

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**

**JUSLEY TALITA GRIMES DE SOUZA**

**LEIS DE POTÊNCIA E ELEIÇÕES**

**DISSERTAÇÃO**

**CAMPO MOURÃO**

**2019**

**JUSLEY TALITA GRIMES DE SOUZA**

**LEIS DE POTÊNCIA E ELEIÇÕES**

Dissertação apresentada ao Programa de Pós Graduação em Inovações Tecnológicas, da Universidade Tecnológica Federal do Paraná, como requisito parcial para a obtenção do grau de Mestre.

Orientador: Prof<sup>a</sup>. Dr<sup>a</sup>. Magda Cardoso Mantovani.

Coorientador: .

**CAMPO MOURÃO**  
**2019**

### Dados Internacionais de Catalogação na Publicação

S726l

Souza, Jusley Talita Grimes de

Leis de potência e eleições / Jusley Talita Grimes de Souza –  
2019.

100f. : il. ; 30 cm.

Texto em português com resumo em inglês  
Orientadora: Magda Cardoso Mantovani

Dissertação (Mestrado) – Universidade Tecnológica Federal do  
Paraná. Programa de Pós-Graduação em Inovações Tecnológicas  
Campo Mourão, 2019.

Inclui bibliografias.

1. Eleições. 2. Estatística. 3. Inovações Tecnológicas –  
Dissertações. I. Mantovani, Magda cardoso, orient. II. Universidade  
Tecnológica Federal do Paraná. Programa de Pós-Graduação em  
Inovações Tecnológicas. III. Título.

CDD: 607

Biblioteca Câmpus Medianeira  
Fernanda Bem 9/1735

## **FOLHA DE APROVAÇÃO**

**JUSLEY TALITA GRIMES DE SOUZA**

**LEIS DE POTÊNCIA E ELEIÇÕES**

Dissertação apresentada ao Programa de Pós Graduação em Inovações Tecnológicas, da Universidade Tecnológica Federal do Paraná, como requisito parcial para a obtenção do grau de Mestre.

Data de aprovação: 04 de março de 2019.

---

Prof<sup>ª</sup>. Dr<sup>ª</sup>. Magda Cardoso Mantovani  
Universidade Tecnológica Federal do Paraná

---

Prof. Dr. Haroldo Valentin Ribeiro  
Universidade Estadual de Maringá

---

Prof. Dr. Wyrllen Everson de Souza  
Universidade Tecnológica Federal do Paraná

---

Prof. Dr. Leandro de Santana Costa  
Universidade Tecnológica Federal do Paraná

A folha de aprovação assinada encontra-se na coordenação do curso ou programa.

## **AGRADECIMENTOS**

A Deus por permitir a conclusão desse trabalho me dando saúde, força e me capacitando.

A minha orientadora Magda Cardoso Mantovani, pelo acolhimento, apoio e orientação com muita paciência e dedicação.

A minha família, pelo amor, incentivo e apoio incondicional.

Aos meus amigos, que sempre me incentivaram, ouviram e aconselheram em muitos momentos durante essa trajetória. Em especial, Adriele, Camila, Danielly, Miriam, Suellen e Fernando.

A Jaiza e Cibele, amigas que fiz no decorrer do mestrado e que me ajudaram muito.

E a todos que diretamente ou indiretamente fizeram parte dessa conquista, o meu muito obrigada.

*Agrada-te do Senhor, e Ele satisfará os desejos do teu coração.*  
*Salmos 37:4*

## RESUMO

Utilizando conceitos e ferramentas de Estatística este trabalho apresenta alguns exemplos de sistemas sociais, nos quais as suas dinâmicas são modeladas por leis de potência ou leis de escala. Inicialmente, foi dado um breve panorama da emergência de leis de escala em fenômenos urbanos de modo geral: características universais nas dinâmicas de crescimento, atividades religiosas, escala urbana e distribuição lei de potência e alometrias para indicadores urbanos. Especificamente, fenômenos eleitorais modelados por relações de escala foram o foco deste trabalho. Neste sentido, foram apresentados resultados em que tais relações emergem de fenômenos eleitorais. Em seguida, são apresentadas as investigações desta dissertação. Essa pesquisa, teve como objetivo investigar resultados de eleições brasileiras para governos dos estados federativos num cenário de apenas dois partidos políticos, tomando, assim, dados de segundo turnos no período de 1994-2018. Obteve-se um modelo analítico dependendo de um único parâmetro, o que possibilitou constatar que na maioria das eleições o número de votos do primeiro colocado tende a crescer com o tamanho da cidade numa proporção menor que o segundo colocado.

**Palavras-chaves:** Sistemas Complexos. Eleições. Lei de Potência.

## **ABSTRACT**

By using concepts and tools from statistics, this paper presents some examples of social systems, in which their dynamics are modeled by power laws or scaling laws. Initially, it was given a brief overview about the scaling laws emerging from urban phenomena in general: universal characteristics in growth dynamics, religion activities, urban scales, power law distribution, and allometric urban indicators. Specifically, electoral phenomena modeled by scaling relations, which was the focus of this work. In this direction, it was presented results in which such relations emerge from electoral phenomena. Next, the investigations of this dissertation are presented. This research aimed to investigate Brazilian election results for federal state governments in a scenario with only two political parties, disputing thus taking data from the second round of elections in the period 1994-2018. This dataset allowed to verify that in the majority of the elections the number of votes for the first placed candidate tends to grow with the city size in a smaller proportion than those of the second placed candidate.

**Key-words:** Complex Systems. Elections. Power Law.

## SUMÁRIO

|                   |   |           |
|-------------------|---|-----------|
| <b>1</b>          | <b>INTRODUÇÃO</b>   | <b>9</b>  |
| <b>2</b>          | <b>LEIS DE POTÊNCIA E FENÔMENOS SOCIAIS</b>   | <b>11</b> |
| 2.1               | Leis de Potência modelam crescimento de atividades religiosas   | 11        |
| 2.2               | Escalas urbanas conectadas com a Lei de Zipf  | 17        |
| 2.3               | Uma relação entre a escala urbana e função de produção  | 25        |
| 2.4               | Fatores que dão origem as relações de escala dos fatores urbanos relacionados ao tamanho da população | 31        |
| 2.5               | Conexão entre distribuição Lei de Potência e Alometrias para indicadores urbanos                      | 35        |
| <b>3</b>          | <b>LEIS DE POTÊNCIA EM ELEIÇÕES</b>   | <b>44</b> |
| 3.1               | Distribuição de Votos   | 44        |
| 3.2               | Simulação de Distribuição de Votos em Rede  | 48        |
| 3.3               | Relevância dos Partidos   | 54        |
| 3.4               | Processo de Candidatura   | 56        |
| 3.5               | Lei de Potência Universal nas Eleições Presidenciais dos Estados Unidos                               | 61        |
| 3.6               | A falta de economia em eleições proporcionais   | 64        |
| <b>4</b>          | <b>LEI DE ESCALA EM SEGUNDO TURNO DE ELEIÇÕES BRASILEIRAS</b>   | <b>72</b> |
| 4.1               | Apresentação dos Dados  | 73        |
| 4.2               | Análise dos Dados   | 76        |
| 4.3               | Resultados e Conclusões   | 86        |
|                   | <b>REFERÊNCIAS</b>  | <b>89</b> |
|                   | <b>APÊNDICES</b>  | <b>93</b> |
| <b>APÊNDICE A</b> | <b>LEI DE POTÊNCIA OU ESCALA</b>  | <b>94</b> |
| <b>APÊNDICE B</b> | <b>LEI DE ZIPF</b>  | <b>95</b> |

|                   |  |            |
|-------------------|--|------------|
| <b>APÊNDICE C</b> | <b>DISTRIBUIÇÃO GAUSSIANA E DISTRIBUIÇÃO LOGNOR-<br/>MAL</b> | <b>97</b>  |
| <b>APÊNDICE D</b> | <b>MODELO DE SZNAJD</b>                                      | <b>98</b>  |
| <b>APÊNDICE E</b> | <b>REDE DE BARABÁSI-ALBERT</b>                               | <b>99</b>  |
| <b>APÊNDICE F</b> | <b>REDE DE ERDÖS-RÉNYI</b>                                   | <b>100</b> |

## 1 INTRODUÇÃO

O número de indivíduos, assim como o números de cidades, estão em geral crescendo ininterruptamente e com isso os desafios sociais também aumentam. Além disso, as interações sociais são reforçadas devido à alta mobilidade das pessoas e grande facilidade de comunicação. Diante desses fatores, é também uma necessidade crescente compreender os variados aspectos da sociedade. Nos últimos anos, muitas pesquisas têm sido realizadas utilizando conceitos e ferramentas de Estatística na tentativa de entender fenômenos coletivos emergentes das interações entre pessoas como unidades elementares em estruturas sociais, isto é, com o objetivo de aumentar a compreensão de sistemas sociais. Nesse contexto, do ponto de vista dos sistemas complexos, muitos sistemas sociais têm suas dinâmicas modeladas por leis de potência, ou leis de escala. Segundo Macau (2002), um sistema é considerado “sistema complexo” quando apresenta um comportamento entrelaçado, no qual é difícil de ser modelado analisando somente suas partes. Como os sistemas complexos aparecem em diversas áreas, alguns exemplos serão apresentados nesse trabalho.

Nesse sentido, no primeiro capítulo apresentou-se alguns estudos sobre pesquisas de fenômenos urbanos modelados por distribuições do tipo lei de potência e alometrias. Começando pelo sistema social da religião, Picoli e Mendes (2008) descobriram que a distribuição das taxas logarítmicas de crescimento anuais exibe a mesma forma funcional para escalas de tamanhos distintas e o desvio padrão das taxas de crescimento escala como uma lei de potência. Gomez-Lievano et al. (2012) construíram uma ferramenta estatística auto-consistente que caracteriza as distribuições de probabilidades conjuntas de indicadores urbanos e tamanho de população da cidade através de um sistema urbano. Lobo et al. (2013) mostraram empiricamente que existe uma dependência sistemática da produtividade urbana no tamanho da população da cidade, resultante do descompasso entre a dependência do tamanho do salário e trabalho, de modo que a produtividade das cidades americanas aumenta cerca de 11% cada vez que dobra a população. Bettencourt (2013) verificou que medidas de eficiência urbana que capturam o equilíbrio entre os resultados socioeconômicos e os custos de infra-estrutura, mostraram-se independentes do tamanho da cidade. Alves et al. (2014) mostraram que um conjunto de 12 indicadores urbanos das cidades brasileiras, são assintoticamente distribuídos como leis de potência.

No segundo capítulo apresentou-se um histórico de resultados de leis de potência em um tipo de fenômeno urbano de grande interesse para a população: as eleições. Inicia-se com Costa Filho et al. (1999;2002), argumentando que o número de candidatos que recebem uma dada fração de votos, em eleições proporcionais

brasileiras, tem distribuição seguindo uma lei de potência com expoente  $-1$ , e que esse comportamento pode ser modelado como um processo multiplicativo; em seguida, os mesmos dados foram melhor ajustados por Lira et al. (2003) usando uma generalização da lei de Zipf. Com o objetivo de simular o comportamento de lei de potência encontrado por Costa Filho et al., Bernardes et al. (2002) construíram um modelo que utiliza a generalização da dinâmica de opinião de Snajd na rede de Barabási-Albert. Travieso e Costa (2006) construíram uma dinâmica de opinião e a aplicaram a cinco modelos de redes. Analisando as eleições proporcionais e, agora considerando a “força” dos partidos, Fortunato e Castellano (2007) apontam que a distribuição do número de votos recebidos por candidatos para três países é ajustada por uma lognormal; em contraste, Arirape e Costa Filho (2009) indicam que os resultados para eleições brasileiras são diferentes. Mantovani et al. (2013), analisando dados de eleições para prefeitos e vereadores, encontraram um comportamento médio lei de potência e um aparente crescimento do expoente com o número de posições. Bokányi et al. (2018) analisaram dados das eleições dos Estados Unidos e do Referendo da União Europeia e concluíram que as curvas de escala são todas universais e dependem de um único parâmetro. Melo et al. (2018) investigaram o crescente custo das campanhas eleitorais, mostrando que os principais candidatos gastam mais dinheiro por voto do que os candidatos menos bem sucedidos, uma superlinearidade que revela uma não economia de escala.

Diante desses resultados para fenômenos urbanos modelados por leis de escala, o capítulo 3 apresenta a pesquisa desenvolvida neste trabalho, que foi direcionada pelos resultados anteriores, em particular, o trabalho que investigou os resultados de eleições para presidentes dos Estados Unidos entre 1948-2016 e do referendo sobre a permanência do Reino Unido na União Europeia em 2016, identificando um comportamento de escala universal com um único parâmetro (BOKÁNYI, 2017). Nesse contexto, essa pesquisa realizou uma investigação similar para os dados do Brasil, optando por investigar eleições para governos dos estados federativos. Além disso, usou-se dados de segundo turnos com o objetivo de obter um modelo para um cenário com apenas dois partidos, obtendo uma lei de escala universal.

## 2 LEIS DE POTÊNCIA E FENÔMENOS SOCIAIS

Com o rápido crescimento das cidades há uma crescente preocupação em compreender os fenômenos urbanos. No entanto, persiste a dificuldade em abordá-los devido a interdependência entre eles. Porém, a existência de uma relação com o tamanho da população é evidente. Sendo  $Y = Y_0 N^\beta$ , onde  $N$  é o tamanho da população,  $Y$  o fenômeno urbano, e neste contexto e, do ponto de vista dos sistemas complexos, foram investigados fenômenos sociais como em religião, economia, propagação de notícias, popularidade, dinâmicas de opinião e política. Neste capítulo serão apresentados alguns exemplos de investigações de fenômenos urbanos, nos quais tais investigações focaram em características universais nas dinâmicas de crescimento, atividades religiosas, escala urbana e distribuições lei de potência e alometrias para indicadores urbanos.

### 2.1 Leis de Potência modelam crescimento de atividades religiosas

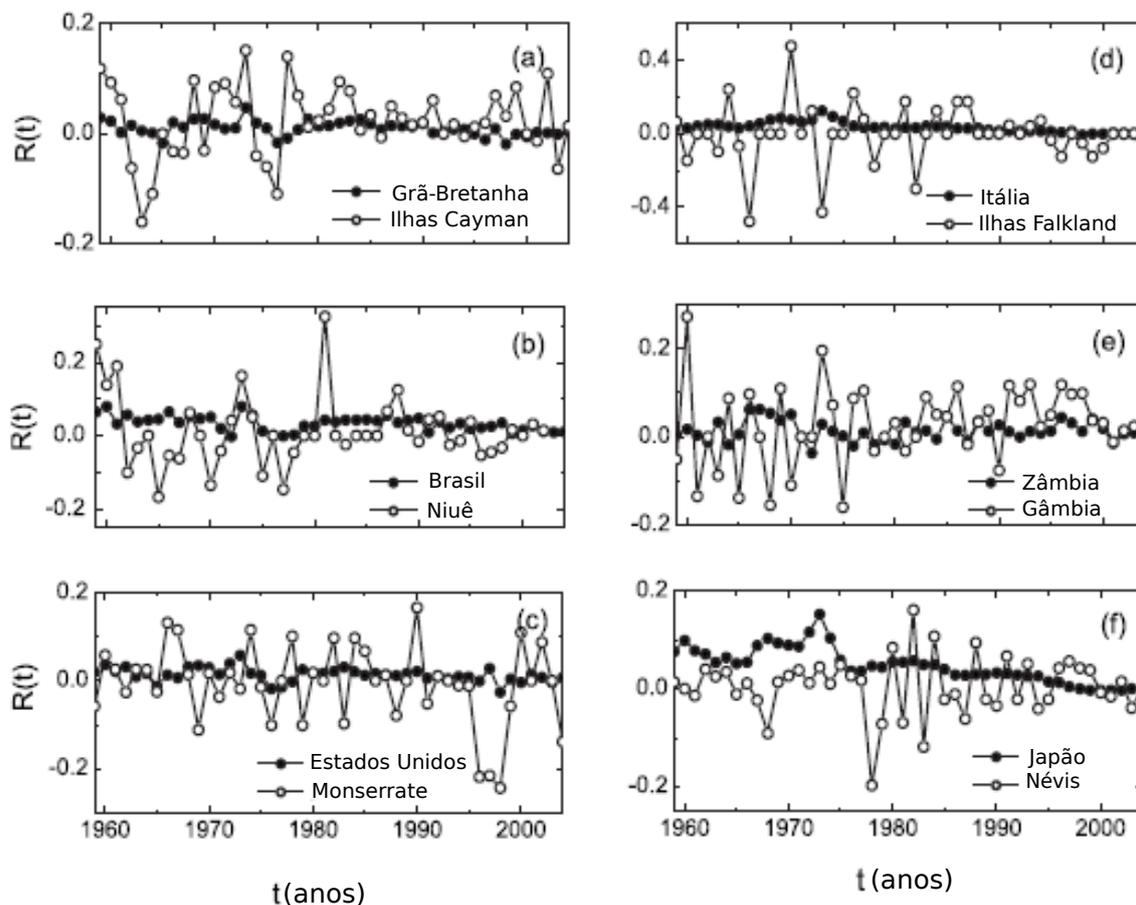
Grupos religiosos são sistemas sociais muito importantes. Nesse sentido, Picoli e Mendes (2008), investigaram a dinâmica de crescimento de um grupo religioso cristão, as Testemunhas de Jeová, em 140 países por um período de 47 anos, desde de 1959 a 2005. Esse grupo compreende mais de 6,9 milhões de membros praticantes em todo o mundo, organizados em cerca de 100.000 congregações em 236 países. Todos os membros editores são envolvidos em um trabalho de pregação pública. As medidas de suas atividades religiosas são registradas pelo país e publicadas nos relatórios anuais do serviço. Os dados analisados foram obtidos a partir desses relatórios anuais de serviços. As investigações focaram em quatro medidas distintas de atividade religiosa: número total de editores; média mensal de tempo integral de editores pioneiros; tempo total investido em horas no trabalho de pregação pública e média mensal de cursos realizados em casas.

Considerou-se a taxa anual de crescimento logarítmica definida como

$$R(t) = \log_{10} \left[ \frac{S(t+1)}{S(t)} \right], \quad (2.1)$$

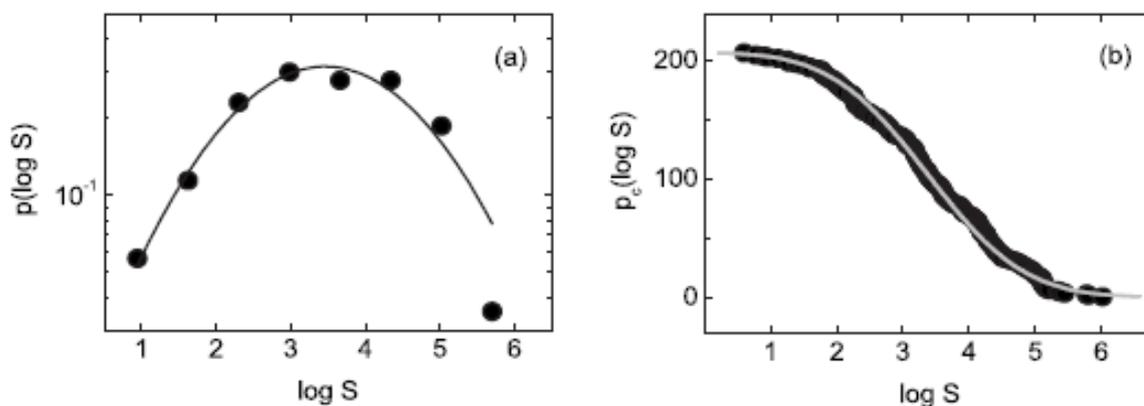
onde  $S(t)$  e  $S(t+1)$  são os números totais de editores em anos sucessivos  $t$  e  $t+1$ , em um dado país. É mostrada na Figura 2.1.1 a evolução temporal de  $R(t)$  para alguns países no período de 1959-2005. Nesse período as médias dos números de editores,  $\bar{S}$ , para esses países são: Grã-Bretanha ( $\bar{S} = 92,006$ ) e Ilhas Cayman ( $\bar{S} = 67$ ), Brasil ( $\bar{S} = 226,682$ ) e Niuê ( $\bar{S} = 20$ ), Estados Unidos ( $\bar{S} = 655,607$ ) e Monserrate ( $\bar{S} = 24$ ), Itália ( $\bar{S} = 112,660$ ) e Ilhas Falkland ( $\bar{S} = 6$ ), Zâmbia ( $\bar{S} = 66,554$ ) e Gâmbia ( $\bar{S} = 42$ ),

Japão ( $\bar{S} = 96,937$ ) e Névis ( $\bar{S} = 40$ ). Note que as flutuações em  $R$  são maiores nos países com pequeno  $S$ .



**Figura 2.1.1 – Taxas de crescimento.** Evolução temporal das taxas de crescimento, dada pela Equação 2.1, para alguns países no período de 1959-2005. Adaptada de Picoli e Mendes (2008).

São mostradas na Figura 2.1.2 (a) a densidade de probabilidade de  $\log_{10}S$ , para 207 países em 2005, em comparação com uma distribuição Gaussiana.  $S$  tem distribuição aproximadamente lognormal, que é reforçada pela distribuição cumulativa de  $\log_{10}S$  na Figura 2.1.2 (b). Resultados semelhantes são válidos para outros anos no período de 1959-2005.



**Figura 2.1.2 – Distribuição de editores.** (a) Densidade de probabilidade do logaritmo do número de editores para 207 países em 2005. A linha sólida é um ajuste Gaussiano dos dados. (b) Distribuição cumulativa do logaritmo do número de editores para os mesmos dados definidos em (a). A linha sólida representa a distribuição cumulativa de uma variável Gaussiana com média e desvio padrão idênticos aos dados empíricos. Adaptada de Picoli e Mendes (2008).

Em seguida, foram analisadas a distribuição das taxas de crescimento de editores usando dados de 140 países de 1949 a 2005. Por ser uma atividade religiosa relativamente recente o número de editores representa uma fração bem pequena da população dos países.

Com o intuito de investigar como a distribuição das taxas de crescimento depende do número de editores, considerou-se grupos de países escolhidos de acordo com a média dos editores. A Figura 2.1.3 (a) mostra a densidade de probabilidade condicional  $p(R|S)$  das taxas de crescimento bem descrita por uma distribuição de Laplace

$$p(R|S) = \frac{1}{\sqrt{2}\sigma(S)} \exp\left(-\frac{\sqrt{2}|R - \mu|}{\sigma(S)}\right), \quad (2.2)$$

onde  $\mu$  e  $\sigma$  são, respectivamente, a média e o desvio padrão de  $R$  calculados sobre um determinado grupo de países. Na Figura 2.1.3 (b) é mostrado a distribuição cumulativa do valor absoluto de  $R$  em comparação com funções exponenciais.

