência do elemento, onde $\int_0^{\infty} f(x)dx = 1$. Se n é o número de objetos no conjunto de dados é $N(x)$ o número de objetos com tamanho maior que x então,

$$N(x) = \int_x^{\infty} nf(u)du \quad (\text{B.3})$$

e a classificação do elemento de tamanho x . Sob a lei de Zipf na Equação B.1, $xN(x) = \text{constante}$. Consequentemente

$$f(x) = -\frac{N'(x)}{n} = \frac{c'}{x^2} \quad (\text{B.4})$$

Onde $c' = \frac{c}{n}$. A equação B.4 é a relação tamanho-frequência correspondente a equação B.2.

Os fenômenos observados por Zipf e justificados pelo raciocínio estatístico levam a uma família de distribuições,

$$n^\alpha x_n = c \quad (\text{B.5})$$

Onde:

$\alpha > 0$ e c levando a seguinte relação tamanho-frequência:

$$f(n) = An^{-(1+\alpha)}, n = 1, 2, \dots, \quad (\text{B.6})$$

Onde $\alpha > 0$, e

$$A = \xi(1 + \alpha) = \sum_{n=1}^{\infty} n^{-(1+\alpha)} \quad (\text{B.7})$$

É a função zeta. A equação B.7 define a distribuição discreta de Pareto, que inclui a equação B.4 como um caso especial quando $\alpha = 1$.

APÊNDICE C - Distribuição Gauciana e Distribuição Lognormal³

³ Adaptado de Souza (2019)

DISTRIBUIÇÃO GAUCIANA E DISTRIBUIÇÃO LOGNORMAL

(BITTENCOURT; VIALLI, 2006). definem a distribuição normal ou curva de Gauss como um modelo que descreve o comportamento de vários fenômenos aleatórios. Dessa distribuição se destaca o Teorema do Limite Central, que diz:

Seja x_1, \dots, x_n , n variáveis aleatórias e independentes e identicamente distribuídas com média μ e desvio padrão σ . A distribuição da soma das n variáveis tende a apresentar um comportamento normal como:

$$x_1 + x_2 + \dots + x_n = \sum x_i \sim normal \quad (C.1)$$

Ou seja, as medias de amostras retiradas de distribuição qualquer tendem a convergir para uma distribuição normal quando o número de observações (tamanho da amostra) aumenta (BITTENCOURT; VIALLI, 2006). A função densidade dessa distribuição é denotada como:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu/\sigma)^2} \quad (C.2)$$

Onde $\mu \in \mathfrak{R}$ é a media $\sigma > 0$ o desvio padrão $x \in \mathfrak{R}$ (BITTENCOURT; VIALLI, 2006).

Uma Variável X tem uma distribuição lognormal quando o seu logaritmo $Y = \log(x)$ tem distribuição normal. A função de densidade desta distribuição é denotada da seguinte maneira:

$$f(x, \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{\ln(x)-\mu^2}{2\sigma^2}\right] \quad (C.3)$$

Com $x > 0$ e 0 caso contrário.

APÊNDICE D – Coeficiente de determinação

COEFICIENTE DE DETERMINAÇÃO

O coeficiente de determinação também chamado de R^2 pode ser interpretado como uma medida de ajuste de um modelo estatístico, expressando a quantidade da variância dos dados que é explicada pelo modelo linear, ou seja, este coeficiente explica o quanto o modelo foi capaz de explicar os dados coletados.

$$SQ_{tot} = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (D.1)$$

Onde n é o número de observações, y_i é o valor observado e \bar{y} é a média das observações.

APÊNDICE E – Rede de barbási-albert⁴

⁴ Adaptado de Souza (2019)

REDE DE BARABÁSI-ALBERT

Este é um modelo de rede progredindo no tempo, inspirado na formação da World Wide Web (WWW) e baseado em dois componentes: crescimento e ligação preferencial (BARABÁSI e ALBERT, 1999). O princípio básico é que na WWW, nós (páginas) com uma alta conectividade alcançam novos nós em uma taxa mais alta do que nós com uma baixa conectividade. O modelo inicia-se com m_0 nós isolados em cada intervalo de tempo $t = 1, 2, 3, \dots, N - m_0$ um novo nó j é adicionado a rede com $m \leq m_0$ ligações. A probabilidade que o nó j seja ligado a um nó i que já existe é proporcional a conectividade atual de i :

$$\frac{k_i}{\sum_l k_l} \quad (\text{E.1})$$

Cada novo nó tem m ligações; no tempo t , a rede deve ter $N = m_0 + t$ nós e $K = mt$ ligações, correspondendo a um grau médio $\langle k \rangle = 2m$ para tempos grandes. No limite de $t \rightarrow \infty$, o modelo produz uma distribuição de grau $P(k) \propto k^{-\gamma}$, com $\gamma = 3$ (MANTOVANI, 2013).

A distribuição de grau segue uma lei de potência $k^{-\gamma}$ com expoente:

$$\gamma = 1 + \frac{2m(1-\rho)+1-\lambda-\rho}{m} \quad (\text{E.2})$$

APÊNDICE F – Série de taylor

SÉRIE DE TAYLOR

Serie de Taylor é uma serie restrita a classe de potência, podendo ser utilizada para aproximar funções ao redor um ponto, bem como, encontrar seus pontos máximos e mínimos.

$$f(x) = \sum_{n=0}^{\infty} a_n (x - a)^n \quad (\text{F.1})$$

Sendo:

$$a_n = \frac{f^n(a)}{n!} \quad (\text{F.2})$$