Foi ainda calculada, Figura 2.1.3 (c), a densidade de probabilidade condicional  $p(r|S)$  da taxa de crescimento normalizada

$$r(t) = \frac{R(t) - \mu}{\sigma}. \quad (2.3)$$

Colapsando em uma única curva,

$$p(r|S) = \frac{1}{\sqrt{2}} \exp(-\sqrt{2}|r|). \quad (2.4)$$

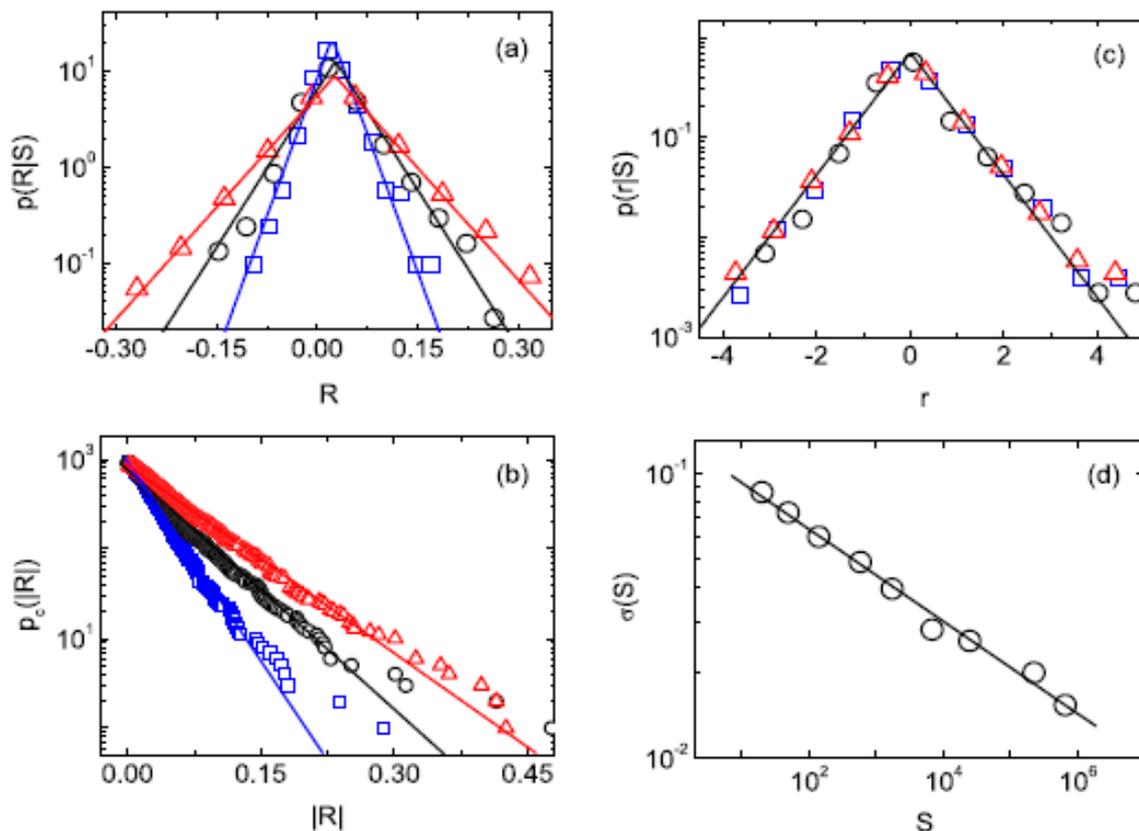
Nas Figuras 2.1.3 (a) e 2.1.3 (c) pode ser observado um tipo de invariância de escala, já que a distribuição das taxas de crescimento exibe a mesma forma funcional da distribuição de Laplace para escalas de tamanhos distintos.

Analisou-se ainda como a largura da distribuição das taxas de crescimento aumenta com o tamanho da escala. Calculou-se  $\sigma$  em função de  $S$ , encontrando-se uma relação aproximadamente lei de potência

$$\sigma(s) \propto S^{-\beta}, \quad (2.5)$$

com  $\beta \simeq 0,16$  como mostra a Figura 2.1.3 (d), indicando que as flutuações nas taxas de crescimento decaem como uma lei de potência. As flutuações de  $R$  são maiores para países com pequenos  $S$ , o que é consistente com a Figura 2.1.1.

Considerou-se ainda o grupo  $76,638 < \bar{S} < 655,607$  mas, para maior clareza, mostrou-se dados apenas para os três primeiros grupos.

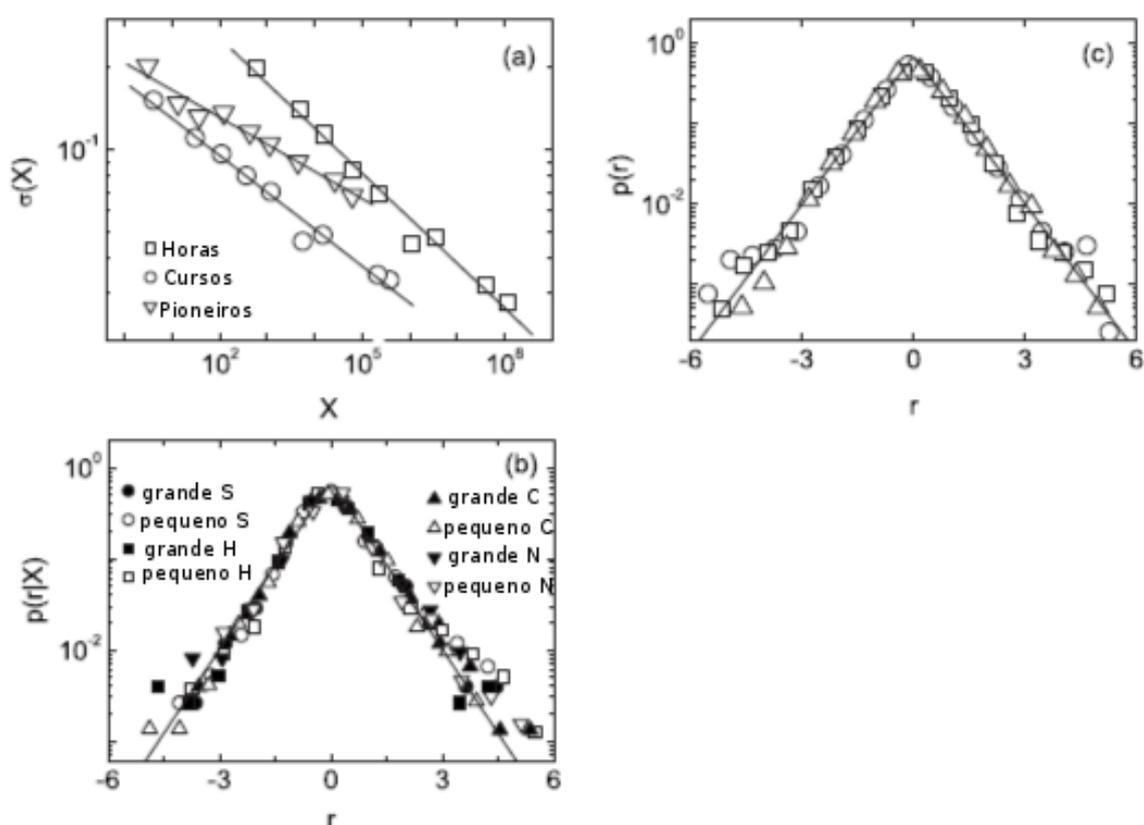


**Figura 2.1.3 – Dinâmica de crescimento dos editores.** Em (a), (b) e (c) os grupos de países selecionados pelo número médio de editores:  $6 < \bar{S} < 105$  (triângulos),  $105 < \bar{S} < 425$  (círculos), e  $980 < \bar{S} < 2,160$  (quadrados). (a) Densidade de probabilidade condicional  $p(R|S)$  das taxas de crescimento  $R$ . As linhas sólidas são dadas pela Equação 2.2 usando  $\mu$  e  $\sigma$  obtido diretamente dos dados. (b) Distribuição cumulativa  $p_c(|R|)$  do valor absoluto de  $R$ . As linhas sólidas são ajustes lineares dos dados (na escala mono-log), representando o decaimento exponencial. (c) Densidade de probabilidade condicional  $p(r|S)$  normalizada das taxas de crescimento  $r$ . (d) Desvio padrão  $\sigma$  das taxas de crescimento  $R$  como uma função de  $S$ . A linha sólida é um ajuste linear para os dados (em escala log-log), dando um expoente de escala  $\beta \simeq 0,16$ . Adaptada de Picoli e Mendes (2008).

Outras medidas importantes também foram analisadas como: média mensal de editores em tempo integral ( $N$ ), o total de tempo gasto em horas no trabalho educativo público ( $H$ ) e a média mensal de cursos realizados nas casas ( $C$ ). Descobriu-se que a densidade da probabilidade de  $\log_{10}N$ ,  $\log_{10}H$  e  $\log_{10}C$  são consistente com uma distribuição Gaussiana, implicando que  $N$ ,  $H$  e  $C$  são distribuições lognormal, o desvio padrão das taxas de crescimento como uma lei de potência, com expoentes  $\beta \simeq 0,16$  para horas,  $\beta \simeq 0,13$  para cursos e  $\beta \simeq 0,10$  para os pioneiros [Figura 2.1.4 (a)], e a densidade de probabilidade condicional das taxas de crescimento normalizada,  $p(r|N)$ ,  $p(r|H)$ , e  $p(r|C)$  são consistentes com a curva universal definida pela Equação 2.4

[Figura 2.1.4 (b)]. Essas distribuições foram obtidas, respectivamente, dos seguintes grupos de países:  $979 < \bar{S} < 2157$  e  $123 < \bar{S} < 424$ ;  $230,508 < \bar{H} < 596,327$  e  $26,431 < \bar{S} < 79,627$ ;  $1,014 < \bar{C} < 2,664$  e  $122 < \bar{C} < 390$ ;  $113 < \bar{N} < 328$  e  $15 < \bar{N} < 44$ .

Encerrando as investigações, analisou-se a distribuição das taxas de crescimento para as seguintes variáveis: cursos por editores ( $C/S$ ), pioneiros por editora ( $N/S$ ), e horas por editores ( $H/S$ ). A Figura 2.1.4 (c) mostra a distribuição das taxas de crescimento normalizadas de 140 países no período de 1959-2005, em comparação com a Equação 2.4.



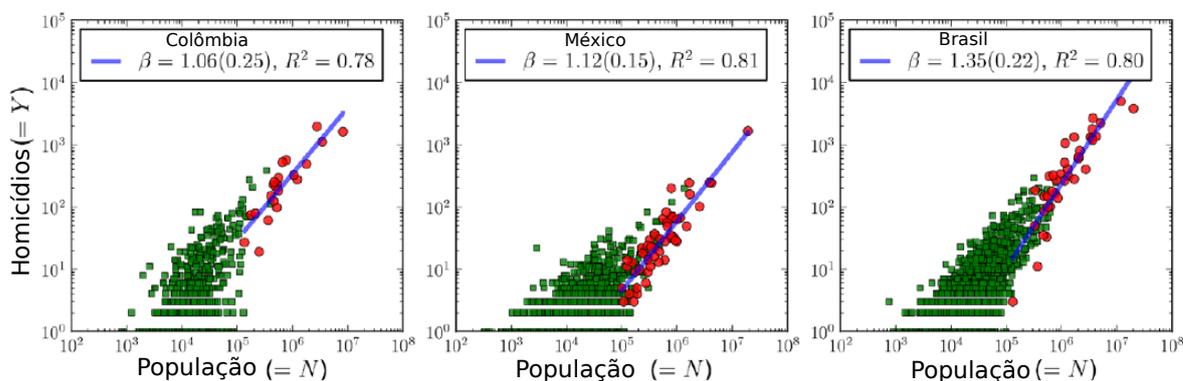
**Figura 2.1.4 – Dinâmica de crescimento de outras atividades religiosas.** (a) Desvio padrão  $\sigma$  das taxas de crescimento em função das seguintes medidas de tamanho ( $X$ ): horas  $H$  (quadrados), cursos  $C$  (círculos), e pioneiros  $N$  (triângulos). As linhas sólidas são ajustes lineares aos dados (na escala  $\log - \log$ ). (b) Função densidade de probabilidade condicional  $p(r|S)$ ,  $p(r|H)$ ,  $p(r|C)$ , e  $p(r|N)$ . A linha sólida é dada pela Equação 2.4. (c) Densidade de probabilidade  $p(r)$  das taxas de crescimento normalizadas  $r$ , calculadas para 140 países, nas seguintes variáveis:  $C/S$  (quadrados),  $N/S$  (triângulos), e  $H/S$  (círculos). A linha sólida é dada pela Equação 2.4. Adaptada de Picoli e Mendes (2008).

Os resultados mostrados na Figura 2.1.4 sugerem que as mesmas leis gerais se aplicam a diferentes medidas de atividade religiosa.

## 2.2 Escalas urbanas conectadas com a Lei de Zipf

Com o objetivo de verificar e identificar relações de escala em indicadores urbanos Gomez-Lievano et al. (2012), construíram uma ferramenta estatística que caracteriza a distribuição de probabilidade conjunta de indicadores urbanos dado o tamanho da população através de sistemas urbanos. Para isso, utilizaram dados sobre homicídios em três países latino-americanos, Brasil (2003-2007), Colômbia (2004-2009) e México (2004-2009), nos quais as taxas nacionais variaram substancialmente. Obtiveram, usando a regra de Bayes juntamente com a probabilidade condicional do número de homicídios por ano, dado o tamanho da população de uma cidade. Em seguida, mostrou-se que as leis de escala emergem como valores de expectativa dessas estatísticas condicionais.

A Figura 2.2.1 mostra o total de homicídios versus o tamanho da população do ano 2007. As variáveis  $Y$  e  $N$ , são respectivamente, número de homicídios e tamanho da população. A linha sólida se encaixa na escala de homicídios apenas para áreas metropolitanas.



**Figura 2.2.1 – Número anual de homicídios nas cidades da Colômbia, do México e do Brasil versus o tamanho da população (2007).** Grandes cidades são definidas em termos de área metropolitana que são agregações de municípios (círculos vermelhos), enquanto municípios não metropolitanos são apresentados separadamente (quadrados verdes). A linha azul ajusta apenas os dados de homicídio para áreas metropolitanas. Adaptada de Gomez-Lievano et al. (2012).

Grandes variações, especialmente entre as unidades populacionais menores, e o fato de que muitos municípios tem  $Y = 0$  (não mostrado) impedem uma análise de escala direta. No entanto, é possível analisar os dados consistentemente através da estimativa de probabilidades condicionais, supondo que as variáveis  $Y$  e  $N$  são variáveis estocásticas e que seus valores em cada cidade específica e tempo são realizações estatísticas.

Assim, pode-se estimar as distribuições de probabilidade, usando a regra de

Bayes em que a função de distribuição da probabilidade condicional de  $Y$ , dada uma cidade de população  $N$ ,  $P(Y|N)$ ,

$$P(Y|N) = \frac{P(N|Y)P(Y)}{P(N)}, \quad (2.6)$$

para calculá-lo, dado o conhecimento da distribuição de probabilidade  $P(Y)$  de homicídios nas cidades, independentemente de sua população e da distribuição de probabilidade condicional  $P(N|Y)$  para o tamanho populacional de cidades com um dado número de homicídios. O denominador é uma constante em  $Y$  e pode ser expresso como o traço do numerador sobre todos os valores de  $Y$ .  $P(N)$  é a função densidade de probabilidade de Zipf para tamanhos populacionais.

Para estimar a distribuição dos homicídios totais nas cidades  $P(Y)$  deve levar em conta o fato de que as propriedades urbanas mudam (super) extensivamente com a população e que há cidades com tamanhos muito diferentes. Devido a esses fatos gerais, as densidades de probabilidade da lei de potência (distribuições Zipf ou Pareto), são comuns entre as métricas urbanas. Mais especificamente, essas distribuições explicam o fato de que um pequeno número de cidades é responsável pela maioria dos homicídios e que um grande número de cidades exibe apenas alguns. No México, por exemplo, aproximadamente 60% dos homicídios vêm de 2% das cidades. Números semelhantes caracterizam a Colômbia e o Brasil nos anos estudados.

Para evitar o ruído da cauda para as grandes cidades adotou-se, na prática, um procedimento de analisar a função distribuição acumulada complementar ao invés da função densidade de probabilidade.

A Figura 2.2.2 mostra as distribuições cumulativas empíricas de homicídios para o ano de 2007 na Colômbia, México e Brasil, as quais apresentam cauda pesada por várias décadas e um corte efetivo menor para valores pequenos de  $Y$ . Com isso supõe-se o modelo

$$P(Y) = \frac{(Y + \kappa)^{-\tau}}{\zeta(\tau, \kappa)}, Y \in \mathbb{N}, \quad (2.7)$$

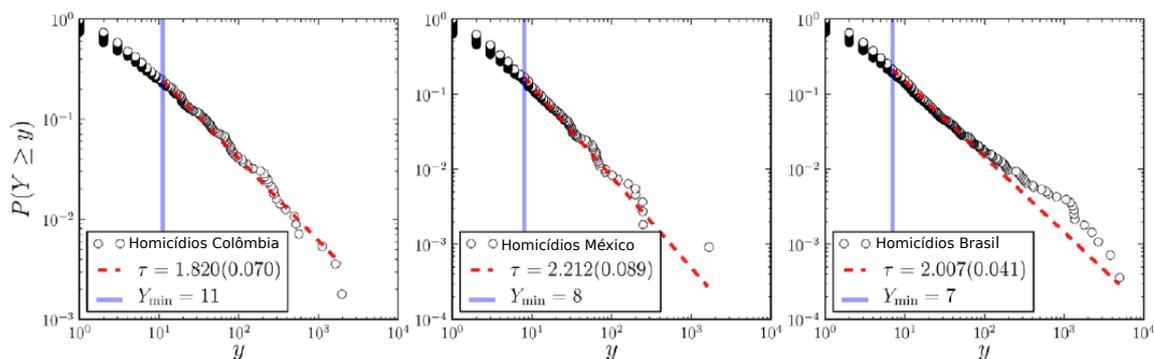
onde  $\tau > 0$  é o expoente da lei de potência e  $\kappa$  é um número real positivo, que permite  $P(Y)$  permanecer analítico com  $Y \rightarrow 0$ , em que

$$\zeta(\tau, \kappa) = \sum_{Y=0}^{\infty} (Y + \kappa)^{-\tau} \quad (2.8)$$

é a função zeta generalizada ou zeta de Hurwitz, que garante a normalização de  $P(Y)$  como uma variável discreta.

As distribuições cumulativas normalizadas de homicídios na Colômbia, México e Brasil são bem descritas por distribuições lei de potência, como pode ser observado na Figura 2.2.2. Para atenuar as flutuações da cauda e facilitar a interpretação visual não analisou-se a função densidade, mas a distribuição cumulativa complementar. Os

melhores ajustes da forma  $P(Y \geq y) = Cy^{-\tau+1}$  foram estimados usando o procedimento para a função densidade. Enquanto a distribuição de homicídios totais é invariante por escala, este é o resultado de traçar distribuições condicionais mais previsíveis para cada cidade sobre uma ampla distribuição de tamanhos de cidades.

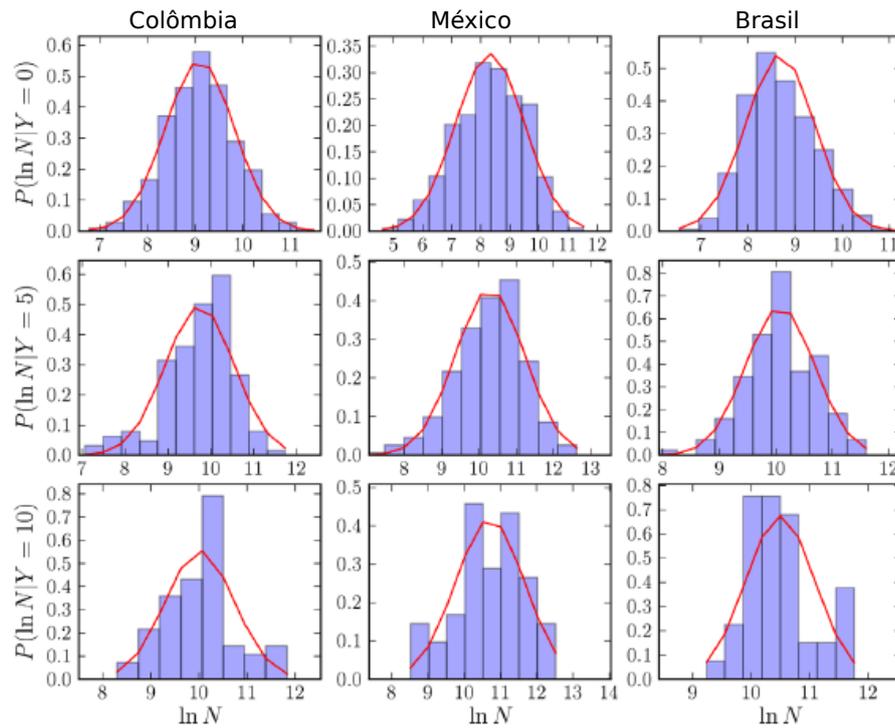


**Figura 2.2.2 – Distribuições normalizadas cumulativas de homicídios na Colômbia, México e Brasil.** A linha vermelha tracejada mostra os ajustes. Os erros padrão são mostrados entre parênteses. A linha azul sólida mostra o valor mínimo de  $Y$  para o qual um ajuste de lei de potência é válido. Adaptada de Gomez-Lievano et al. (2012).

Para calcular  $P(N|Y)$ , fixou-se o valor de  $Y$  e estimou-se a distribuição de probabilidade sobre a população. A Figura 2.2.3 mostra os histogramas das frequências de homicídios para Colômbia, México e Brasil, para uma faixa de  $Y$ . Esses números dão uma impressão do tipo da distribuição de probabilidade que descreve os dados. Observa-se que todas as distribuições, em cada valor de  $Y$ , mostram um pico distinto com média e variância definidas. A hipótese nula da distribuição de Poisson foi rejeitada com alta confiança por um método de máxima verossimilhança. Em vez disso, os dados se encaixam bem em termos de uma distribuição lognormal:

$$P(N|Y) = \frac{1}{N\sqrt{2\pi\sigma_Y^2}} e^{-\frac{(\ln N - \mu_Y)^2}{2\sigma_Y^2}}, \quad (2.9)$$

onde  $\mu_Y$  e  $\sigma_Y$  são parâmetros dependendo de  $Y$ .



**Figura 2.2.3 – Histogramas de frequência normalizada do logaritmo da população da cidade para um número variável de homicídios observados  $Y$ .** Uma distribuição lognormal (note que o eixo  $x$  é expresso em termos de  $\ln N$ ) é mostrada como uma linha vermelha sólida, com parâmetros obtidos via estimativa máxima verossimilhança. Adaptada de Gomez-Lievano et al. (2012).

Segue-se que a média e a variância são dadas por:

$$\langle N \rangle_Y = e^{\mu_Y + \sigma_Y^2/2} \quad (2.10)$$

$$(\Delta N_Y)^2 = (e^{\sigma_Y^2} - 1)e^{2\mu_Y + \sigma_Y^2}. \quad (2.11)$$

Os estimadores de máxima verossimilhança dos parâmetros lognormal são:

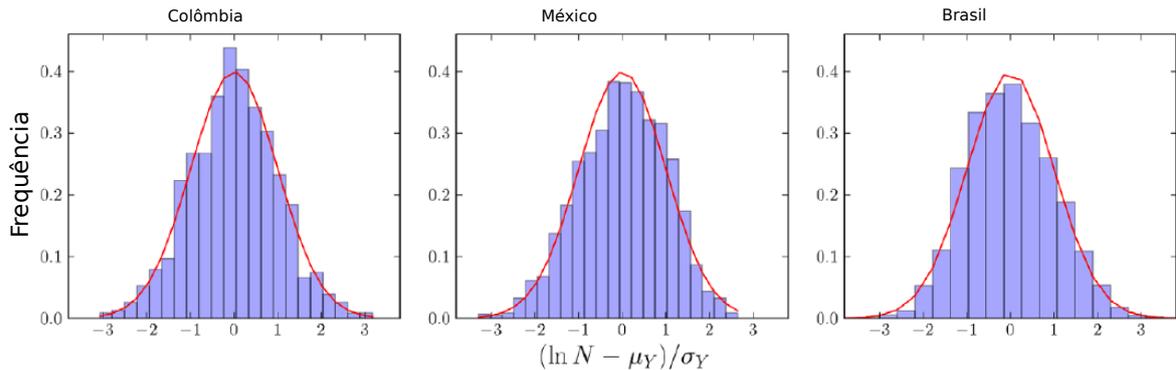
$$\widehat{\mu}_Y = \frac{1}{n_Y} \sum_{i \in S_Y} \ln N_i \quad (2.12)$$

$$\widehat{\sigma}_Y^2 = \frac{1}{n_Y} \sum_{i \in S_Y} (\ln N_i - \widehat{\mu}_Y)^2 \quad (2.13)$$

onde  $S_Y$  é o conjunto de cidades com  $Y$  homicídios e  $n_Y$  o número de cidades em  $S_Y$ .

Se a distribuição normal mantém em termos as variáveis logarítmicas da população, dado diferentes valores de  $Y$ , pode-se colapsar os diferentes histogramas da Figura 2.2.3 padronizando as log-variáveis. Isso é possível calculando os estimadores de máxima verossimilhança da média e da variância para cada valor de  $Y$  e, em

seguida, construindo no mesmo histograma a distribuição para vários valores de  $Y$ . A Figura 2.2.4 mostra essas distribuições padronizadas. Esse procedimento tem suas limitações devido ao fato de que à medida que aumentamos  $Y$ , o número de cidades decresce, até que haja apenas uma cidade com dado  $Y$  e  $N$  e a estimativa estatística se torne impossível. Por outro lado, tem a vantagem de que a forma da distribuição  $P(N|Y)$  para vários valores de  $Y$  pode ser exibida em uma única figura.



**Figura 2.2.4 – Colapsos dos histogramas de  $P(N|Y)$  através de valores de  $Y$  em 2007.** As funções densidade de probabilidade lognormal para as três nações são mostradas como linhas vermelhas sólidas. Adaptada de Gomez-Lievano et al. (2012).

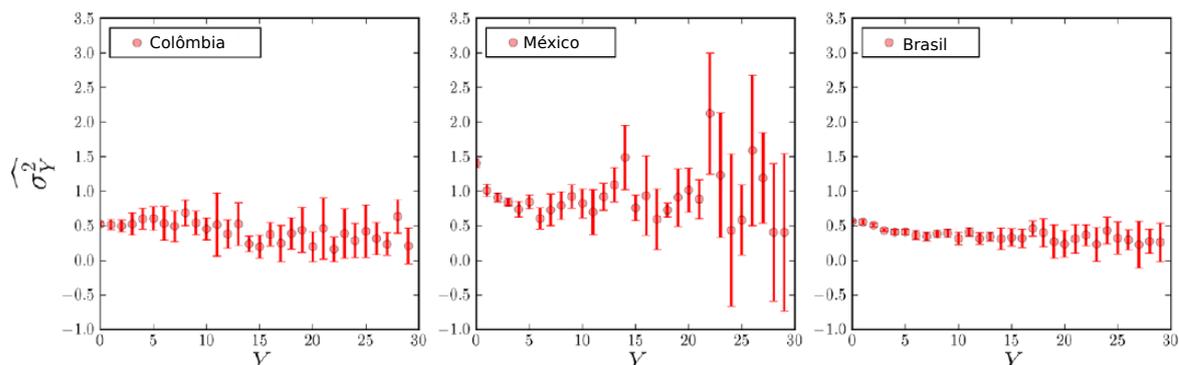
Note que, as distribuições da lei de potência que descreve o número total de homicídios nos sistemas urbanos têm, de fato, estatísticas mais previsíveis quando condicionadas ao tamanho populacional da cidade.

Agora, pode-se estimar os parâmetros da Equação 2.9 usando as Equações 2.12 e 2.13, e plotando  $\widehat{\sigma_Y^2}$  e  $\widehat{\mu_Y}$  inferindo sua dependência funcional de  $Y$ .

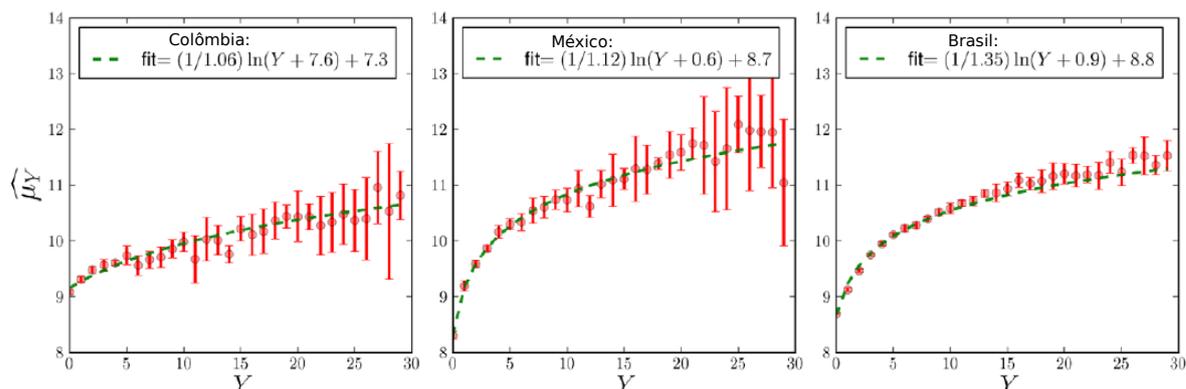
O comportamento de  $\widehat{\sigma_Y^2}$  mostrado na Figura 2.2.5 é estável e supôs-se constante a partir daqui. As curvas mostradas na Figura 2.2.6 sugerem um crescimento logarítmico de  $\widehat{\mu_Y}$  em função de  $Y$ . A função logarítmica mais geral que pode ser ajustada a  $\widehat{\mu_Y}$  é:

$$\widehat{\mu_Y} = f(Y) = b \ln(Y + r) + \ln A \quad (2.14)$$

onde  $r$  é uma constante positiva que permite o logaritmo permanecer finito (e positivo) quando  $Y \rightarrow 0$  e  $A$  é um número positivo. Adiante, a constante  $b$  será identificada com o expoente de escala  $1/\beta$ . Note que, cada país tem uma variação característica de seus indicadores condicionados a outras quantidades urbanas. A esse respeito, é interessante notar as semelhanças entre a Colômbia e o Brasil.



**Figura 2.2.5 – Estimativas de  $\sigma_Y^2$  (via máxima verossimilhança) para diferentes valores de  $Y \in \{0, \dots, 29\}$ , para a Colômbia, México e Brasil.** Uma curva diferente foi construída para cada ano da análise. Os pontos mostrados são as médias ao longo de vários anos. As barras de erro representam intervalos de um desvio padrão em torno da média (nível de confiança de 67%). As figuras não mostram uma dependência sistemática clara de  $Y$  em  $\sigma_Y^2$ . Adaptada de Gomez-Lievano et al. (2012).



**Figura 2.2.6 – Estimativas de  $\mu_Y$  (via máxima verossimilhança) para diferentes valores de  $Y \in \{0, \dots, 29\}$ , para a Colômbia, México e Brasil.** Uma curva diferente foi construída para cada ano da análise, e os pontos traçados são as médias ao longo de vários anos. As barras de erro representam um intervalo de um desvio padrão sobre a média. Adaptada de Gomez-Lievano et al. (2012).

Os gráficos mostram uma dependência logarítmica de  $Y$ , da qual uma relação de escala surge em termos de valores esperados. Os melhores ajustes foram obtidos usando um algoritmo de Levenberg-Marquardt, ponderando cada ponto por seu erro.

Usando a Equação 2.6, deriva-se a função de probabilidade condicional  $P(Y|N)$ . Se a Equação 2.9 é válida para todos  $Y \geq 0$ , usando a Equação 2.7, obtém-se:

$$P(Y|N) \propto (1/\tilde{Y}) \exp \left[ -\frac{1}{2\sigma_Y^2} (\ln N - \mu_Y)^2 + (1 - \tau) \ln \tilde{Y} \right], \quad (2.15)$$

onde  $\tilde{Y} \equiv Y + \kappa$ .

Usando a Equação 2.14 para substituir  $\mu_Y$  por  $\widehat{\mu}_Y$ , obtém-se

$$P(Y|N) \propto (1/\widetilde{Y}) \exp \left[ -\frac{1}{2\sigma_Y^2} ((\ln N - \ln A(Y^*)^b)^2 + 2\sigma_Y^2(\tau - 1)\ln \widetilde{Y}) \right]. \quad (2.16)$$

Expandindo os termos ao quadrado, os logaritmos e agrupando alguns dos termos, essa equação se transforma em:

$$P(Y|N) \propto (1/\widetilde{Y}) \exp \left[ -\frac{1}{2\sigma_0^2} (\ln^2 Y^* - 2\ln(\frac{(Y^*)^P}{\widetilde{Y}\sigma_0^2(\tau - 1)}) + P^2) \right], \quad (2.17)$$

onde  $Y^* = Y + r$ ,  $P = \frac{1}{b} \ln(N/A)$  e  $\sigma_0 = (\sigma_Y/b)$ .

Lembrando que  $r$  em  $Y^* = Y + r$  e  $\kappa$  em  $\widetilde{Y} = Y + \kappa$  foram introduzidos para explicar o limite quando  $Y \rightarrow 0$ . Essas constantes geram limites esperados e nos impedem de dividir por zero na distribuição da lei de potência e de tomar o logaritmo de zero em  $\widehat{\mu}_Y$ . Não há restrições que impeçam de considerá-los iguais e pequenos, pois ambos introduzem uma escala característica que se manifesta como uma mudança de regime no comportamento de escala quando as cidades são muito pequenas e as realizações de homicídios nulos (ou outras medidas discretas) começam a ocorrer. Portanto, não é razoável supor que eles são os mesmos, assim  $Y^* \approx \widetilde{Y}$ . Sob esta suposição, pode-se completar o quadrado e calcular a distribuição posterior. Percebendo que  $P(Y|N) = P(Y^*|N)$  porque  $\Delta Y^*/\Delta Y = 1$ , e mantendo apenas termos dependentes de  $Y$  (os outros serão finalmente absorvidos pela constante de normalização), chega-se a

$$P(Y^*|N) \propto (1/Y^*) \exp \left[ -\frac{1}{2\sigma_0^2} (\ln Y^* - (P - \sigma_0^2(\tau - 1)))^2 \right], \quad (2.18)$$

que é uma distribuição lognormal para  $Y^*$  dado  $N$ , com parâmetros  $\mu_N = P - \sigma_0^2(\tau - 1)$  e  $\sigma_N = \sigma_0$ . Expressando os parâmetros de distribuição nas variáveis originais e introduzindo a constante de normalização adequada, obtém-se

$$P(Y^*|N) = \frac{1}{Y^* \sqrt{2\pi\sigma_N^2}} e^{-\frac{(\ln Y^* - \mu_N)^2}{2\sigma_N^2}}, \quad (2.19)$$

$$\mu_N = \frac{1}{b} \ln\left(\frac{N}{A}\right) - \sigma_N^2(\tau - 1) \quad (2.20)$$

$$\sigma_N^2 = \sigma_Y^2/b^2. \quad (2.21)$$

As expressões acima relacionam as estatísticas lognormal da distribuição condicional  $P(Y|N)$  com a escala e a lei de Zipf para a distribuição do tamanho das cidades. Como mostra abaixo, isso leva a um relacionamento entre o escalonamento e os expoentes de Zipf.

Com essas distribuições condicionais é possível calcular seus momentos, como a média e a variância. Tomando a Equação 2.14 e a Equação 2.20 para derivar  $\langle N \rangle_Y$  e  $\langle Y \rangle_N$  explicitamente em termos de  $Y$  e  $N$ , isto é:

$$\langle Y + r \rangle_N = \left( \frac{e^{(3/2-\tau)\sigma_N^2}}{A^\beta} \right) N^\beta \quad (2.22)$$

$$\langle N \rangle_Y = A e^{\sigma_Y^2/2} (Y + r)^{1/\beta}. \quad (2.23)$$

Da mesma maneira, os desvios padrão  $\Delta Y_N^* = \Sigma_N N^\beta$  e  $\Delta N_Y$  podem ser expressos como

$$\Delta Y_N^* = \Sigma_N N^\beta \quad (2.24)$$

$$\Delta N_Y = \Sigma_Y (Y + r)^{1/\beta}, \quad (2.25)$$

onde  $\Sigma_N$  e  $\Sigma_Y$  são coeficientes de proporcionalidade.

Pode-se mostrar como uma distribuição da lei de potência emerge ao derivar a distribuição da probabilidade de  $N$ . Na Equação 2.6,  $P(N)$  é chamado de “evidência”, e age na prática como uma constante de normalização. Podendo ser calculado a partir do conhecimento do numerador

$$P(N) = \Sigma_{Y^*=r}^\infty P(N|Y^*)P(Y^*) \quad (2.26)$$

$$\propto N^{-\alpha}, \quad (2.27)$$

que é uma distribuição da lei de potência. Segue-se então que os vários expoentes são obrigados a obedecer à relação

$$\beta = \frac{\alpha - 1}{\tau - 1}. \quad (2.28)$$

Se a escala superlinear ( $\beta > 1$ ) se aplica a algum indicador urbano  $Y$ , pode-se prever que os tamanhos populacionais sejam distribuídos pela lei de potência com expoente  $\alpha > \tau$ , e vice-versa, se a escala for sublinear. Se  $\alpha \approx 2$ , significa que a escala superlinear  $\tau < 2$  e, portanto, a quantidade  $Y$  pode não ter média e variância definidas. Nestes casos, as referências as “cidades médias” não têm significado matemático. Note, no entanto, que também é possível que  $\alpha > \tau > 2$  desde que o expoente de Zipf é suficientemente maior que 2. Em geral, essas propriedades podem ser usadas para restringir o valor do expoente de Zipf a partir da observação das estatísticas de muitos indicadores urbanos diferentes e do conhecimento de suas propriedades médias de escala.

### 2.3 Uma relação entre a escala urbana e função de produção

Essa seção apresenta uma investigação de Lobo et al. (2013) na qual a partir de relações de escala urbana de quantidades econômicas versus população derivou-se um modelo de saída econômica na forma de uma função de produção do tipo Cobb-Douglas. A principal contribuição deste estudo foi abordar a questão de como formas específicas de funções de produção, comuns a todas as cidades, surgem como modelos eficazes de produção econômica como resultado da observação das relações de escala urbana e seus fundamentos teóricos.

Um resultado observado para as cidades é que a maioria de suas propriedades não são proporcionais ao tamanho da população. Por exemplo, as cidades maiores tendem a exibir maiores produções per capita, desde crime a salários, e precisam de uma infraestrutura com menos material por pessoa, embora a utilizem mais intensamente. Essas propriedades, e sua detalhada expressão quantitativa observada em termos de relações de escala, podem ser derivadas de uma teoria microscópica que descreve as cidades como co-localizadas misturando redes sociais, sujeitas a certas restrições gerais de eficiência.

As relações de escala caracterizam-se como uma determinada quantidade de interesse,  $Y$ , dependendo de uma medida do tamanho de um sistema,  $N$ , de acordo com a relação

$$Y(N) = Y_0 N^\beta, \quad (2.29)$$

onde  $Y_0$  é uma constante de normalização e  $\beta$  o expoente de escala. A relevância desta relação lei de potência torna-se clara quando considera-se uma mudança de escala arbitrária por um fator  $\lambda$  de  $N$  para  $\lambda N$ . Isso induz uma mudança em  $Y$  de  $Y(N)$  para  $Y(\lambda N)$  que pode ser expressa como

$$Y(\lambda N) = Z(\lambda, N)Y(N). \quad (2.30)$$

Se o fator de escala  $Z$  depende apenas de  $\lambda$ , isto é,  $Z(\lambda N) = Z(\lambda)$ , a Equação 2.30 pode ser resolvida de forma única fornecendo o resultado invariante de escala da Equação 2.29, com  $Z(\lambda) = \lambda^\beta$ . Invariância de escala implica que a relação  $Y(\lambda N)/Y(N)$  é parametrizada por um único número adimensional,  $\beta$ . A relação  $Y(\lambda N)/Y(N)$  é independente do tamanho do sistema particular  $N$ , mas depende da relação entre os tamanhos,  $\lambda$ ; tais sistemas são frequentemente referidos como autossimilares.

A Equação 2.29 assemelha-se com uma função de produção, sendo  $Y$  a produção econômica total e  $N$  o tamanho da população urbana ou grupos de trabalho. Em uma base per capita, essa mesma equação implica  $y \equiv Y/N = Y_0 N^{\beta-1}$ , que pode ser interpretada, por exemplo, como uma equação para a saída por pessoa, em função do número máximo de pessoas compartilhando ideias entre si.

As relações de escala e funções de produção podem expressar expectativas médias para saídas (econômicas) em termos de conjuntos de entradas. Porém, a interpretação estatística correta das leis de escala são valores de expectativa para a quantidade  $Y$ , condicionada ao tamanho da população de uma cidade; essa é a média associada à densidade de probabilidade  $P(Y|N)$ .

As flutuações estatísticas em torno da lei de escala média, com os valores dos parâmetros de escala, podem ser determinadas aplicando  $\ln$  na Equação 2.29.

$$\ln Y_i = \ln Y_0 + \beta \ln N_i + \xi_i, \quad (2.31)$$

com áreas urbanas indexadas por  $i$ . Aqui, as flutuações estão representando desvios locais (específicos de uma cidade) da escala em forma de variável. Tendo como exemplo uma métrica urbana que exibe um comportamento de escala, considere o total de salários, definido como a soma total de salários e vencimentos auferidos pelos residentes em uma área urbana. A estimativa de mínimos quadrados ordinários da Equação 2.31 - corrigindo para heterocedasticidade, usando dados sobre salários totais ( $TW$ ) e população para as 943 áreas urbanas dos Estados Unidos durante o período de 2009-2011 fornece o seguinte resultado:

$$\ln(TW_i) = 1,404 + 1,146 \ln(Pop_i), R^2 = 0,97, \quad (2.32)$$

com p-valores virtualmente zero. A Figura 2.3.1 mostra o gráfico de dispersão dos dados e a linha de regressão ajustada; o gráfico da Figura 2.3.2 mostra que eles são independentes de escala. Assim, um aumento de 1% na população está associado, em média, a um aumento de 1,15% na produção, independentemente do tamanho da cidade, em geral de acordo com as expectativas teóricas para  $\beta \sim 7/6$ . Esses retornos de escala semelhantes e crescentes estabelecem quantitativamente as vantagens econômicas das grandes cidades.

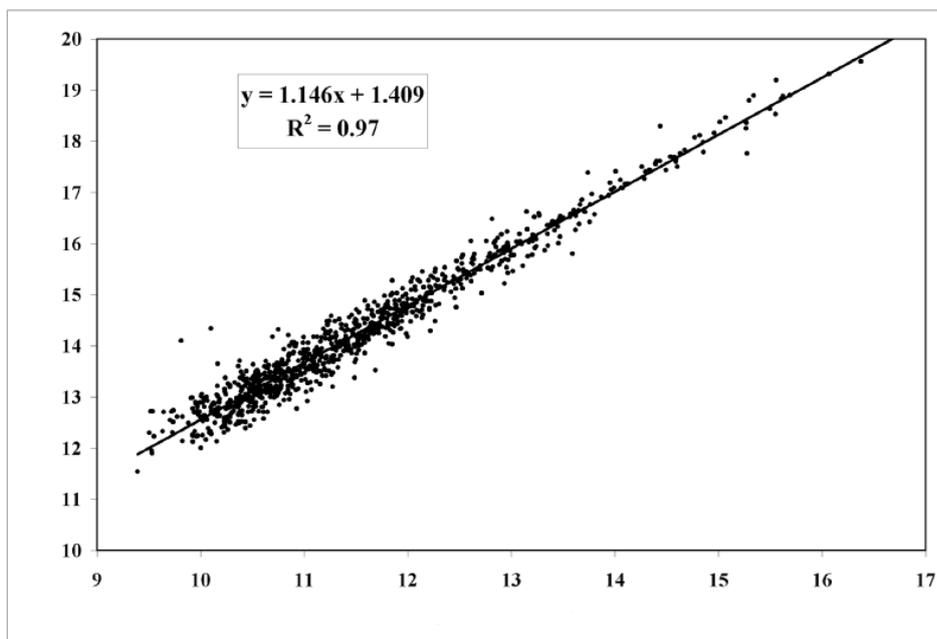
A Equação 2.32 indica a produtividade média de uma cidade com tamanho  $N$ . Os desvios médio desse comportamento detêm as características de cada área urbana individual não esclarecidas pelos efeitos gerais de aglomeração do tamanho da população. Estes desvios podem ser quantificados escrevendo a equação residual em 2.31 como

$$\xi_i = \ln \frac{Y_i}{Y(N_i)} = \ln \frac{Y_i}{Y_0 N_i^\beta} \quad (2.33)$$

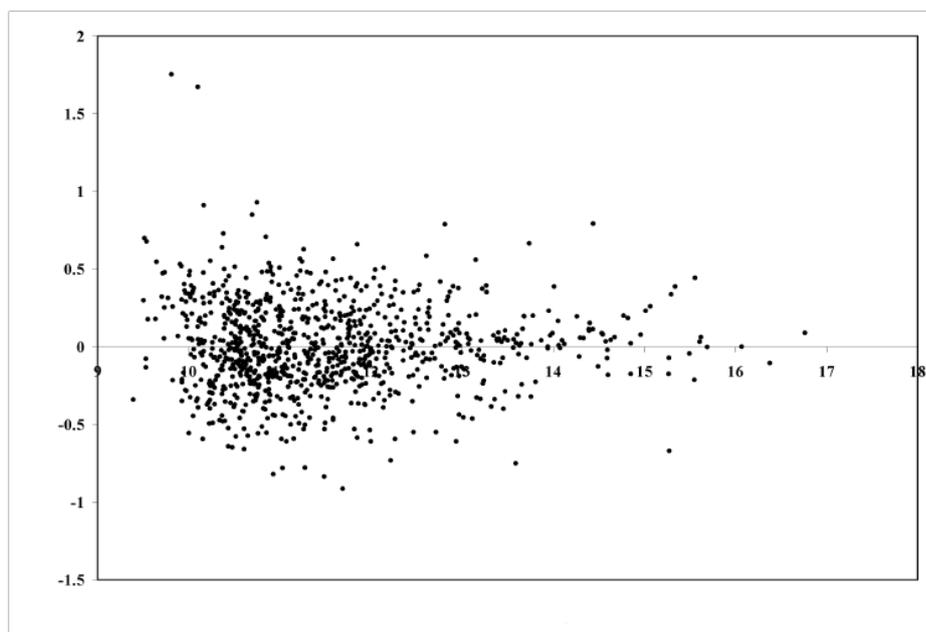
onde  $Y_i$  é o valor observado de saída para cada área metropolitana e  $\xi$  é um indicador metropolitano ajustado à escala (SAMI). Como os SAMIs podem ser construídos para quaisquer características variáveis de captura da vida urbana que estão sujeitas a efeitos de aglomeração em escala, pode-se escrever qualquer indicador urbano estocástico, como

$$Y_i = Y_0 N_i^\beta e^{\xi_i}. \quad (2.34)$$

Agora pode-se derivar a função de produção econômica das cidades a partir de suas propriedades probabilísticas de escala.



**Figura 2.3.1 – Gráfico de dispersão dos dados e linha de regressão ajustada.** Escala do total de salários usando dados para todas as 943 áreas urbanas dos Estados Unidos (suavizadas no período de 2009-2011) mostrando escala superlinear. Adaptada de Lobo et al. (2013).



**Figura 2.3.2 – Resíduos da regressão de  $\ln$  (total de salários) em  $\ln$  (população).** Na regressão foi usados dados para todas as 943 áreas urbanas dos Estados Unidos (suavizadas no período de 2009-2011). Adaptada de Lobo et al. (2013).

Prosseguindo-se com uma relação contábil onde em algum momento,  $t$ ,

$$Y_i(t) = W_i(t) + R_i(t) \quad (2.35)$$

sendo  $Y$ , o valor pecuniário da saída total gerada na  $i$ -ésima área metropolitana,  $W$  a renda total do trabalho e  $R$  a renda total de capital. As ações do fator de produção são definidas como:

$$1 - \alpha = \frac{W_i(t)}{Y_i(t)}, \alpha = \frac{R_i(t)}{Y_i(t)}. \quad (2.36)$$

Em geral,  $\alpha = \alpha_i(t, N_i)$  é específica da cidade e de uma função de tempo e tamanho da população  $N$ . Diferenciando a Equação 2.35 em relação ao tempo (ou em relação a  $N$ ) e dividindo por saída,  $Y$ , obtém-se

$$\frac{1}{Y_i(t)} \frac{dY_i(t)}{dt} = \frac{1}{Y_i(t)} \frac{dW_i(t)}{dt} + \frac{1}{Y_i(t)} \frac{dR_i(t)}{dt} = \frac{[1 - \alpha(t)]}{W_i(t)} \frac{dW_i(t)}{dt} + \frac{\alpha(t)}{R_i(t)} \frac{dR_i(t)}{dt}. \quad (2.37)$$

Integrando 2.37, tem-se

$$\ln Y_i(t) = \int (1 - \alpha) d \ln W_i(t) + \int \alpha d \ln R_i(t). \quad (2.38)$$

Integrando por partes, obtém-se o seguinte resultado:

$$\ln Y_i(t) = (1 - \alpha) \ln W_i(t) + \alpha \ln R_i(t) + \int \ln \left( \frac{W_i(t)}{R_i(t)} \right) d\alpha. \quad (2.39)$$

A última integral pode ser escrita como

$$\int \ln \left( \frac{1 - \alpha}{\alpha} \right) d\alpha = \ln [c(1 - \alpha)^{\alpha-1} \alpha^{-\alpha}], \quad (2.40)$$

onde  $c$  é uma constante de integração, de modo que,

$$Y_i(t) = c(1 - \alpha)^{\alpha-1} \alpha^{-\alpha} W_i(t)^{1-\alpha} R_i(t)^\alpha. \quad (2.41)$$

Note que a Equação 2.41 é uma instanciação de uma relação mais geral, que pode ser derivado algebricamente e que para o fator livre ser independente dos fatores de produção  $\alpha$  deve ser uma constante. Essa derivação destaca que  $Y$ ,  $W$  e  $R$  são funções do tempo.

Restringindo a solução na Equação 2.39 para ser consistente com a Equação original 2.35 determina  $c = 1$ . Essa solução é geral na medida em que não exige, por exemplo, que o fator compartilhado,  $\alpha$ , seja constante no tempo ou no tamanho da população. Assim, a derivação de uma função de produção do tipo Cobb-Douglas segue diretamente das definições 2.35 e 2.36 e não possui significado econômico mais específico além do contido nessas relações. De fato, a função de produção Cobb-Douglas é basicamente uma identidade trivial que decorre de um simples argumento

dimensional, já que  $Y$ ,  $W$  e  $R$  devem ter as mesmas dimensões, e supondo que  $Y$  é unicamente por  $W$  e  $R$ , deve ser expressável como a Equação 2.41, com expoentes somando a unidade. No entanto, esse formalismo assume uma utilidade potencialmente maior quando  $\alpha$  é uma constante independente do tempo e do tamanho da população; isto é, quando

$$\frac{\partial \alpha}{\partial N} \Big|_{t=0} = 0, \frac{\partial \alpha}{\partial N} \Big|_{N=0} = 0. \quad (2.42)$$

A Equação 2.39, sob a suposição de  $\alpha$  constante, pode ser relacionada à função de produção Cobb-Douglas, que é um modelo amplamente utilizado para economias nacionais e urbanas. Isto requer a introdução dos fatores de conversão relativos a salários,  $W_i(t)$ , à mão-de-obra,  $L_i(t)$ , e renda do capital,  $R_i(t)$ , à entrada do capital,  $K_i(t)$ :

$$w_i(t) = \frac{W_i(t)}{L_i(t)}, r_i(t) = \frac{R_i(t)}{K_i(t)}, \quad (2.43)$$

$w$  é o salário médio e  $r$  é o preço médio de locação do capital. Assim, pode-se escrever,  $Y$  como

$$Y_i(t, N) = C(\alpha) W_i(t, N)^{1-\alpha} R_i(t, N)^\alpha = A_i(t, N) L_i(t, N)^{1-\alpha} K_i(t, N)^\alpha, \quad (2.44)$$

com  $C(\alpha) \equiv (1 - \alpha)^{-(1-\alpha)} \alpha^{-\alpha}$ . O pré-fator  $A(t, N)$  é frequentemente referido como a produtividade total de fatores (PTF) da área urbana e é medida preferida da produtividade econômica. Uma maior ou menor PTF multiplica os mesmos fatores de entrada de mão-de-obra e capital para produzir maior ou menor produção econômica, respectivamente. A partir da Equação 2.44, obtém-se a seguinte expressão para a PTF urbana em função da produtividade do trabalho e capital:

$$A_i(t, N_i) = C(\alpha) \left( \frac{W_i(t, N_i)}{L_i(t, N_i)} \right)^{1-\alpha} \left( \frac{R_i(t, N_i)}{K_i(t, N_i)} \right)^\alpha = C(\alpha) w_i(t, N_i)^{1-\alpha} r_i(t, N_i)^\alpha. \quad (2.45)$$

Em seguida, mostra-se como a existência de relações de escala determina a forma de  $A$ , resultando em sua parametrização sistemática como uma função explícita do tamanho da população,  $N_i$ , e desvios locais,  $\xi_i$ .

Até aqui, explorou-se as consequências de uma relação contábil, Equação 2.35, e a definição de ações fatoriais, Equação 2.36, juntamente com as leis de conservação expressas na Equação 2.42, para obter uma função do tipo Cobb-Douglas comum para todas as cidades. Mostra-se agora, que a constância de  $\alpha$  é uma consequência das relações de escala urbana e a dinâmica microscópica subjacente, e usa-se essas relações para obter uma nova expressão para  $A_i(t, N_i)$ . Note que, com

$$W_i(t, N_i) = W_0 e^{\xi_i^w(t)} N_i^{\beta_w}(t), R_i(t, N_i) = R_0 e^{\xi_i^R(t)} N_i^{\beta_R}(t), \quad (2.46)$$

segue que

$$\alpha = \frac{R_0(t)}{Y_0(t)} e^{\xi_i^R(t) - \xi_i^Y(t)} N_i^{\beta_R - \beta_Y}(t) = 1 - \frac{W_0(t)}{Y_0(t)} e^{\xi_i^W(t) - \xi_i^Y(t)} N_i^{\beta_W - \beta_Y}(t). \quad (2.47)$$

Assim, para  $\alpha$  independente de  $N$  é equivalente a exigir que ambos os salários e rendas tenham o mesmo expoente escalar, de modo que  $\beta_W = \beta_R$ . Consequentemente, a observação da escala urbana superlinear socioeconômica universal e seus fundamentos teóricos implicam a conservação de  $\alpha$  versus  $N$  e Cobb-Douglas para a produção econômica das cidades versus o tamanho da população.

A constância de  $\alpha$  no tempo requer que os pré-fatores  $W_0$  e  $R_0$  compartilhem a mesma dependência de tempo, e que as diferenças entre os SAMIs para os salários totais específicos e renda total de capital, e o SAMI para a produção total também seja independente do tempo. O primeiro refere-se ao crescimento econômico (nacional) do sistema urbano e, como tal, espera-se que variem lentamente no tempo. O último muda lentamente no tempo, mas a análise de suas estatísticas revela que sua variância (lembrando que os SAMIs têm média zero) é aproximadamente independente do tempo, como tal, pode-se esperar que a média de  $\alpha$  sobre os SAMIs também é aproximadamente independente do tempo.

Assumindo a constância de  $\alpha$  a partir dos argumentos anteriores, agora deriva-se uma expressão explícita para a PTF das cidades. Em primeiro lugar, nota-se que tanto o numerador quanto o denominador nas expressões salário por trabalhador e renda média de capital exibem comportamento de escala para que a produtividade marginal dos dois fatores de produção possa ser reformulada usando seus SAMIs associados como:

$$w_i(t) = \frac{W_i(t, N_i)}{L_i(t, N_i)} = \frac{W_0 e^{\xi_i^W(t)} N_i(t)^{\beta_W}}{L_0 e^{\xi_i^L(t)} N_i(t)^{\beta_L}} = \frac{W_0}{L_0} e^{\xi_i^W(t) - \xi_i^L(t)} N_i(t)^{\beta_W - \beta_L}, \quad (2.48)$$

$$r_i(t) = \frac{R_i(t, N_i)}{K_i(t, N_i)} = \frac{R_0 e^{\xi_i^R(t)} N_i(t)^{\beta_R}}{K_0 e^{\xi_i^K(t)} N_i(t)^{\beta_K}} = \frac{R_0}{K_0} e^{\xi_i^R(t) - \xi_i^K(t)} N_i(t)^{\beta_R - \beta_K}. \quad (2.49)$$

O termo para PTF assume, então, a forma geral:

$$A_i(t) = A_0(t) e^{\xi_i^A} N_i^{\beta_A}(t), \quad (2.50)$$

com

$$A_0(t) = C(\alpha) \left[ \frac{W_0(t)}{L_0(t)} \right]^{1-\alpha} \left[ \frac{R_0(t)}{K_0(t)} \right]^\alpha, \quad (2.51)$$

$$\xi_i^A = (1 - \alpha)(\xi_i^W - \xi_i^L) + \alpha(\xi_i^R - \xi_i^K), \quad (2.52)$$

$$\beta_A = (1 - \alpha)(\beta_W - \beta_L) + \alpha(\beta_R - \beta_K). \quad (2.53)$$

As Equações 2.50 e 2.53 explicitam como a PTF urbana depende do tamanho da população, através dos expoentes de escala, quanto de flutuações locais independentes da escala por meio das SAMIs. A Equação 2.50 difere de uma formulação

padrão da PTF na qual os efeitos de aumento de produtividade da população são explicitamente controlados e os efeitos neutros da população explicitamente representados pela Equação 2.52. Como consequência, qualquer propriedade urbana adicional proposta para explicar uma maior ou menor produtividade de cidades específicas não vinculadas ao seu tamanho deve ser expressa em termos de sua contribuição para os SAMIs para  $W, L, R$  e  $K$ .

Avaliar  $A$  requer conhecimento de como  $K$ , o estoque de capital metropolitano, escala com tamanho urbano. Dados os valores observados para os coeficientes de escala para salários totais e trabalho,  $\beta_W \approx 1.15$  e  $\beta_L \approx 1$ , e com  $(1 - \alpha) \approx 0.7$  o primeiro termo à direita do sinal de igualdade na Equação 2.52 tem um valor de 0.11 sobre o valor do termo  $\alpha(\beta_R - \beta_K)$ . Sob a suposição de que o aluguel de capital,  $r$ , é constante, ou quase, em todas as áreas metropolitanas, e dado que  $R = r \times K$ , ou equivalente,  $R_0 N^{\beta_K} = r K_0 N^{\beta_K}$ . Para  $r$  ser constante, deve-se ter  $\beta_R = \beta_K$ . Portanto,  $\beta_A \approx 0.11$  implica que a produtividade urbana, medida pela PTF, aumenta cerca de 11% em cada duplicação da população.

## 2.4 Fatores que dão origem as relações de escala dos fatores urbanos relacionados ao tamanho da população

Na presente seção será contemplado resultados obtidos por Bettencourt (2013) em um estudo que busca os fatores que originam as relações de escala dos fenômenos urbanos sobre o tamanho da população. Segundo Bettencourt (2013) as grandes dificuldades de uma abordagem científica para as cidades são resultados de suas várias características interdependentes, como social, econômico, infraestrutural, e sistemas complexos espaciais que existem em formas semelhantes, no entanto mudando sobre um conjunto de escalas. Nesse sentido essa investigação teve como objetivo mostrar como todas as cidades podem evoluir de acordo com um pequeno conjunto de princípios básicos que operam localmente.

Ao longo de 40 anos, várias pesquisas empíricas alegam que não há nenhum tamanho especial para as cidades, de modo que a maioria das propriedades urbanas,  $Y$ , variam continuamente com o tamanho da população e são descritas matematicamente em termos de relações de escala  $Y = Y_0 N^\beta$ , onde  $Y_0$  e  $\beta$  são constantes em  $N$  (população). As cidades de tamanhos diferentes não possuem as mesmas propriedades. Especificamente, observa-se geralmente que as taxas de quantidades sociais (tais como salários ou novas invenções) aumentam per capita com o tamanho da cidade (escala superlinear,  $\beta = 1 + \delta > 1$ , com  $\delta \simeq 0.15$ ), enquanto o volume ocupado pela infraestrutura urbana per capita (estradas, cabos, etc.) diminui (escala sublinear,  $\beta = 1 - \delta < 1$ ). Assim, esses dados resumem as expectativas familiares de que cidades

maiores além de ter um custo mais elevado e congestionamentos, são também mais criativas quando comparadas a cidades pequenas.

Os resultados empíricos sugeriram, também, que, apesar da aparente complexidade, as cidades podem ser simples: as propriedades globais médias podem ser determinadas por apenas alguns parâmetros-chave. No entanto, a origem dessas relações de escala observadas e uma explicação para as interdependências entre as facetas espaciais, infra-estruturais e sociais da cidade permaneceram desconhecidas.

Para uma compreensão teórica de como as cidades operam e como essas interdependências surgem, o autor desenvolve uma estrutura unificada e quantitativa. Considerando primeiro o modelo mais simples de uma cidade delimitada pela área  $A$  e população  $N$ , escreve-se as interações entre pessoas  $i, j$  em termos de uma rede social  $F_k^{ij}$ , e assume-se que as interações sociais (amizade, emprego, conhecimento, etc.) são locais, ocorrendo em uma área de interação  $a_0$ , e tem força  $g_k$ , onde  $k$  descreve tipos de links sociais. O parâmetro,  $g_k$ , pode ser positivo (atraente, expressando um benefício social, por exemplo, relações econômicas mutuamente benéficas) ou negativo (repulsivo, expressando um custo social, por exemplo, crime). Todos estes processos compartilham a mesma dinâmica subjacente média dos encontros sociais no espaço e tempo, no contexto da cidade e suas redes de infra-estrutura.

O número médio de interações locais por pessoa é dado pelo produto do alcançado por seu movimento,  $a_0 l$ , multiplicado pela densidade populacional  $n = N/A$ , onde  $l$  é o comprimento típico percorrido por pessoas, bens e informações. A produção social média total de uma cidade pode ser obtida multiplicando-se o número total de interações pelo resultado médio por interação,  $\bar{g}$ , levando a  $Y = G \frac{N^2}{A}$ , com parâmetro  $G \equiv \bar{g} a_0 l$  medindo o produto da média da área de tempos de saída social, ambos per capita (Figura 2.4.1). Cada saída socioeconômica,  $Y$ , tem unidades físicas definidas por  $g_k$ , porém é útil pensar em todas as quantidades expressas em termos de energia por unidade de tempo (potência).

Além disso, outra propriedade fundamental das cidades é que elas estão misturando populações, ou seja, mesmo que as pessoas na cidade exploram locais diferentes em momentos diferenciados, qualquer pessoa pode, em princípio, ser alcançada por outra pessoa. Este conceito, é a base das definições de cidades funcionais como áreas metropolitanas. Na prática, isso significa que o custo por pessoa de uma população mista é proporcional à dimensão transversal (diâmetro),  $L$ , da cidade  $L \sim A^{1/2}$ . Assim, a potência de gastos em processos de transporte para manter a cidade misturada é dada por,  $T = \varepsilon L N = \varepsilon A^{1/2} N$ , onde  $\varepsilon$  é uma força por unidade de tempo. Este custo deve ser coberto pelo orçamento de cada indivíduo,  $y = Y/N$ , requerendo  $y \simeq T/N$ , o que implica  $A(N) = a N^\alpha$  com  $\alpha = 2/3$  e  $\alpha = (G/\varepsilon)^\alpha$ . A área da linha de

base,  $a$ , aumenta com interações mais produtivas, por exemplo, devido ao crescimento econômico, os custos no transporte é diminuído, como pode-se observar nos padrões mundiais de expansão urbana ao longo do tempo. Assim, obtém-se  $Y = Y_0 N^\beta$ , em que  $\beta = 2 - \alpha = 1 + 1/3 > 1$  e  $Y_0 = G^{1-\alpha} \varepsilon^\alpha$ . Este modelo simples leva a área,  $A$ , variando sublinearmente com  $N$  ( $\alpha = 2/3 < 1$ ), e os resultados socioeconômicos,  $Y$ , variando superlinearmente ( $\beta = 4/3 > 1$ ). No entanto, isso superestima  $\beta$  porque, à medida que as cidades crescem, o espaço é ocupado e o transporte de pessoas, bens e informações é canalizado para as redes. O espaço criado por essas redes fornece a medida correta das interações sociais que podem ocorrer nas cidades.

No entanto, o autor propôs um modelo mais realista generalizando essas ideias em quatro premissas:

1) Juntar população. A cidade se desenvolve para que os cidadãos possam explorá-la com recursos à sua disposição. Formaliza-se este princípio como uma condição de entrada, exigindo que os recursos mínimos acessíveis a cada indivíduo,  $Y_{min}/N \sim GN/A$ , correspondam ao custo de chegar a qualquer ponto da cidade. Como as trajetórias de viagem não precisam ser lineares, generaliza-se a geometria através de uma dimensão fractal,  $H$ , de modo que a distância percorrida  $\propto A^{H/D}$ . Combinando densidade de interação aos custos, obtém-se uma relação de escala de área generalizada,  $A(N) = aN^\alpha$ , com  $a$  como antes e  $\alpha = \frac{2}{2+H}$  [ $\alpha = \frac{D}{D+H}$  em dimensões  $D$ ].  $H = 1$  permite que os indivíduos explorem totalmente a cidade dentro da menor distância percorrida, o que implica a escala  $N$  como um volume físico.

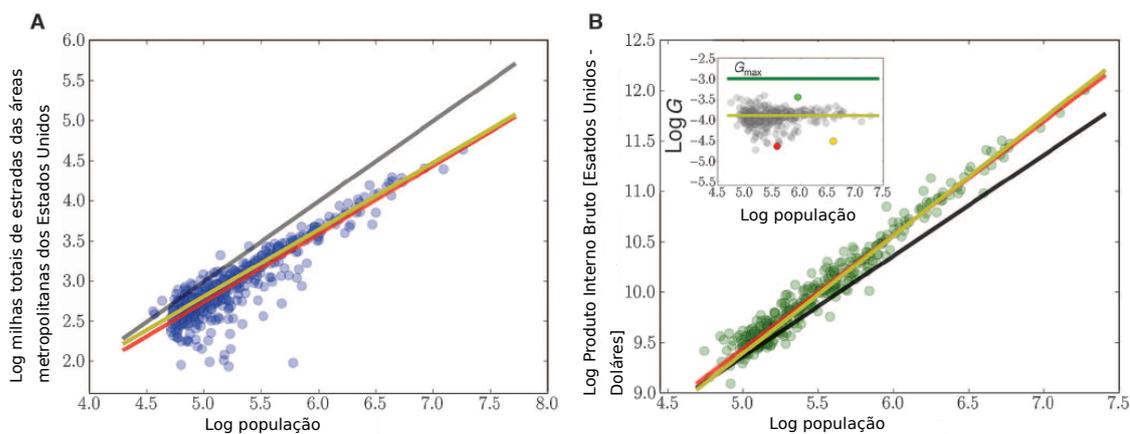
2) Crescimento incremental de rede. Esta suposição exige que as redes de infra-estrutura se desenvolvam gradualmente para conectar as pessoas à medida que elas se unem, levando a redes descentralizadas. Especificamente, a escala da Figura 13 2.4.1 (A) é obtida quando a distância média entre os indivíduos  $d = n^{-1/2} = (A/N)^{1/2}$  é igual ao comprimento médio da rede de infra-estrutura per capita, de modo que a área total da rede,  $A_n(N) \sim Nd = A^{1/2} N^{1/2}$ . Juntamente com primeira premissa, isso implica que  $A_n \sim a^{1/2} N^{1-\delta}$  com  $\delta = 1/6$  [ $A_n \sim A^{1/2} N^{(D-1)/D} = a^{1/D} N^{1-\delta}$ , com  $\delta = \frac{H}{D(D+H)}$  em  $D$  dimensões].

3) O esforço humano é limitado, o que requer que  $G$  seja, em média, independente de  $N$ , isto é,  $dG/dN = 0$  (Figura 2.4.1 (B), inserção). As cidades em crescimento exigem de seus habitantes um esforço maior tanto mental como físico e, isso tem sido uma preocupação generalizada para os cientistas sociais. Assim, essa premissa é necessária para levantar uma importante objeção a qualquer conceituação de cidades como sistemas invariantes de escala.

4) Os resultados socioeconômicos são proporcionais às interações sociais locais, de modo que  $Y = GN^2/A_n \sim N^{1+\delta}$ . Nessa perspectiva, as cidades não são

apenas concentrações de pessoas mas de interações sociais. Este ponto foi enfatizado por Jacobs (1969) mas, tem sido difícil quantificar. A previsão de que a escala de interações sociais com  $\beta = 1 + \delta \simeq 7/6$ , foi observada recentemente em redes de telecomunicações urbanas por Schläpfer (2013). Juntos, esses pressupostos preveem expoentes de escala para uma ampla variedade de indicadores urbanos, desde de padrões de comportamento humano e propriedades da infraestrutura até o valor da terra.

A Figura 2.4.1 (A) mostra milhas de pista total (volume) de estradas nas áreas metropolitanas dos Estados Unidos em 2006. Os dados para as 415 áreas urbanas foram obtidos do Escritório de Informações sobre Políticas Rodoviárias da Administração Federal de Rodovias. O melhor ajuste para uma relação de escala  $Y(N) = Y_0 N^\beta$ , obteve  $\beta = 0,849 \pm 0,038$  (intervalo de confiança de 95%,  $R^2 = 0,65$ ); a previsão teórica,  $\beta = 5/6$  e a escala linear  $\beta = 1$ . Na Figura 2.4.1 (B) é mostrado o produto metropolitano das áreas metropolitanas dos Estados Unidos em 2006. Os dados para as 363 áreas metropolitanas foram obtidos pelo Departamento de Análise Econômica. O melhor ajuste aos dados obteve,  $\beta = 1,126 \pm 0,023$  (intervalo de confiança de 95%,  $R^2 = 0,96$ ); a previsão teórica,  $\beta = 7/6$  e escala proporcional,  $\beta = 1$ .



**Figura 2.4.1 – Escala da infra-estrutura urbana e produção socioeconômica -** (A) Os pontos azuis representam milhas de pista total (volume) de estradas nas áreas metropolitanas dos Estados Unidos. As linhas mostram o melhor ajuste para uma relação de escala (vermelho); a previsão teórica (amarelo); e a escala linear (preto). (B) Os pontos verdes representam o produto metropolitano das áreas metropolitanas dos Estados Unidos. As linhas descrevem o melhor ajuste (vermelho) aos dados; a previsão teórica (amarelo); e escala proporcional, (preto). Adaptada de Bettencourt (2013).

Os dois parâmetros de melhor ajuste em cada relação de escala foram estimados por meio da minimização de mínimos quadrados ordinários para a relação linear entre as variáveis transformadas logaritmicamente. A inserção mostra a estimativa de  $G$  para as 313 áreas metropolitanas dos Estados Unidos e a lei de conservação

$\frac{d \ln G}{d \ln N} = 0$  ( $R^2 = 0,003$ ).  $G$  é medido como o produto do produto interno bruto e do volume da estrada, ambos per capita. Como previsto, os valores observados de  $G$  para diferentes cidades se agrupam em torno de seu valor mais provável, que fornece uma ótima estimativa para  $G^*$  e são limitados pelo  $G_{max} \simeq 8G^*$ .

## 2.5 Conexão entre distribuição Lei de Potência e Alometrias para indicadores urbanos

Essa seção apresenta a existência de conexões entre distribuições leis de potência e alometrias encontradas nas investigações de Alves et al. (2014). Sistemas urbanos possuem várias propriedades, entre as quais duas são notáveis e onipresentes: o surgimento de distribuições assintóticas da lei de potência e de alometrias para uma ampla gama de variáveis que de alguma forma caracterizam os sistemas urbanos (indicadores urbanos ou métricas). Por um lado a população [MARSILI et al. (1998); BATTY (2008)], tamanhos de edifícios [BATTY (2008)], fortunas pessoais [NEWMAN (2005)], tamanhos de empresa [AXTELL (2001)], e outros indicadores foram encontrados para seguir sinteticamente as distribuições da lei de potência. Por outro lado, as alometrias com o tamanho da população foram relatadas para o crime [BETTENCOURT et al. (2010); ALVES et al. (2013)], suicídio [MELO et al. (2014)], várias métricas urbanas, incluindo patentes, postos de gasolina, produto interno bruto [BETTENCOURT et al. (2007); ARBESMAN (2009); BETTENCOURT (2010)], entre outros. Estas alometrias foram modeladas por Bettencourt (2013) através de um pequeno conjunto de princípios simples e baseados localmente, como visto na seção 2.4. No entanto, a conexão entre essas duas características foi elucidada por Gomez-Lievano et al. (2012). Neste trabalho descrito na seção 2.2, os autores mostraram, para a relação entre homicídios e população, que quando ambos os indicadores urbanos são assintoticamente distribuídos como leis de potências, também podem exibir uma relação alométrica específica. Particularmente, o expoente da alometria pode ser totalmente determinado a partir dos expoentes das distribuições da lei de potência. Diante disto, Alves et al. (2014) estenderam empiricamente os resultados através de uma ampla caracterização de todas as possíveis relações entre pares a partir de 12 indicadores urbanos de cidades brasileiras.

Para essa análise foram usados dados de cidades brasileiras, disponíveis no sítio do Sistema Único de Saúde - DATASUS no ano de 2000, selecionando 2862 cidades (cerca de 51% das cidades) para as quais todos os valores dos indicadores urbanos estavam disponíveis. A base de dados é composta por doze indicadores urbanos  $Y_i$  ao nível da cidade. São eles: população total ( $i = 0$ ), número de casos de trabalho infantil ( $i = 1$ ), população com mais de 60 anos ( $i = 2$ ), população feminina

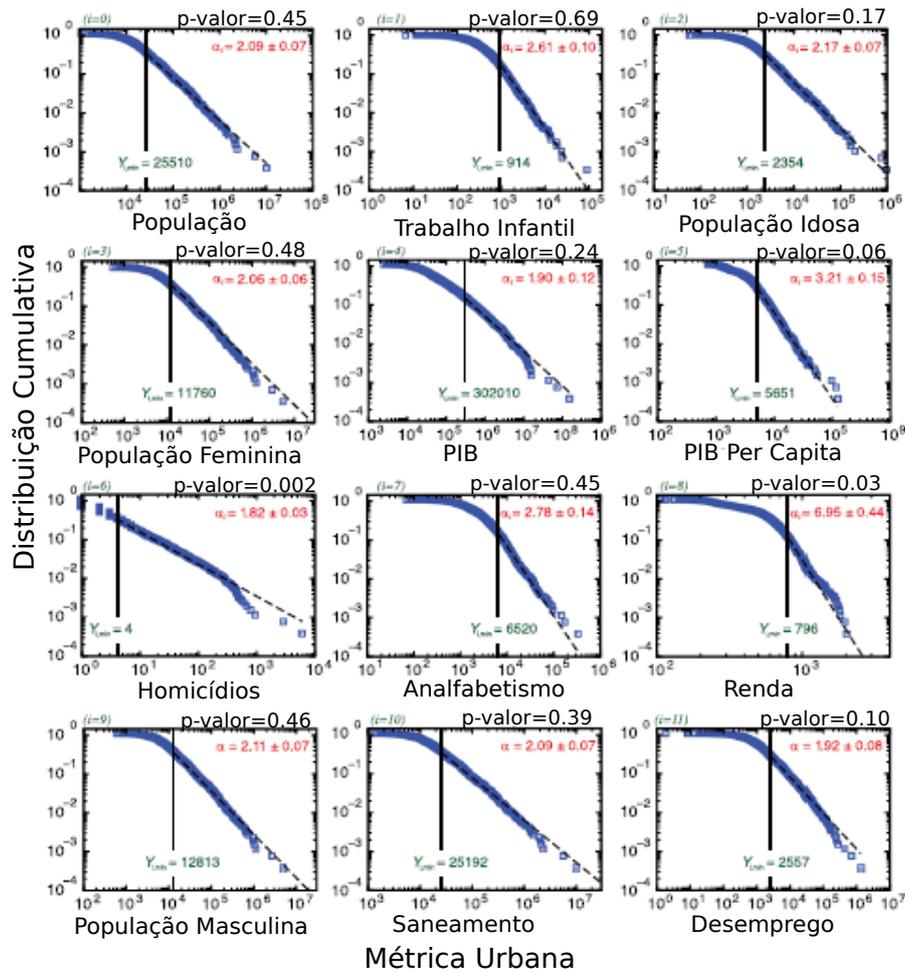
( $i = 3$ ), produto interno bruto - PIB ( $i = 4$ ), PIB per capita ( $i = 5$ ), número de homicídios ( $i = 6$ ), número de analfabetos com mais de 15 anos ( $i = 7$ ), renda familiar média ( $i = 8$ ), população masculina ( $i = 9$ ), número de instalações sanitárias ( $i = 10$ ) e número de desempregados com mais de 16 anos ( $i = 11$ ). Considerando que cidades são as menores unidades administrativas com um governo local.

Iniciou-se as investigações por construir os gráficos das distribuições acumuladas para os 12 indicadores urbanos, como por ser visto na Figura 2.5.1. Pode-se observar um comportamento aproximadamente linear da dispersão dos dados em log-log o que permite supor que as formas das distribuições podem ser aproximadas por leis de potências assintótica [ $P_i(Y_i) \propto Y_i^{-\alpha_i}$ ], das distribuições para grandes valores de indicadores urbanos. Com o objetivo de tornar esse resultado mais robusto, empregou-se o procedimento estatístico proposto por Clauset et al. (2009), que é uma maneira sistemática de testar estatisticamente as distribuições empíricas lei de potência. Devido à natureza discreta da maioria desses indicadores urbanos, considerou-se a versão discreta do procedimento, aplicando-se o ajuste de máxima verossimilhança

$$P_i(Y_i) = \frac{Y_i^{-\alpha_i}}{\zeta(\alpha_i, Y_{i,min})}, \quad (2.54)$$

onde  $\zeta(\alpha_i, Y_{i,min}) = \sum_{n=0}^{\infty} (Y_{i,min} + n)^{-\alpha_i}$  é a função zeta de Hurwitz, para normalizando  $P_i(Y_i)$  quando  $Y_i$  é discreto.  $Y_{i,min}$  é um parâmetro que representa o início do regime de lei de potência, e  $\alpha_i$  é o expoente da lei de potência. Depois de encontrar os parâmetros de melhor ajuste, avaliou-se a adequação do ajuste por meio do teste de Crámer-von Mises.

Os valores de  $\alpha_i$  e  $Y_{i,min}$ , bem como o p-valor do teste de Crámer-von Mises, também são mostrados na Figura 2.5.1. Não se pode rejeitar a hipótese da lei de potência em um nível de confiança de 99% para quase todos indicadores, exceto o número de homicídios, para os quais p-valor é 0,002.



**Figura 2.5.1 – Distribuições cumulativas de 12 indicadores urbanos das cidades brasileiras no ano 2000.** Adaptada de Alves et al. (2014).

As distribuições exibem decaimentos da lei de potência assintótica que são bem descritas pela função lei de potência  $P_i(Y_i) \sim Y_i^{-\alpha_i}$  para  $Y_i \geq Y_{i,min}$ .

Os autores também verificaram se as distribuições da lei de potência para indicadores urbanos implicam em relações alométricas entre pares de indicadores. Os seguintes desenvolvimentos analíticos são apenas uma revisão dos resultados presentes no artigo de Gomez-Lievano et al. (2012). Eles propuseram que, se a forma da alometria estiver bem estabelecida, pode-se escrever

$$P_i(Y_i) = \sum_{Y_j=Y_j^*}^{\infty} P_{i,j}(Y_i|Y_j)P_j(Y_j), \quad (2.55)$$

onde  $P_i(Y_i)$  é a distribuição de probabilidade do indicador urbano  $Y_i$  e  $P_{i,j}(Y_i|Y_j)$  é a probabilidade condicional representando as flutuações em torno da alometria entre os indicadores  $Y_i$  e  $Y_j$ . Já se sabe que  $P_i(Y_i)$  tem a forma da lei de potência da Equação 2.54 e após os resultados empíricos, assume-se que a probabilidade

condicional é uma distribuição lognormal

$$P_{i,j}(Y_i|Y_j) = \frac{1}{Y_i \sqrt{2\pi\sigma_{i,j}^2(Y_j)}} \exp \left\{ -\frac{[\ln Y_i - \mu_{i,j}(Y_j)]^2}{2\sigma_{i,j}^2(Y_j)} \right\}, \quad (2.56)$$

em que  $\mu_{i,j}(Y_j)$  e  $\sigma_{i,j}^2(Y_j)$  representam possíveis dependências dos parâmetros lognormal no indicador  $Y_j$ . Em particular, porque também está se assumindo que existe uma alometria entre  $Y_i$  e  $Y_j$  (isto é,  $Y_i \propto Y_j^{\beta_{i,j}}$ ), a forma funcional de  $\mu_{i,j}(Y_j)$  é

$$\mu_{i,j}(Y_j) = \ln[A_{i,j}Y_j^{\beta_{i,j}}], \quad (2.57)$$

se considerar o desvio padrão não dependente de  $Y_j$  (isto é,  $\sigma_{i,j}^2(Y_j) = \sigma_{i,j}^2$ ).

Com essas considerações e substituindo o somatório na Equação 2.55 por uma integral e considerando  $Y_j^*$  suficientemente pequeno. Após alguns cálculos obtém-se

$$P_i(Y_i) \propto Y_i^{-\frac{(\alpha_j - 1)}{\beta_{i,j}} - 1}. \quad (2.58)$$

Lembrando que  $P_i(Y_i) \sim Y_i^{-\alpha_i}$ , obtém a relação entre o expoente alométrico ( $B_{i,j}$ ) e os expoentes da lei de potência ( $\alpha_i$  e  $\alpha_j$ ) como

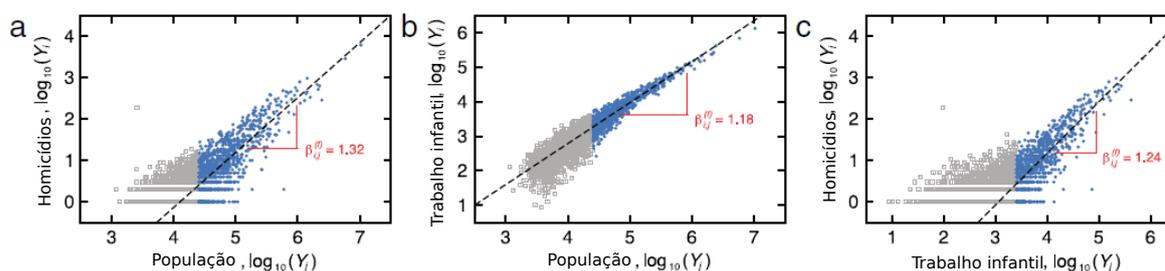
$$\beta_{i,j} = \frac{\alpha_j - 1}{\alpha_i - 1}. \quad (2.59)$$

Nas investigações de Gomez-Lievano et al. (2012) indicadores urbanos apresentam distribuições lei de potência e podem exibir relações alométricas com expoentes específicos. Para testar empiricamente este resultado, investigou-se todas as possíveis alometrias entre pares de indicadores urbanos da base de dados, considerando o logaritmo dos indicadores urbanos e ajustando (via método dos mínimos quadrados ordinários) a função linear

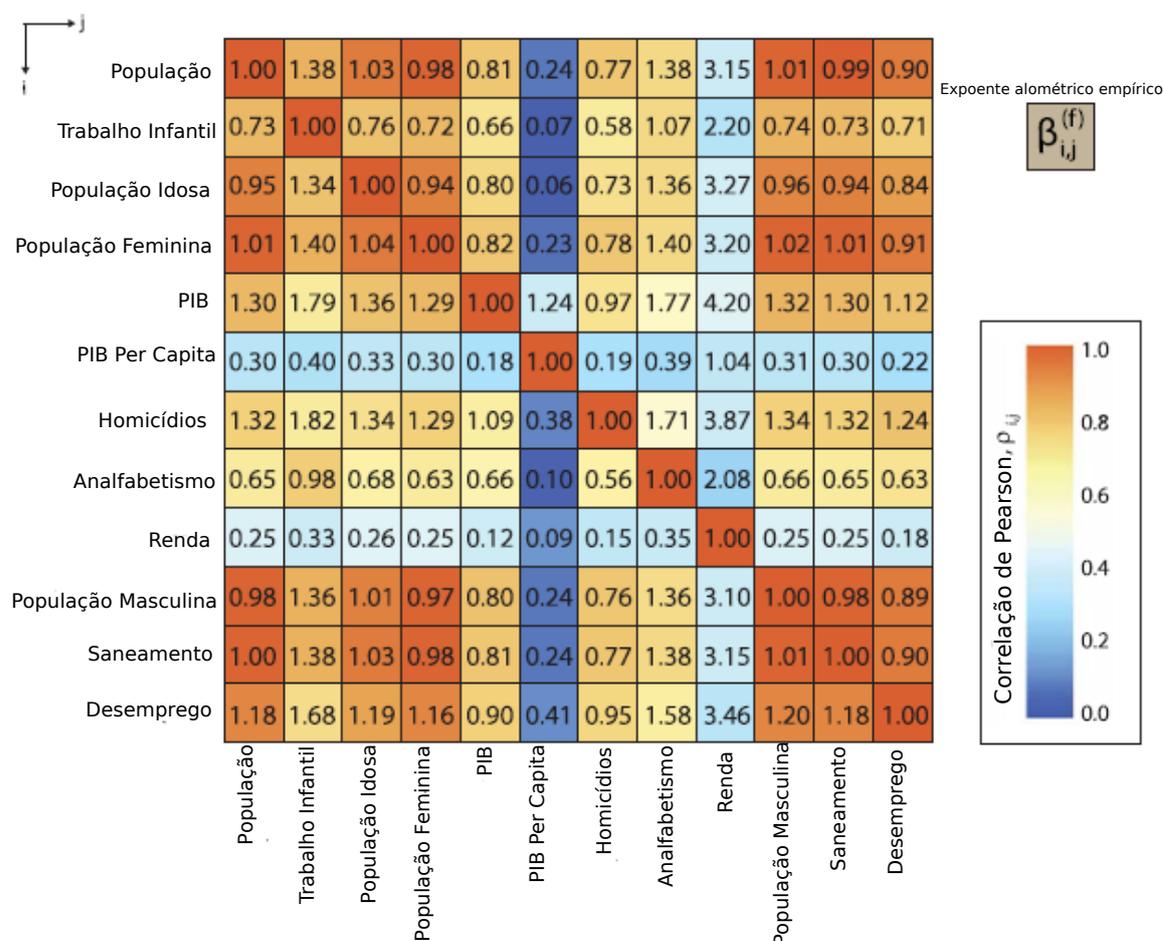
$$\log_{10}Y_i = A_{i,j} + \beta_{i,j}^{(f)} \log_{10}Y_j \quad (2.60)$$

para todos os indicadores. Aqui,  $A_{i,j} = \log_{10}A_{i,j}$  é uma constante empírica e  $\beta_{i,j}^{(f)}$  é o valor empírico do expoente alométrico (obtido após o ajuste da alometria). O procedimento é mostrado na Figura 2.5.2 para homicídio versus população, trabalho infantil versus população e homicídio versus trabalho infantil. Para encontrar os valores de  $\beta_{i,j}^{(f)}$ , aplicou-se um ponto de corte no eixo das abscissas considerando apenas os valores de  $Y_j$  que são maiores que  $Y_{j,min}$  (o início da lei de potência na distribuição de  $Y_j$ ). Os pontos azuis na Figura 2.5.2 são os pontos que obedecem a essa condição e as linhas tracejadas representam as funções lineares ajustadas aos dados. Repete-se este procedimento para todas as possíveis alometrias e a Figura 2.5.3 mostra um gráfico da matriz com todos os valores de  $\beta_{i,j}^{(f)}$ . Caracterizou-se, ainda, a qualidade dessas

alometrias calculando o coeficiente de correlação de Pearson  $\rho_{i,j}$ , para as relações alométricas linearizadas (isto é,  $\log_{10}Y_i$  versus  $\log_{10}Y_j$ ), que é mostrado na Figura 2.5.3 através do código de cores. Observe que a maioria das relações apresentam valores  $\rho_{i,j}$  maiores que 0.5 ( $\approx 70\%$ ). No entanto, também observa-se duas exceções sistemáticas: alometrias com indicadores urbanos PIB per capita e renda. O motivo desse comportamento desviante está relacionado ao fato de que esses dois indicadores são definidos per capita e por família. Este resultado também sugere que a principal fonte de correlação entre os dois indicadores  $Y_i$  e  $Y_j$  ( $i \neq j \neq 0$ ) vem de suas relações com o tamanho da população  $Y_0$ .



**Figura 2.5.2 – Exemplos de alometrias entre indicadores urbanos:** (a) homicídios versus população, (b) trabalho infantil versus população e (c) homicídios versus trabalho infantil. Os marcadores são logaritmos de base 10 dos valores dos indicadores urbanos  $Y_i$  versus  $Y_j$  para cada cidade. Os pontos azuis representam aquelas cidades que possuem  $Y_j \geq Y_{j,min}$  e quadrados cinza são aquelas para as quais  $Y_j < Y_{j,min}$ . As linhas tracejadas são ajustes lineares ordinários para os pontos que obedecem à condição  $Y_j \geq Y_{j,min}$  e a inclinação de cada reta é igual ao expoente alométrico  $\beta_{i,j}^{(f)}$ . Adaptada de Alves et al. (2014).



**Figura 2.5.3 – Valores empíricos dos expoentes alométricos  $\beta_{i,j}^{(f)}$  obtidos de todos os possíveis pares de relações (em escala logarítmica) entre os 12 indicadores.** O código de cores mostra o valor do coeficiente de correlação de Pearson  $\rho_{i,j}$  para cada alometria. Adaptada de Alves et al. (2014).

Para comparar os parâmetros  $\beta_{i,j}^{(f)}$  empíricos e os parâmetros  $\beta_{i,j}$  do modelo analítico, a Figura 2.5.4 (a) mostra um gráfico de dispersão de  $\beta_{i,j}^{(f)}$  versus  $\beta_{i,j}$  para alometrias com correlação de Pearson  $\rho_{i,j}$  maior que 0,5. Note que, uma função linear (sem constante aditiva) descreve muito bem a relação entre  $\beta_{i,j}^{(f)}$  e  $\beta_{i,j}$ , especialmente se tomar os erros padrão dos valores  $\beta_{i,j}^{(f)}$  (área sombreada) em conta. Também, note que a qualidade dessa relação linear se deteriora quando começa-se a considerar as alometrias com valores menores de  $\rho_{i,j}$ . Assim, além de serem distribuídos como leis de potência, os indicadores  $Y_i$  e  $Y_j$  também devem apresentar uma alometria de boa qualidade (grande valor de  $\rho_{i,j}$ ) para que a relação da Equação 2.59 seja válida.

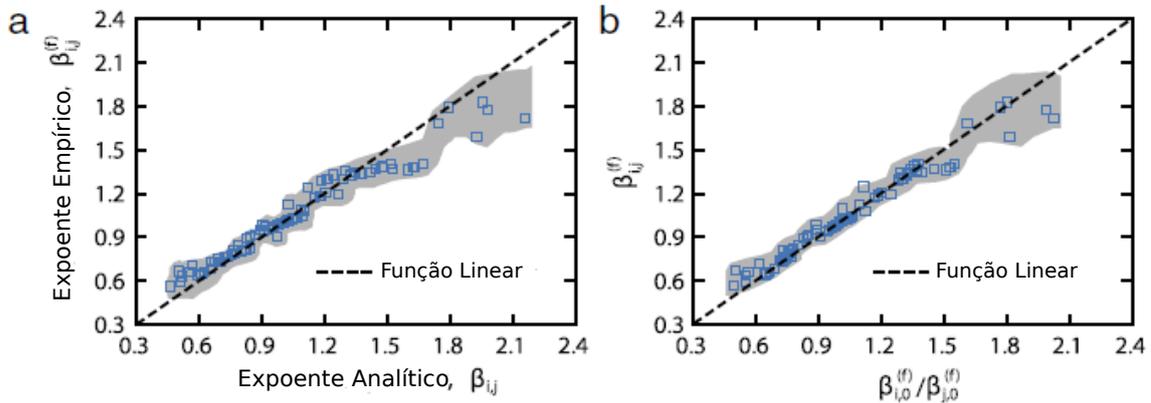
Vale a pena notar que, embora tenha-se considerado todos os expoentes da lei de potência  $\beta_{i,j}^{(f)}$  independentes, eles são realmente dependentes uns dos outros. Em primeiro lugar, o expoente alométrico da relação entre  $Y_i$  e  $Y_j$  é supostamente o inverso daquele obtido para a relação entre  $Y_j$  e  $Y_i$ . Para os valores empíricos, observa-se pequenos desvios nesta relação devido aos diferentes pontos de corte aplicados ao

eixo das abscissas. No entanto, ao considerar os intervalos de confiança para os valores de  $\beta_{i,j}^{(f)}$ , verifica-se que essa relação inversa entre os valores de expoentes alométricos se mantém na maioria dos casos. Em segundo lugar e, mais importante, quando se tem  $Y_i \sim Y_k^{\beta_{i,k}^{(f)}}$  e  $Y_j \sim Y_k^{\beta_{j,k}^{(f)}}$  segue diretamente que  $Y_i \sim Y_j^{\beta_{i,j}^{(f)}}$  com  $\beta_{i,j}^{(f)} = \beta_{i,k}^{(f)} / \beta_{j,k}^{(f)}$ . Este resultado implica que as alometrias entre todos os pares de indicadores  $Y_i (i \neq 0)$  podem ser entendidas como uma consequência da relação alométrica entre o indicador  $Y_i$  e o tamanho da população  $Y_0$ . No entanto, relações alométricas ruidosas e o fato de  $P_i(Y_i)$  não ser uma lei de potência perfeita podem ter um papel não trivial nos valores empíricos de  $\beta_{i,j}$ . Na Figura 2.5.4 (b), testa-se empiricamente a relação  $\beta_{i,j}^{(f)}$  e  $\beta_{i,0}^{(f)} / \beta_{j,0}^{(f)}$  se deteriora quando começa-se a considerar as alometrias com valores menores de  $\rho_{i,j}$ .

Apesar de, as alometrias entre  $Y_i$  e  $Y_j (i, j \neq 0)$  serem uma consequência de suas relações alométricas com o tamanho da população  $Y_0$ , não significa necessariamente que o indicador  $Y_j$  não tem potencial explicativo para descrever  $Y_i$ . Na tentativa de testar essa possibilidade, ajusta-se o seguinte modelo linear generalizado:

$$\log_{10} Y_i = a_{i,j} + b_{i,0} \log_{10} Y_0 + b_{i,j} (\log_{10} Y_j) / (\log_{10} Y_0)^{\beta_{i,j}^{(f)}}, \quad (2.61)$$

considerando todos os possíveis pares de relações entre  $\log_{10} Y_i$  e  $\log_{10} Y_j$ , mas excluindo os indicadores população idosa, população feminina, população masculina e saneamento, uma vez que formam uma relação linear quase perfeita com a população. Assim, todo  $b_{i,j}$  estatisticamente diferente de zero indica que o  $\log_{10} Y_j$  apresenta um potencial explicativo para descrever  $\log_{10} Y_i$ , o qual não está relacionado à sua própria relação com  $\log_{10} Y_0$ .



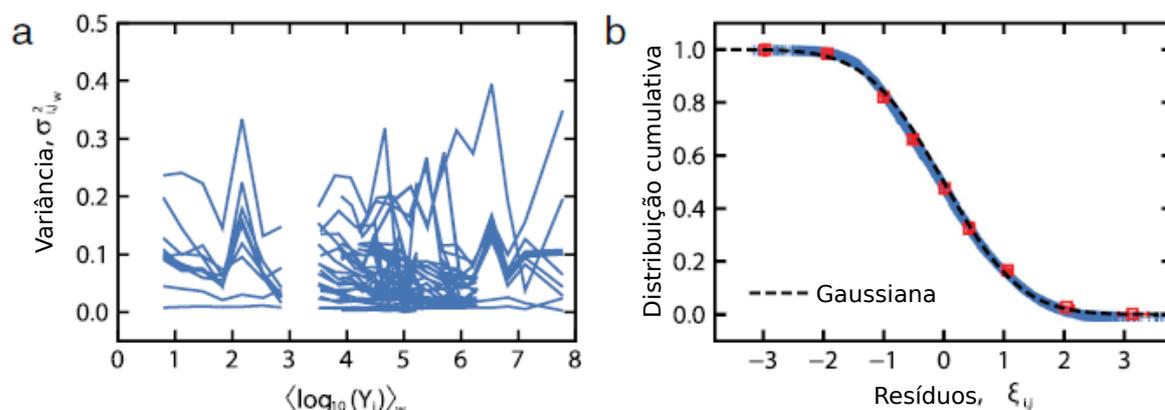
**Figura 2.5.4 – Comparação dos parâmetros empíricos e os parâmetros do modelo analítico.** (a) Comparação entre os valores empíricos dos expoentes alométricos  $\beta_{i,j}^{(f)}$  (obtidos através de ajustes lineares para dados empíricos) e os analíticos  $\beta_{i,j}$  (obtidos da Equação 2.59). Os quadrados são os valores de  $\beta_{i,j}^{(f)}$  versus  $\beta_{i,j}$  onde o coeficiente de correlação de Pearson  $\rho_{i,j}$  que caracteriza a alometria é maior que 0,5. (b) Verificação empírica da relação linear entre  $\beta_{i,j}^{(f)}$  e  $\beta_{i,0}^{(f)}/\beta_{j,0}^{(f)}$  (quadrados) quando se considera  $\rho_{i,j} \geq 0.5$ . As áreas sombreadas representam os erros padrão de  $\beta_{i,j}^{(f)}$  após suavizar os dados com um filtro médio móvel de tamanho de janela 5 e as linhas tracejadas são funções lineares sem constante aditiva, em ambos os gráficos. Adaptada de Alves et al. (2014).

Abordou-se, ainda, a questão das flutuações em torno das alometrias com o objetivo de verificar as duas hipóteses remanescentes nos cálculos de Gomez-Lievano et al. (2012): o comportamento constante do parâmetro lognormal ( $\sigma_i^2 Y_i$ ) e a distribuição lognormal das flutuações. Para verificar essas hipóteses usou-se um procedimento de agrupar as variáveis  $\log_{10} Y_i$  e  $\log_{10} Y_j$  com  $Y_j > Y_{j,min}$  em  $\omega$  janelas igualmente espaçadas, e calculou-se a variância  $\sigma_{i,j\omega}^2$ , em cada janela. Observe que nessa escala log-log o valor da variância  $\sigma_{i,j\omega}^2$  é aproximadamente igual ao valor do parâmetro lognormal  $\sigma_{i,j}^2(Y_j)$ ; portanto, verificando que  $\sigma_{i,j\omega}^2$  não é dependente de  $\log_{10} Y_j$  o que equivale a mostrar que  $\sigma_{i,j}^2(Y_j)$  é constante. A Figura 2.5.5 (a) mostra o comportamento de  $\sigma_{i,j\omega}^2$  no valor médio da janela de  $\log_{10} Y_j$ . Observe que não há dependência clara nessas relações e que o comportamento pode ser aproximado por uma função constante, onde o valor do platô muda de indicador para indicador. Por fim, considerou-se as flutuações normalizadas em torno das relações alométricas, isto é,

$$\xi_{i,j} = \frac{\log_{10} Y_i - \langle \log_{10} Y_i \rangle_\omega}{\sigma_{i,j\omega}}, \quad (2.62)$$

onde  $\langle \log_{10} Y_i \rangle_\omega$  é o valor médio na janela  $\omega$  de  $\log_{10} Y_i$ . A Figura 2.5.5 (b) mostra a distribuição acumulada de  $\xi_{i,j}$  para todos os indicadores urbanos, a qual é bem ajustada pela distribuição acumulada da Gaussiana. Consequentemente  $\xi_{i,j}$  é normalmente distribuído,  $P_{i,j}(Y_i|Y_j)$  (que representa as flutuações na escala usual) deve seguir uma

distribuição normal lognormal. Analogamente aos expoentes alométricos, nem todas estas distribuições são independentes porque conhecendo  $P_{i,0}(Y_i|Y_0)$  e  $P_{j,0}(Y_j|Y_0)$  também pode-se calcular  $P_{i,j}(Y_i|Y_j)$ .



**Figura 2.5.5 – Variâncias e distribuições cumulativas.** (a) Variâncias  $\sigma_{i,j\omega}^2$  das flutuações que envolvem as alometrias (na escala logarítmica de base 10) para cada indicador urbano. (b) Distribuições cumulativas das flutuações normalizadas  $\xi_{i,j}$  em torno das alometrias (cruzamentos). Os quadrados são os valores médios de todas as distribuições, as barras de erro são intervalos de confiança de 95% e a linha tracejada é uma distribuição Gaussiana padrão (média zero e variância um). Em ambas as parcelas, considerou-se apenas as relações alométricas onde a correlação de Pearson  $\rho_{i,j}$  é maior que 0,5. Adaptada de Alves et al. (2014).

### 3 LEIS DE POTÊNCIA EM ELEIÇÕES

Em uma sociedade democrática um dos processos mais fundamentais são as eleições, além de representar um ato de cidadania. O desenvolvimento futuro de uma nação democrática depende muito dos resultados das eleições pois, são os indivíduos que escolhem representantes e governantes que farão e executarão leis que interferem diretamente na vida dos cidadãos. Vários esforços têm sido feitos para entender o modo como as pessoas fazem suas escolhas e quais aspectos sociais, econômicos e políticos são relevantes para influenciar, ou mesmo determinar, o resultado geral de um processo eleitoral particular. Além disso, os processos eleitorais podem ser considerados como um sistema complexo, pois ao tentar encontrar uma regularidade é necessário analisar vários resultados de eleições. E ainda, encontra-se disponível ao público uma enorme quantidade de dados de eleições, em formato eletrônico. Diante desse fato, este capítulo apresenta um histórico de resultados sobre os dados de eleições.

#### 3.1 Distribuição de Votos

No contexto da política, várias pesquisas têm sido realizadas utilizando-se ferramentas de sistemas complexos. Esses estudos têm mostrado que em muitos processos eleitorais as dinâmicas dos sistemas sociais vistos como sistemas complexos se comportam como leis de potência. Dois dos trabalhos que analisaram processos eleitorais obtendo comportamentos leis de potência para os dados, foram Costa Filho et al. (1999; 2002) investigando os resultados das eleições brasileiras nos anos de 1998 e 2002, indicando que a distribuição de votos entre os candidatos segue uma lei de potência nas eleições proporcionais (deputados estaduais e federais).

Eleições proporcionais são usadas, no Brasil, para eleger vereadores, deputados estaduais e federais. Nesse sistema é previsto o uso de um sistema de lista aberta, na qual são reunidos os votos gerais dos candidatos de cada partido. Este ranking origina às listas partidárias que são compostas pelos candidatos mais votados de determinado partido naquela eleição. Nesses sistemas, cada partido obtém um número de vagas proporcionais à soma dos votos em todos os seus candidatos, e estas vagas são distribuídas, pela ordem, aos candidatos mais votados dentro do partido (GOVERNO DO BRASIL).

Para esta análise, os autores consideraram os resultados das eleições em cada estado para a posição de deputado estadual e federal e, normalizaram os votos

de cada candidato pelo número total de eleitores do respectivo estado. O histograma do número de candidatos  $N$  que receberam uma fração de votos  $v$  é mostrado nas Figuras 3.1.1 (a) e 3.1.1 (b). Assim, há um comportamento hiperbólico, ou seja,  $N(v)$  segue bem próximo a uma lei de potência

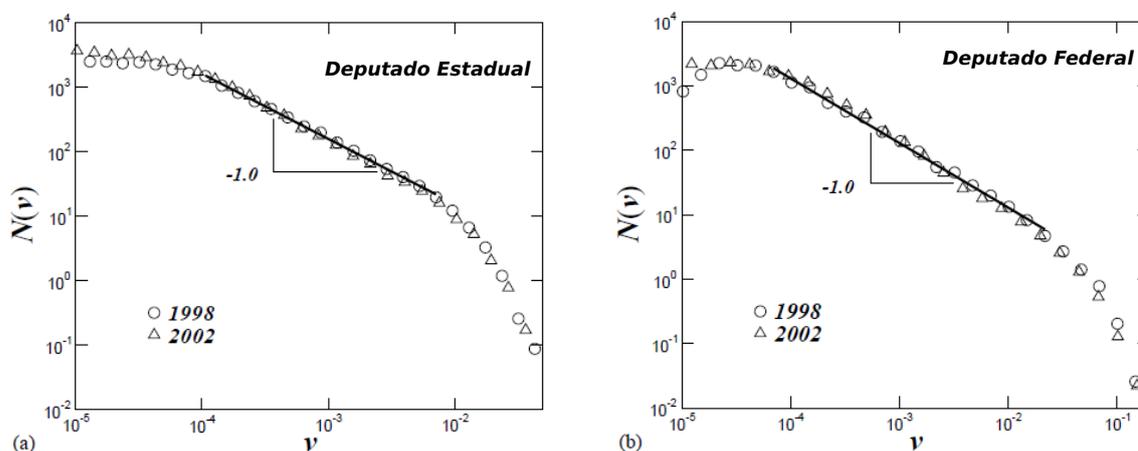
$$N(v) \propto v^{-\alpha}, \quad (3.1)$$

com  $\alpha \approx 1$ , por duas ordens de magnitude.

As leis de potência encontradas podem ser interpretadas como um “grande processo” multiplicativo. Nesse caminho, associa-se a cada candidato fatores  $p'_i$ s, no qual, cada  $p'_i$  está intrinsecamente associado aos atributos do candidato que realiza o subprocesso  $i$  entre os eleitores e, ainda mais, assume-se que os  $p'_i$ s são positivos e independentes, de maneira que a sua fração de voto deve ser

$$v \propto p_1 p_2 \dots p_n \quad (3.2)$$

em que  $n$  é a quantidade de subprocessos. Para  $n$  suficientemente grande e pelo teorema do limite central, a distribuição de  $v$  é aproximadamente lognormal. Como a lognormal é proporcional a  $\frac{1}{v} \exp(-a(\ln v)^2 + b \ln v)$  e caso seu expoente, em módulo, for muito pequeno, a lognormal se aproxima de  $v^{-1}$ , que é o padrão encontrado na Figura 3.1.1.

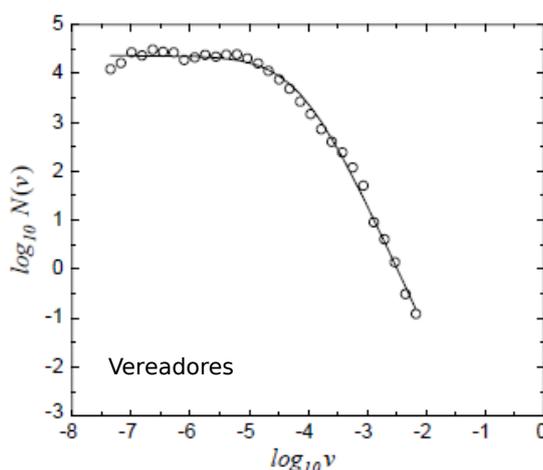


**Figura 3.1.1 – Histogramas das frações de votos** - Os gráficos em log-log mostram a distribuição de votos para as eleições de 1998 (círculos) e de 2002 (triângulos). As linhas sólidas representam os ajustes para as regiões de escala (a) para deputados estaduais e (b) para deputados federais. Adaptada de Costa Filho et al. (2002).

Porém, analisando os histogramas da Figura 3.1.1, nota-se que as curvas, aparentemente, são bem comportadas. No entanto, uma parte relevante da distribuição dos dados não é ajustado pela lei de potência sugerida, surgindo assim a seguinte questão: Um outro modelo ajustaria melhor esses dados? Com base nos resultados das

eleições proporcionais do Brasil nos anos de 1998 e 2000, essa questão foi respondida para as posições de vereadores, deputados estaduais e federais. Lyra et al. (2003) mostraram que as distribuições de votos podem ser bem descritas por uma lei de Zipf generalizada, introduzida em uma ferramenta estatística não-extensiva.

Os dados das eleições brasileiras, realizadas no ano de 2000, e os resultados das 15 maiores cidades para a posição de vereador, foram agrupados em 15 histogramas por meio das médias e mostrados em log-log na Figura 3.1.2.



**Figura 3.1.2 – Ajuste via Lei de Zipf - Vereadores** - Gráfico em log-log mostra distribuição de votos para vereadores. Os círculos são referentes aos dados dos resultados de eleições proporcionais brasileiras de 2002, considerando as 15 maiores capitais. A linha sólida é o melhor ajuste via lei de Zipf, equação (11), cujo o expoente é  $\alpha = 2,63$ . Adaptada de Lyra et al. (2003).

A curva resultante segue bem próximo da lei de Zipf

$$N(v) = \frac{A}{(1 + Cv)^\alpha}, \quad (3.3)$$

onde  $A$  é uma constante de normalização e  $C$  regula a transição entre o platô inicial e o regime de lei de potência caracterizada pelo expoente  $\alpha$ . Em particular, a exponencial  $q$ ,  $exp_q(t)$  é solução da equação

$$\frac{dW(t)}{dt} = \lambda_q W^q \quad (3.4)$$

onde  $\lambda_q = 1$  e  $W(0) = 1$ , ou seja,  $W(t) = exp_q(t) = [1 + (1 - q)t]^{1/(1-q)}$ . Quando  $q \rightarrow 1$  corresponde à exponencial usual. Considerando  $W = N$  e  $t = v$  na Equação 3.4, obtém-se a solução

$$N(v) = \frac{N(0)}{[1 + (q - 1)\lambda'_q v]^{1/(q-1)}}, \quad (3.5)$$

com  $\lambda'_q = -\lambda_q N(0)^{q-1}$ . O melhor ajuste de  $N(v)$  para as eleições de vereadores para a Equação 3.5 é mostrado como uma linha sólida contínua na Figura 3.1.2 e, quando

comparada a Equação 3.3, fornece,  $\alpha = 1/(q - 1) = 2.63$  ( $q = 1.68$ ), um expoente consideravelmente maior que um.

Na Figura 3.1.3 é mostrado  $N(v)$  para dados de eleições para deputados estaduais (a) e federais (b) para as 15 maiores capitais do Brasil, assim como na análise para vereadores. Nota-se que depois do platô inicial, um regime de escala lei de potência emerge com expoente bem próximo de  $-1.0$  em ambos os casos. Porém, observa-se que o regime de escala muda para grandes frações de votos e ocorre um rápido decaimento. Esta quebra de invariância de escala não é vista nos dados das eleições para vereadores, como pode-se notar na Figura 3.1.2.

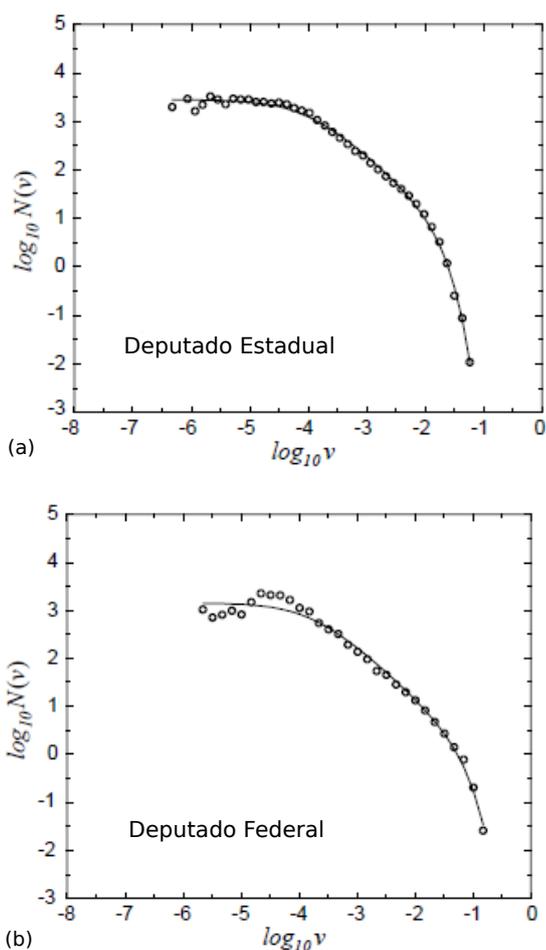
Portanto, é notório, que os gráficos da Figura 3.1.2 podem ser composições entre a exponencial usual ( $q = 1$ ) e a exponencial  $q$  ( $q \neq 1$ ). Logo, é admissível supor que  $N(v)$  pode ser a solução de uma equação que é a interpolação entre a Equação 3.4 com ( $q = 1$ ) e a Equação 3.4 com ( $q \neq 1$ ), ou seja, considera-se que  $N(v)$  satisfaz a equação

$$\frac{dN(v)}{dv} = -\lambda - (\lambda_q - \lambda)N^q, \quad (3.6)$$

cuja a solução é

$$N(v) = \frac{N(0)}{[1 - \lambda'_q/\lambda + (\lambda'/\lambda)e^{(q-1)\lambda v}]^{1/(q-1)}} \quad (3.7)$$

com  $\lambda'_q = \lambda - (\lambda - \lambda_q)N(0)^{q-1}$ . Na Figura 3.1.2, a linha sólida contínua representa o melhor ajuste da Equação 3.7. Especificamente, nesse caso,  $\lambda'_q = 12227, 3$ ,  $\lambda_q = 135, 1$  e  $\alpha = 1/(q - 1) = 1, 03$  ( $q = 1, 97$ ) para deputados estaduais e  $\lambda'_q = 6774, 4$ ,  $\lambda_q = 33, 1$  e  $\alpha = 1/(q - 1) = 1, 07$  ( $q = 1, 93$ ) para deputados federais. Observe que o expoente  $\alpha$  está de acordo com o valor estimado anteriormente. Os processos de votação não-locais mostram um comportamento de escala universal. Porém, a mudança para o comportamento exponencial ocorre em diferentes frações (não universal) de votos. A diferença nos expoentes de escala que aparecem na lei de Zipf para representantes vereadores e deputados estaduais e federais propõem distintas dinâmicas subjacentes.



**Figura 3.1.3 – Ajuste via Lei de Zipf generalizada - Deputados Estaduais e Federais -** Os gráficos em log-log mostram a distribuição de votos para deputados estaduais e federais. Os círculos são referentes às eleições brasileiras no ano de 1998, considerando os estados com as 15 maiores capitais. A linha sólida é o melhor ajuste via lei de Zipf generalizada, Equação 3.7. O expoente de escala da lei de potência é  $\alpha = 1,03$  para (a) deputados estaduais e  $\alpha = 1,07$  para (b) deputados federais. Adaptada de Lyra et al. (2003).

### 3.2 Simulação de Distribuição de Votos em Rede

Eleições são processos dinâmicos e convincentes, onde, ao mesmo tempo, tem a interação entre os indivíduos e a influência externa (propaganda política, campanhas, etc). Levando em conta essas interações, pesquisadores utilizam modelos de rede na tentativa de simular os comportamentos observados na seção 3.1. No Brasil, nas eleições proporcionais (vereadores ou deputados) os eleitores votam diretamente para os candidatos e não para os partidos. Levando em conta que cada voto obtido é resultado de argumentos e convencimentos, Bernardes et al. (2002) introduziram um modelo para eleições proporcionais, com o objetivo de simular o comportamento

hiperbólico em distribuição de votos. O modelo é baseado no modelo de Sznajd, proposto para simular o processo de formação de opinião.

No modelo de Sznajd, pequenos grupos de pessoas influenciam seus vizinhos próximos, se e somente se, todas as pessoas dentro do grupo concordam. Uma generalização desse modelo foi aplicada em uma rede cúbica, que é uma versão tridimensional, e na rede de Barabási-Albert.

Uma rede cúbica de tamanho  $L \times L \times L$  representa o conjunto dos eleitores. Um número  $N_{tot}$  de candidatos,  $N_{tot} \ll L^3$ , é fixado no início da simulação. O valor de  $n = 1, 2, \dots, N_{tot}$  de um nó na rede  $S$  representa que o eleitor prefere aquele candidato  $n$ . O modelo tem dois estágios: primeiro, produz a condição inicial e depois realiza a simulação da campanha eleitoral (somente eleitores podem influenciar outros eleitores). Como nas eleições reais, não se espera por um tipo de estado de equilíbrio, mas conta-se os votos em algum momento intermediário. Basicamente, o que se faz é a análise durante o tempo transitório. Como nas eleições reais, os candidatos têm diferentes chances iniciais de serem votados (representando mais dinheiro para campanhas eleitorais, mais visibilidade inicial, etc). Isto é modelado por uma probabilidade  $P_c$  de convencimento, calculado a partir do rótulo  $n$  do candidato

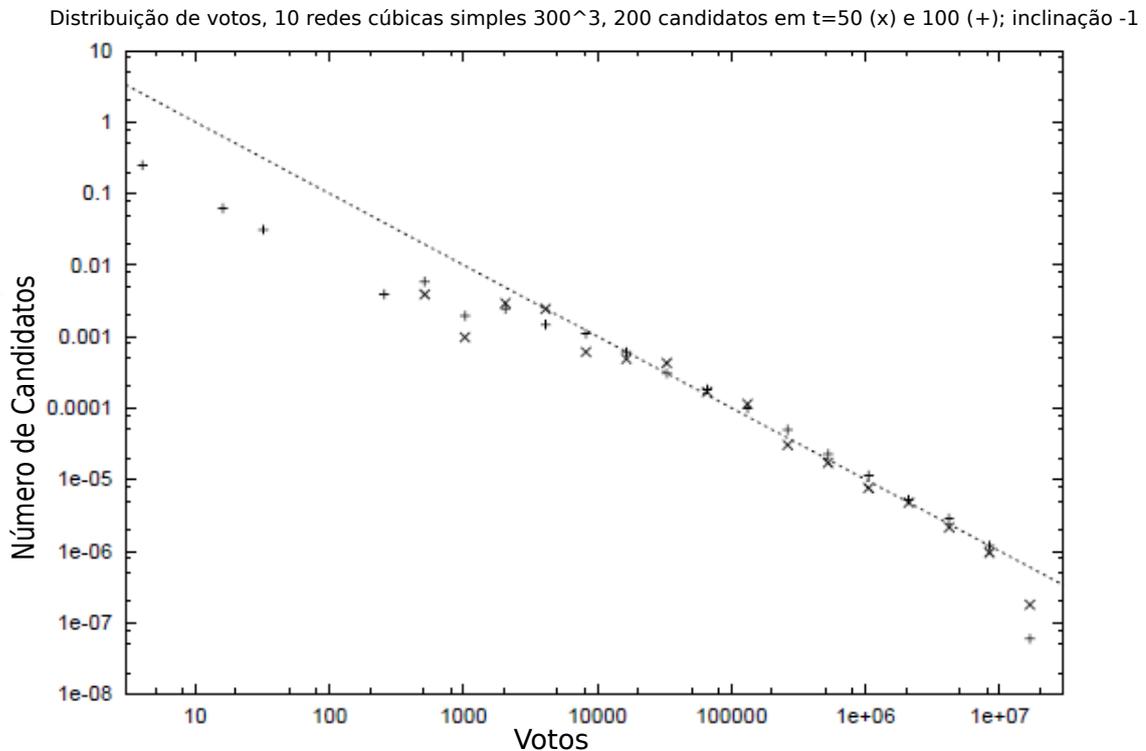
$$P_c = (n/N_{tot})^2. \quad (3.8)$$

Quanto maior o rótulo  $n$  de um candidato, maior é a probabilidade de convencer um eleitor.

Na primeira etapa, começa-se com todos os nós tendo valor zero, o que significa que não há eleitores comprometidos. Em seguida, visita-se todos os nós uma única vez em ordem aleatória. Para cada visita, tenta-se convencer o eleitor a adotar um candidato, escolhido aleatoriamente. Um número aleatório  $r$  é gerado e comparado com  $P_c$ . Se  $r \leq P_c$  o candidato é aceito por aquele eleitor. Se o candidato convencer o eleitor, este eleitor tentará convencer os vizinhos locais. Outra vez lança-se os dados e compara-se um novo número aleatório com  $P_c$ . Se for bem sucedido,  $r \leq P_c$ , o eleitor tentará convencer o bairro da seguinte maneira: verifica-se todos os seis locais vizinhos; para cada um que tenha o mesmo valor do candidato escolhido anteriormente, todos os dez nós vizinhos deste vínculo de dois nós assumirá o mesmo valor (como na prescrição usual de Sznajd). Se ninguém escolheu o mesmo candidato, somente o eleitor originalmente selecionado está comprometido com este candidato.

No segundo estágio, é realizado um processo usual de Sznajd sem a complicação da probabilidade  $P_c$  (assume-se que todos os eleitores são iguais e restringem a probabilidade  $P_c$  para descrever o convencimento dos candidatos). Um nó vizinho é escolhido aleatoriamente e, verifica-se os dois nós têm o mesmo valor, ou seja, preferem o mesmo candidato. Neste caso, todos os dez vizinhos votam naquele candidato.

A Figura 3.2.1 mostra desvios de uma lei de potência simples para números pequenos e grandes de votos.



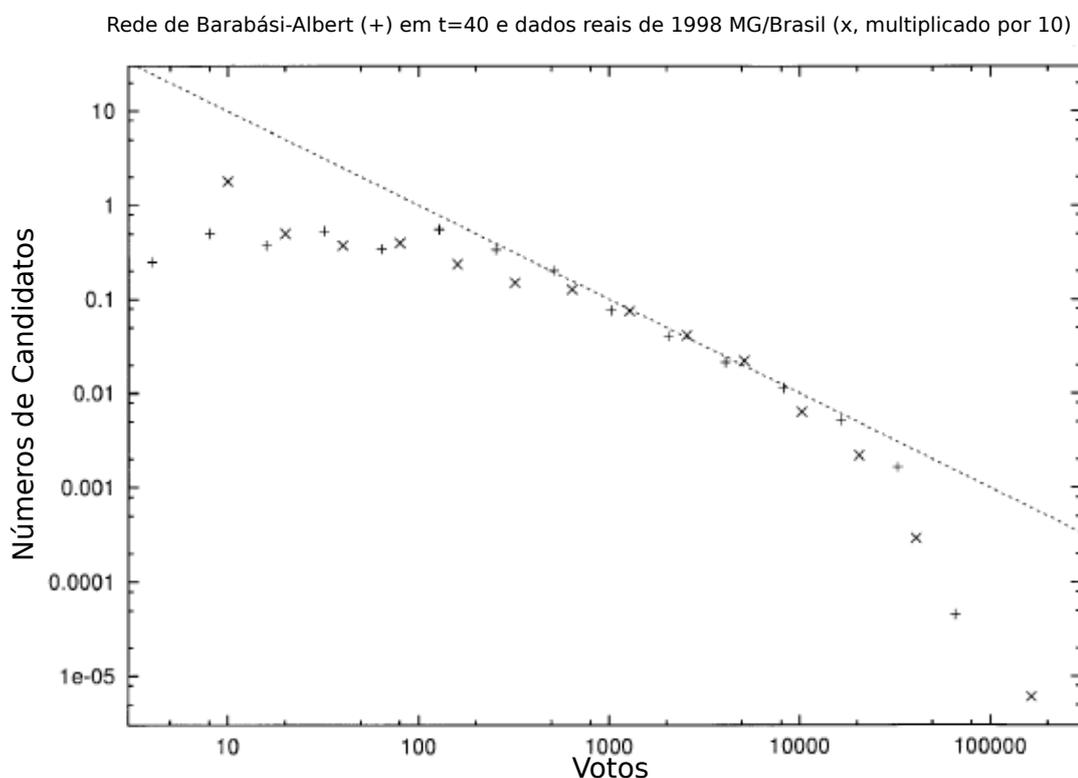
**Figura 3.2.1 – Rede Cúbica** - O gráfico mostra a distribuição do número de candidatos  $N(v)$  recebendo votos  $v$  na rede cúbica simples, após 50 e 100 interações. Eleições de 200 candidatos por 27 milhões de eleitores. Adaptada de Bernardes et al. (2002).

Para aplicar o modelo na rede de Barabási-Albert, a rede foi construída partindo de 6 nós e  $m = 5$ . Mil candidatos foram distribuídos aleatoriamente desconsiderando o número de conexões de um nó, e a partir daí iniciou a campanha. Em determinado espaço de tempo, todos os nós são visitados. Em cada nó, o procedimento é da seguinte maneira:

- Se um nó  $i$  já tem a preferência por um candidato, escolhe-se um nó conectado  $j$  aleatoriamente. Se o nó  $i$  não tem candidato ( $n = 0$ ), direciona-se aleatoriamente para outro nó selecionado.
- Se o nó  $j$  tem o mesmo candidato do nó  $i$ , eles tentam convencer todos os nós conectados a eles. A probabilidade de convencer os outros nós, para cada um dos dois nós, é inversamente proporcional ao número de nós conectados a ele, significando que cada nó convence, em média, um outro nó em cada processo.
- Se o nó  $j$  não tem candidato, o nó  $i$  tenta convencê-lo a aceitar o seu candidato com a mesma probabilidade descrita acima.

- Se o nó  $j$  tem um candidato diferente do nó  $i$ , segue-se para outro nó  $i$ .

A Figura 3.2.2 mostra que, exceto para valores pequenos e grandes de  $v$ , a lei hiperbólica concorda bem com o modelo para tempos intermediários como  $t = 40$ .



**Figura 3.2.2 – Rede Barabási-Albert** - O gráfico mostra a distribuição de  $N(v)$  para meio milhão de nós na rede Barabási-Albert, em que cada nó adicionado é ligado a 5 nós já existentes. Adaptada de Bernardes et al. (2002).

Existem várias possibilidades de se criar modelos que simulam dinâmicas de eleições, visto que uma dinâmica de opinião pode ser aplicada a vários tipos de redes. Nesse sentido, Travieso e Costa (2006) abordaram quatro modelos para simular o regime hiperbólico, sendo que, uma mesma dinâmica de opinião é aplicada a cinco tipos de rede.

Para a simulação do modelo de formação de opinião, adotaram-se as redes de Erdős-Rényi, Barabási-Albert, uma generalização do modelo de Barabási-Albert, uma malha bidimensional e uma malha bidimensional com conexões aleatórias. A rede Erdős-Rényi é caracterizada por uma distribuição de grau de Poisson e pela presença da propriedade de “mundo pequeno” (a distância média entre os nós é pequena e cresce lentamente com o número de nós na rede). O modelo de Barabási-Albert também possui propriedades de “mundo pequeno”, porém a distribuição de grau segue uma lei de potência. É perceptível que na rede regular a adição de nós aleatórios direciona à propriedade de “mundo pequeno” de maneira controlada. A generalização

do modelo de Barabási-Albert é descrita por uma distribuição de lei de potência, cujo o expoente é ajustável.

No modelo de opinião, com  $N$  eleitores (nós), começa-se distribuindo os candidatos  $C$  entre os nós escolhidos aleatoriamente com probabilidade uniforme, ou seja, cada candidato é atribuído para apenas um nó na rede (isso reflete o fato de que os candidatos também são eleitores). Os eleitores restantes começam como indecisos. O processo, a seguir, é repetido subsequentemente  $SN$  vezes: escolhe-se aleatoriamente um eleitor  $i$  que já tenha um candidato  $C_i$  associado; para todos os vizinhos do eleitor  $i$ , se são indecisos, são associados ao candidato  $C_i$ ; caso contrário, eles mudam de opinião, com probabilidade  $p$ . Este modelo de opinião é motivado pelo seguinte:

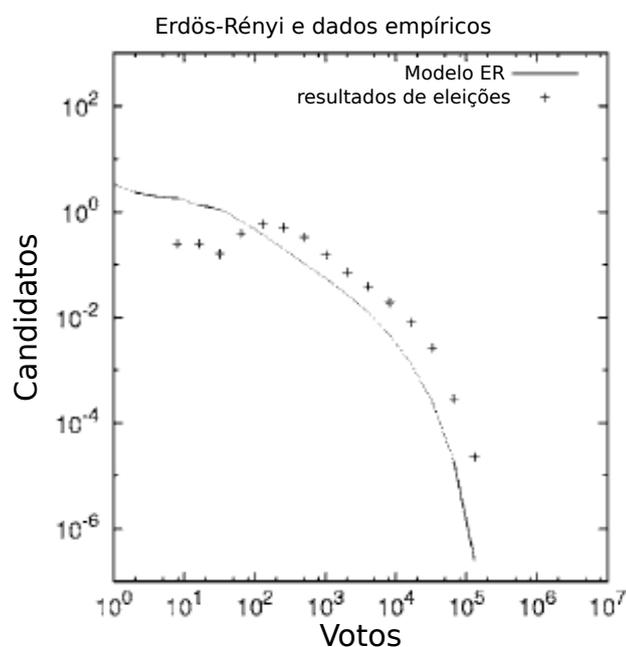
1. eleitores indecisos são passivos, no sentido de que não espalham a falta de opinião para outros eleitores;
2. eleitores indecisos são facilmente convencidos por eleitores que já têm a opinião formada;
3. a flexibilidade de mudar de opinião devido a uma interação, quantificada por um parâmetro  $p$ , é a mesma para todos os eleitores.

As simulações do modelo nas cinco redes consideradas geraram os seguintes resultados:

- O resultado para a rede de Erdős-Rényi mostra uma lei de potência para números intermediários de votos com expoente  $-1$ , um platô para pequenos números de votos e um corte para grandes números de votos.
- Para o modelo de Barabási-Albert, embora tenha apresentado dois regimes de lei de potência, nenhum deles corresponde ao valor experimental  $-1$ . A cauda da curva segue uma lei de potência com inclinação  $-1,45$ .
- Na malha bidimensional simples, não existe indício de um regime de lei de potência e um pico em torno de 1000 votos pode ser notado, não obtendo acordo com a natureza livre de escala.
- O efeito da adição de conexões aleatórias na malha bidimensional simples mostrou que a adição de apenas um pequeno número de novas conexões é o suficiente para obter um resultado semelhante ao do modelo Erdős-Rényi.
- Para verificar se o expoente da lei de potência da distribuição de candidatos com um determinado número de votos é influenciado pela inclinação da distribuição de

grau para redes com lei de potência, usou-se o modelo de Barabási-Albert generalizado. O modelo original de Barabási-Albert, tem expoente  $-3$  na distribuição de grau. Para a simulação do modelo de Barabási-Albert generalizado foram escolhidos dois conjuntos de parâmetros, para os quais os expoentes obtidos foram  $-2$  e  $-2,5$ . Os resultados mostraram que a inclinação da lei de potência da distribuição de votos quase não é afetada pela inclinação da distribuição de grau. De fato, ajustando para uma lei de potência na região entre 100 e 200.000 votos, obtém-se para o modelo original de Barabási-Albert (expoente igual  $-3$ ) um expoente de  $-1,45$ ; para o modelo generalizado (expoente igual  $-2,5$ ) um expoente de  $-1,44$  e (expoente igual  $-2$ ) um expoente de  $-1,40$ .

Uma comparação do modelo na rede de Erdős-Rényi com resultados de eleições reais é mostrado, na Figura 3.2.3, apresentando a distribuição do número de votos para eleições brasileiras em 1998 para deputado estadual no estado de São Paulo. Segundo Travieso e Costa (2006), os efeitos de “mundo pequeno” parecem ser de importância central para modelar o comportamento lei de potência dos dados reais. O modelo aplicado à rede Erdős-Rényi foi o que apresentou o melhor ajuste aos dados reais. Para os modelos de Barabási-Albert e da malha sem a adição aleatória de trajetórias curtas os resultados são inconsistentes com os de eleições reais como a de São Paulo.



**Figura 3.2.3 – Comparação do modelo Erdős-Rényi com resultados de eleições reais** - Dados referentes a eleições no estado de São Paulo em 1998, com 23.321.034 eleitores. Adaptada de Travieso e Costa (2006).

### 3.3 Relevância dos Partidos

Os resultados referentes as eleições proporcionais, apresentados até o momento, levam em conta apenas a distribuição do número de candidatos em função do número de votos. A seguir, apresenta-se resultados, em eleições proporcionais, que consideram a relevância dos partidos. Assim, o número de votos é reescalado de acordo com a “força” do partido para qual cada candidato pertence.

Para as discussões que seguem sobre eleições proporcionais, considera-se:

- $v_i$  o número de votos do candidato  $i$ ;
- $l_i$  a lista de candidatos do partido ao qual  $i$  pertence;
- $Q_i$  o número de candidatos da lista  $l_i$ ;
- $N_{l_i}$  o número total de votos recebidos pelos  $Q_i$  candidatos.

Fortunato e Castellano (2007), utilizando dados da Itália (1958, 1972 e 1987), Polônia (2005) e Finlândia (2003), obtiveram que a distribuição de probabilidade do número de votos recebidos por candidato  $P(v; Q; N)$  pode ser reescrita como uma função de uma única variável reescalada,  $F(vQ/N)$ .

Comparando as funções de escala  $F$  para os cinco conjuntos de dados obtém-se um resultado de universalidade ainda mais notável, a função  $F(vQ/N)$  é bem produzida pela função lognormal

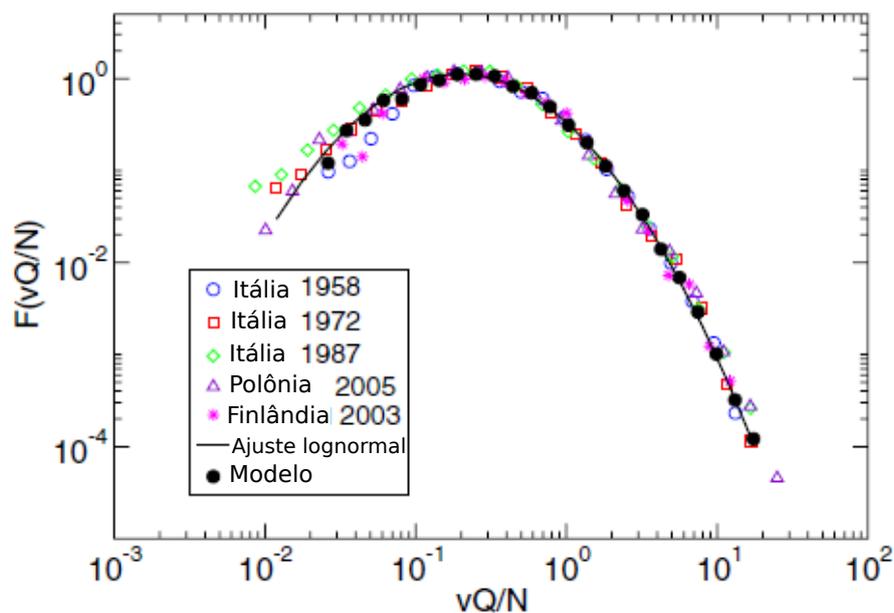
$$F(vQ/N) = \frac{N}{\sqrt{2\pi\sigma vQ}} e^{-(\log(vQ/N) - \mu)^2 / 2\sigma^2}, \quad (3.9)$$

com  $\mu = 0,54$  e  $\sigma^2 = -2\mu = 1,08$ .

Porém, o caráter universal desse resultado foi contestado por Araripe e Costa Filho (2009), já que para dados de eleições no Brasil e Finlândia apontam outro comportamento. Sendo assim, os autores utilizaram o mesmo método de Fortunato e Castellano (2007) para três eleições brasileiras, e o melhor ajuste para os dados é a seguinte função exponencial

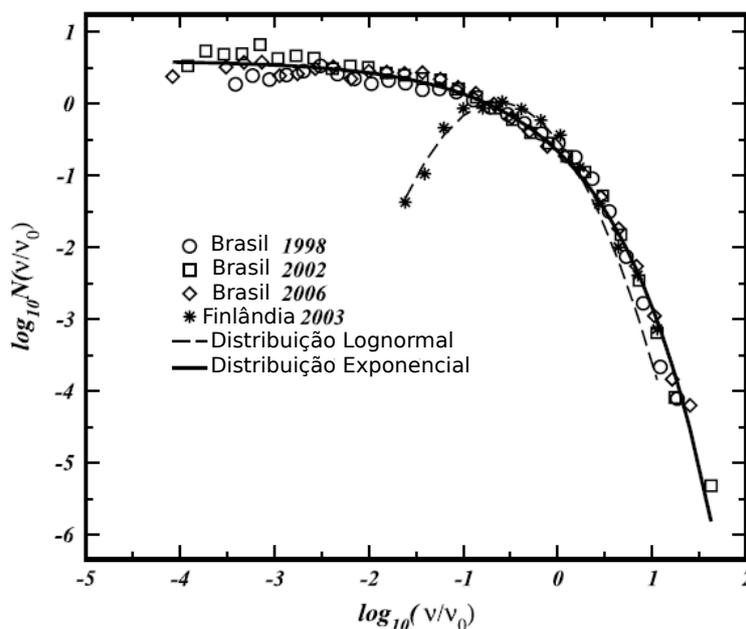
$$F(v/v_0) = \gamma e^{-\beta v/v_0} \quad (3.10)$$

em que  $v_0$  é o número médio de votos dos candidatos de cada partido,  $\gamma = 1,86$  e  $\beta = 1,262$ . No caso da Finlândia, a distribuição também é ajustada por uma lognormal, porém com parâmetros diferentes.



**Figura 3.3.1 – Universalidade da função de escalonamento  $F(vQ/N)$  em diferentes países e anos.** Adaptada de Fortunato e Castellano (2007).

Note na Figura 3.3.1 que o ajuste lognormal, realizado na curva polonesa, descreve muito bem os dados. A curva universal é bem produzida pelo modelo, onde a dinâmica das opiniões dos eleitores reflete a disseminação do boca a boca no eleitorado do partido.



**Figura 3.3.2 – Gráfico log-log de um histograma mostrando a distribuição de votos onde a variável foi normalizada de acordo com o número médio de votos dos candidatos do partido pertencente.** As curvas são ajustadas por diferentes funções. Para o Brasil a distribuição segue um decaimento exponencial para todos os anos, enquanto na Finlândia o melhor ajuste é uma lognormal. Adaptada de Araripe e Costa Filho (2009).

### 3.4 Processo de Candidatura

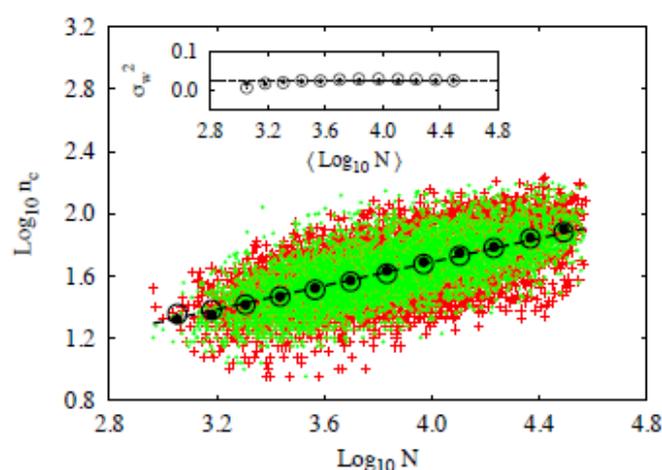
Diferentemente dos resultados apresentados até aqui, os quais analisaram resultados de eleições, os estudos apresentados a seguir investigaram dinâmicas de processos de candidatura eleitoral. Mantovani et al. (2013), abordaram uma questão aberta relacionada às eleições: a escolha dos candidatos à eleição (processo de candidatura). Consideraram os seguintes dados: 16 eleições do Brasil (1996, 2000, 2004, 2008 - eleições para prefeito e vereadores), Itália (2010 - eleições para prefeito), Inglaterra (2006, 2010 - eleições para vereadores), Canadá (2006, 2008 - eleições parlamentares) e Austrália (2004, 2007, 2010 - eleições parlamentares). Portanto, foram 5 eleições para prefeito, 6 eleições para vereadores e 5 eleições nacionais. Os dados referem-se ao número de candidatos e ao número de eleitores em cada divisão política local, sendo que, para o número de vereadores, foram consideradas cidades com menos de 47.000 habitantes, nas quais o número de vagas para vereadores é 9.

Investigaram como o número de candidatos ( $n_c$ ) aumenta com o número de eleitores ( $N$ ). A Figura 3.4.1 mostra o gráfico de dispersão das duas variáveis, em logaritmo de base 10, das eleições do Brasil para prefeito de 2008. Note que, uma tendência emerge do comportamento linear dos gráficos de dispersão dos dados em logaritmo, subjacentes às flutuações, e evidenciada pelo coeficiente Pearson  $r = 0,63$ .

Isto sugere que  $n_c$  aumenta com  $N$  como uma lei de potência, ou seja,  $n_c \sim N^\alpha$ . Um possível método para superar as flutuações e evidenciar a tendência a lei de potência dos dados, é construir janelas logaritmicamente espaçadas em  $N$  e tomar o valor médio dos pontos internos. Estes valores médios também são mostrados na Figura 3.4.1 por círculos abertos, e resultados empíricos mostrou-se muito robusto em relação ao número de janelas  $\omega$ . Com isso a relação da lei de potência torna-se evidente onde se tem

$$\langle \log_{10} n_c \rangle = A + \alpha \langle \log_{10} N \rangle, \quad (3.11)$$

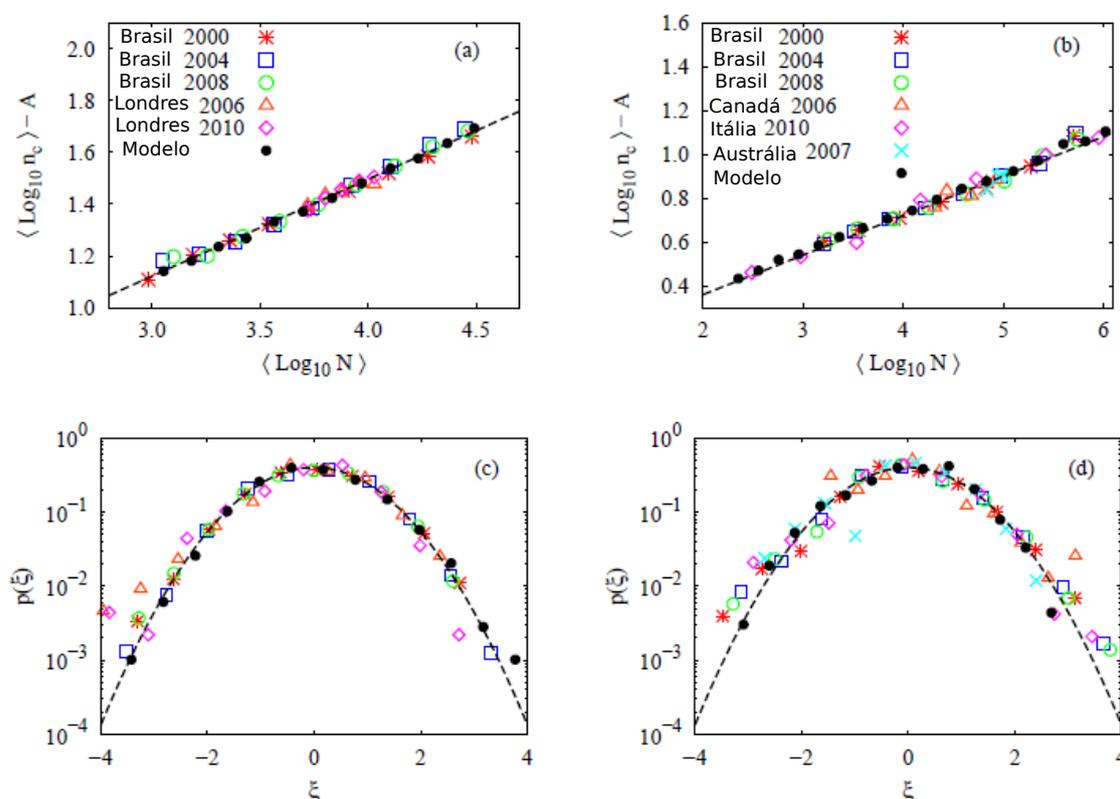
com  $\langle \dots \rangle$  representando os valores médios nas janelas e  $\alpha = 0,36$  para essa eleição.



**Figura 3.4.1 – 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 é 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  $\sigma_w^2$ . Adaptada de Mantovani et al. (2013).

A Figura 3.4.2 (a) mostra os valores médios para cinco eleições a prefeito, e em (b) faz o mesmo para quatro eleições a prefeito e duas eleições parlamentares. Note que, traçou-se estas relações descontando o valor constante de  $A$ . O colapso de boa qualidade e os valores de  $\alpha$  sugerem duas classes de universalidade: uma para eleições de prefeitos (único membro) caracterizada por  $\alpha \approx 0,18$  e outra para as eleições de vereadores (eleições de vários membros) com  $\alpha \approx 0,36$ . O resultado anterior indica que um eleitor também pondera o número de assentos em um processo de candidatura, ou seja, quanto mais representativo o cargo, maior o expoente.

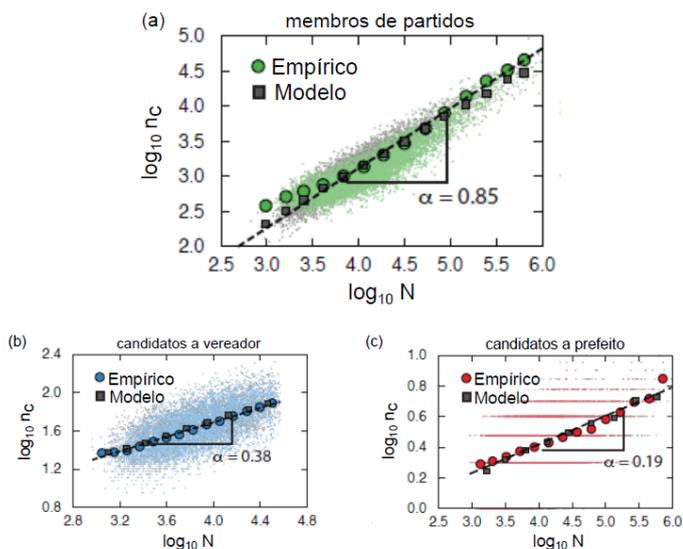
Agora aborda-se a questão da flutuação considerando a variância do logaritmo do número de candidatas  $\sigma_w^2$ . Para tal, emprega-se o mesmo procedimento utilizados para valores médios, ou seja, janelas logaritmicamente espaçada. Os resultados encontrados indicam basicamente que a variância não depende do número de eleitores  $N$ , que pode ser observado a partir da inserção da Figura 3.4.1. Além da variância, pode-se investigar as flutuações em torno da relação de lei de potência. Para isso, considere a variável  $\xi = \frac{\log_{10} n_c - f_w(N)}{\sigma_w}$  onde  $f_w(N) = A + \alpha \langle \log_{10} N \rangle_w$  representa a função ajuste para os valores médios considerando cada janela  $w$ . As Figuras 3.4.2 (c) e 3.4.2 (d) mostram a função de distribuição da probabilidade para as mesmas eleições das Figuras 3.4.2 (a) e 3.4.2 (b). A partir dessas figuras, observe que  $\xi$  segue muito perto do padrão gaussiano.



**Figura 3.4.2 – Valores médios** - Os gráficos superiores mostram os valores médios, em log-log, do número  $n_c$  de candidatas 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 o valor médio do expoente  $\alpha$ , onde  $\alpha = 0,36$  para eleições de multi-membros e  $\alpha = 0,18$  para eleições de um membro. Nos gráficos debaixo são mostradas as distribuições de flutuação para (c) as mesmas eleições com múltiplos membros da Figura (a) e (d) para as mesmas eleições de um único membro da Figura (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 BA com  $\beta = 0,32$  ( $\beta = 0,41$ ) para a eleição de múltiplos membros (único membro). As outras eleições apresentam o mesmo padrão. Adaptada de Mantovani et al. (2013).

Mantovani et al. (2013) analisaram, ainda, os dados de filiados a partidos políticos com o objetivo de tentar compreender como o eleitor pondera sua candidatura a uma posição política. Os dados eleitorais consistem do número de candidatos em eleições para prefeito e vereador em todas as cidades brasileiras nos anos de 2000, 2004 e 2008, considerando 9 o número de vereadores. Para o número de membros de partidos políticos, os dados coletados são referentes aos anos de 2006, 2008 e 2011 para todos os 27 partidos políticos brasileiros em cada cidade.

Dando continuidade as investigações de candidatura e acrescentando agora a variável número total de filiados a partidos políticos por cidade, para dados do Brasil, investigou-se a hierarquia da posição política entre as posições de filiados, vereadores e prefeitos. Para isso, de maneira análoga a utilizada no primeiro trabalho, investigou-se a relação entre  $\log_{10}n_c$  e  $\log_{10}N$ , no qual  $N$  representa o número de eleitores e  $n_c$ , neste caso, o número de membros de partidos políticos candidatos a vereador e candidatos a prefeito. A Figura 3.4.3 (a) mostra o espalhamento dessas relações quando se considera o número de membros de partidos políticos, em (b) o número de candidatos a vereador e em (c) o número de candidatos a prefeito.



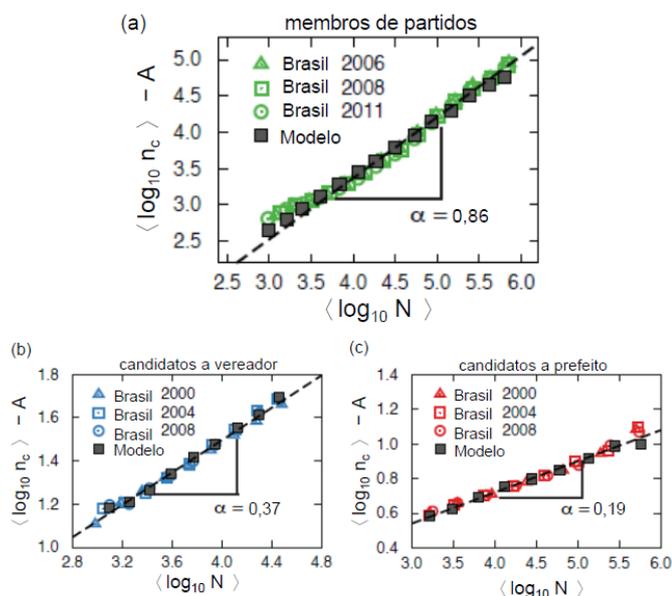
**Figura 3.4.3 – Espalhamento dos Dados** - Os gráficos mostram a relação entre  $\log_{10}n_c$  e  $\log_{10}N$ , no qual  $N$  é o número de eleitores e  $n_c$  o número total filiados a partidos políticos em (a), candidatos a vereador em (b) e candidatos a prefeito em (c). Os dados são referentes ao ano de 2008. Os pontos pequenos coloridos representam o gráfico de dispersão dos dados empíricos ((a) filiados em verde, (b) vereadores, (c) prefeitos em vermelho), enquanto os pontos pequenos em cinza representam os dados simulados pelo modelo analítico. Os círculos coloridos, de mesma cor dos pontos de dispersão, representam os valores médios dos dados empíricos em janelas igualmente espaçadas da variável  $\log_{10}N$  e os quadrados cinza representam os valores médios, nas mesmas janelas, dos dados simulados. As linhas tracejadas são os ajustes lineares para os valores médios e os expoentes das leis de potências são mostrados em cada gráfico. Adaptada de Mantovani et al. (2013).

Subjacentes à flutuações, observa-se uma tendência a um regime de lei de potência. Tendo em vista superar as flutuações, ou seja, identificar um comportamento médio para os dados, novamente  $\log_{10}N$  foi tomada em janelas  $\omega$  igualmente espaçadas e calculados os valores médios de  $\log_{10}n_c$  e  $\log_{10}N$  em cada janela. Os círculos na Figura 3.4.3 representam esses valores médios para os quais nota-se uma relação alométrica (lei de potência) que pode ser escrita como a relação linear

$$\langle \log_{10}n_c \rangle = A + \alpha \langle \log_{10}N \rangle, \quad (3.12)$$

a qual descreve muito bem esses valores médios em escalas log-log. Ainda, na Figura 3.4.3, verificam-se os diferentes valores de  $\alpha$  (expoente da lei de potência) em cada um dos casos: membro de partidos políticos, candidato a vereador e candidato a prefeito. A Figura 3.4.4 mostra os valores médios obtidos por esse procedimento para toda a base de dados, identificando que os valores de  $\alpha$  não mudam com o passar do tempo para membros de partidos políticos, candidatos a vereador e a prefeito. De fato, nos anos de 2006, 2008 e 2011, encontrou-se  $\alpha = 0,87 \pm 0,02$ ,

$\alpha = 0,86 \pm 0,02$  e  $\alpha = 0,85 \pm 0,02$ ; para candidatos a vereador, nos anos 2000, 2004 e 2008,  $\alpha = 0,37 \pm 0,02$ ,  $\alpha = 0,37 \pm 0,02$  e  $\alpha = 0,38 \pm 0,02$  e para candidatos a prefeito, nos anos de 2000, 2004 e 2008,  $\alpha = 0,19 \pm 0,01$ ,  $\alpha = 0,20 \pm 0,01$  e  $\alpha = 0,19 \pm 0,01$ . Portanto, não se pode rejeitar a hipótese de que os valores de  $\alpha$  são constantes no tempo para cada posição política depois de levar em conta os intervalos de confiança.



**Figura 3.4.4 – Leis de potências dos dados -** Relações alométricas médias entre  $\log_{10}n_c$  e  $\log_{10}N$ , sendo  $N$  números de eleitores,  $n_c$  o número de membros de partidos em (a), o número de candidatos a vereador em (b) e a prefeito em (c). Para cada posição política, os valores de  $\alpha$  (a inclinação das retas ajustadas) praticamente não mudam com o tempo. O valor médio de  $\alpha$  é mostrado em cada gráfico. As linhas pontilhadas representam ajustes dos valores médios dos dados empregando os expoentes médios. Adaptada de Mantovani et al. (2013).

### 3.5 Lei de Potência Universal nas Eleições Presidenciais dos Estados Unidos

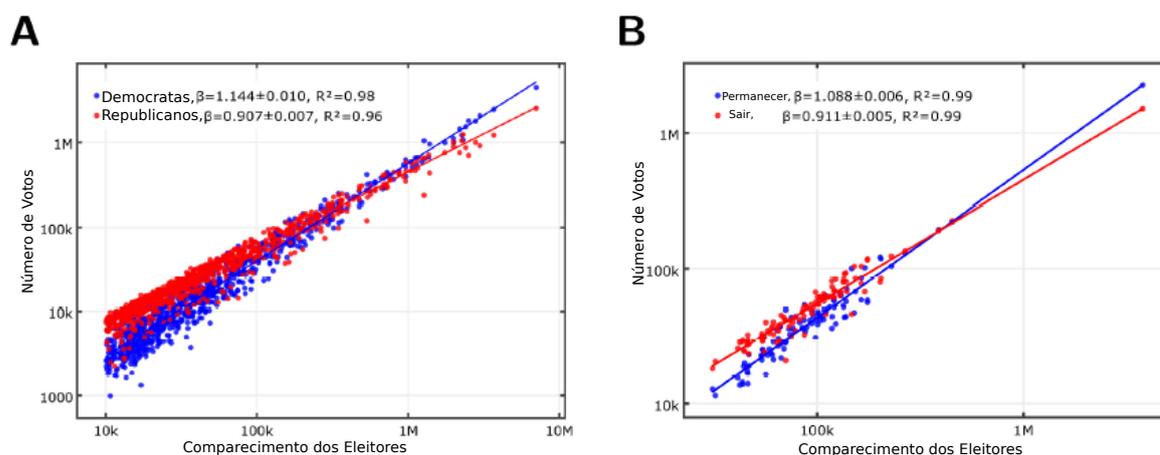
Bokányi et al. (2018), analisaram os resultados das eleições presidenciais para os dois partidos (Democrata e Republicano) nos Estados Unidos no período de 1948 a 2016 e, o resultado do referendo da União Europeia em 2016 no Reino Unido, mostrando que as curvas de escala são universais e dependem apenas de um único parâmetro:

$$Y = Y_0 \cdot N^\beta, \quad (3.13)$$

sendo  $\beta > 1$  a escala é superlinear e  $\beta < 1$  escala sublinear.

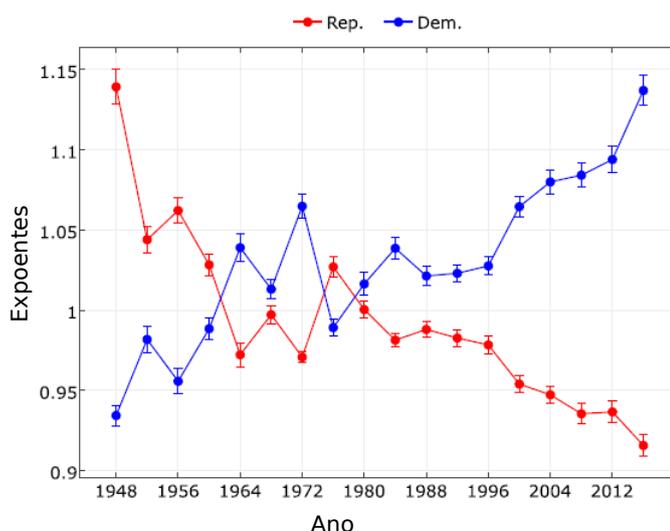
A Figura 3.5.1 (A) mostra a quantidade de votos para os Democratas e Republicanos em função do comparecimento dos eleitores e, a Figura 3.5.1 (B) mostra

a quantidade de votos para a permanência e saída do Reino Unido em função do comparecimento dos eleitores. Observe que os votos para Democratas e “Permanecer na União Europeia” tem escala superlinear com expoentes  $\beta_D \approx 1,14$  e  $\beta_{rem} \approx 1,09$ , enquanto que, os votos para Republicanos e “Deixar a União Europeia” tem escala sublinear com expoentes  $\beta_R \approx 0,92$  e  $\beta_{rem} \approx 0,91$  e o coeficiente de regressão  $R^2 \geq 0,9$  para ambos os partidos. Mesmo as eleições ocorrendo em duas situações políticas diferentes, mostram expoentes semelhantes.



**Figura 3.5.1 – Escala urbana nas eleições presidenciais nos Estados Unidos e no 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.  $\beta$  é o ajuste de inclinação linear e  $R^2$  é o coeficiente de regressão nas inserções. Em (B) os votos para Permanecer (azul) e Sair (vermelho) em função do comparecimento dos eleitores para os distritos eleitorais do Reino Unido. Adaptada Bokányi et al. (2017).

A Figura 3.5.2 mostra os expoentes dos Republicanos e Democratas para as 18 eleições presidenciais no período de 1948 à 2016. Note que quando um expoente dos Republicanos é superlinear, o expoente dos Democratas é sublinear e vice-versa. O expoente dos Democratas tem um crescimento, enquanto o expoente dos Republicanos decresce.



**Figura 3.5.2 – Gráfico dos expoentes** - Escala dos expoentes para os Republicanos (vermelho) e Democratas (azul) com barras de erro para as 18 eleições presidenciais de 1948 a 2016. Adaptada Bokányi et al. (2017).

A relação dos dois expoentes torna-se aparente quando o expoente Republicano é traçado em função do expoente Democrata na Figura 3.5.3 (A).

Para cada eleição e partido determina-se o expoente de escala  $\beta$  e a constante  $Y_0$  independentes dos ajustes. Na Figura 3.5.3 (B) mostra o gráfico  $\log Y_0$  em função  $\beta$ , obtendo:

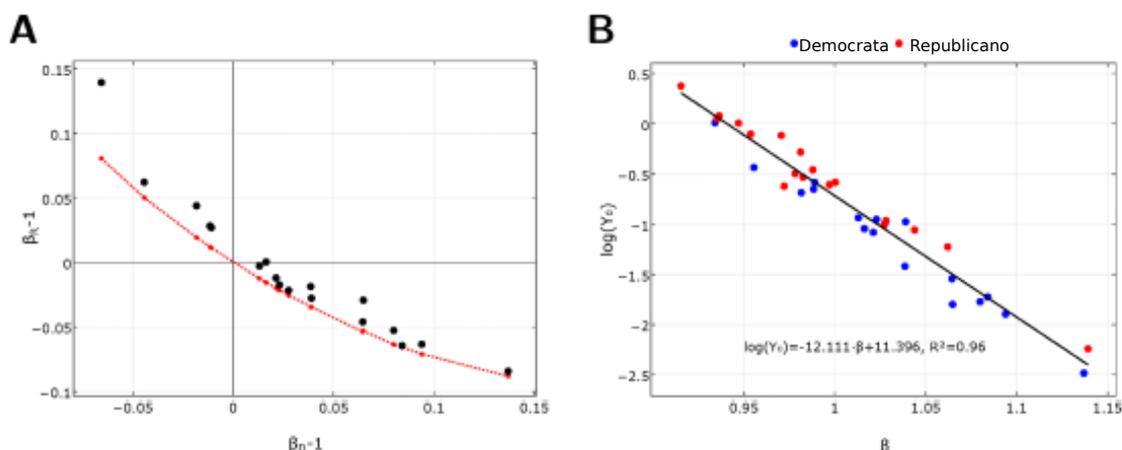
$$\log Y_0 = -\alpha\beta + \delta, \quad (3.14)$$

sendo  $\alpha = 12,111$  e  $\delta = 11,396$ .

Um dos partidos sempre mostra a escala superlinear ( $\beta > 1$ ) e orienta o processo, enquanto o expoente sublinear do outro partido é apenas conservação da probabilidade:

$$\frac{1}{2}\langle(N/N^*)^{\beta_D-1}\rangle + \frac{1}{2}\langle(N/N^*)^{\beta_R-1}\rangle = 1. \quad (3.15)$$

Esta equação garante que um expoente será superlinear, enquanto o outro é sublinear. A solução numérica é mostrada na Figura 3.5.3 (A). Portanto, um modelo e um expoente derivado para os resultados de um dos partidos determinará os resultados do outro partido através da conservação de probabilidade.



**Figura 3.5.3 – Correlação dos parâmetros de escala urbana nas eleições dos Estados Unidos** - Em (A), escala urbana dos expoentes 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  $\log Y_0$  em função dos expoentes  $\beta$  para Republicanos (vermelho) e Democratas (azul) para eleições presidenciais no período 1948 a 2016. Linha com com parâmetros e coeficiente de regressão no intercepto. Adaptada Bokányi et al. (2017).

Em seguida analisou-se os resultados no período de 2000-2016, onde os expoentes dos Democratas são superlineares e os expoentes dos Republicanos são sublineares. Mostra-se que a distribuição estatística dos resultados para os Democratas está de acordo com o modelo GLPLH, enquanto que a distribuição dos votos para os Republicanos é simplesmente a consequência da probabilidade de conservação.

Um indicador metropolitano ajustado à escala (SAMI) é o desvio logaritmo do valor  $Y_i$  da curva de escala média para uma cidade com população  $N_i$

$$\xi_i = \log Y_i - \log Y_0 - \beta \log N_i. \quad (3.16)$$

Os SAMIs para um determinado tamanho de cidade são, normalmente, distribuídos se a medida investigada obedece as leis de escala urbana. No modelo GLPLH a variância pode ser expressa com o parâmetro de complexidade  $q$  e o número de fatores complementares  $M$  como  $\sigma_{SAMI}^2 = q^2 M b (\log N_0 - \langle \log N \rangle)$ , onde  $\langle \log N \rangle$  é a média do logaritmo dos tamanhos das cidades. Também pode ser expresso com o expoente de escala

$$\sigma_{SAMI}^2 = q(\beta - 1)(\log N_0 - \langle \log N \rangle). \quad (3.17)$$

### 3.6 A falta de economia em eleições proporcionais

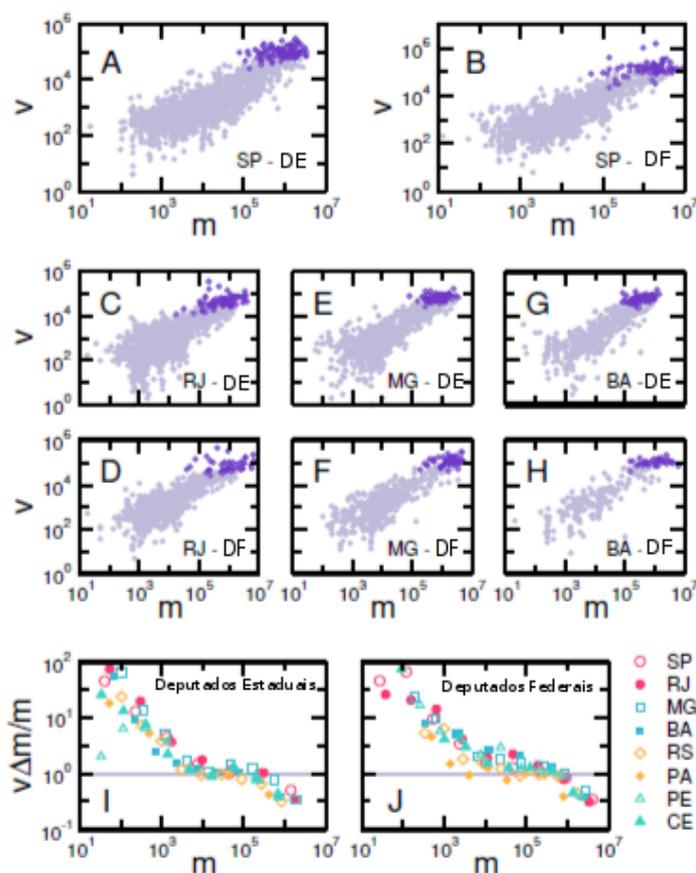
A presente seção aborda a investigação de Melo et al. (2018) sobre como as despesas das campanhas eleitorais influencia no resultado da eleição. O custo das

campanhas eleitorais aumenta cada vez mais e, com isso, aumenta-se a necessidade da campanha ter um planejamento eficaz e um retorno do investimento. Mesmo diante do impacto das eleições na vida diária da sociedade, a forma como esse investimento é traduzido em votos continua desconhecido.

Os dados analisados são das eleições proporcionais no Brasil realizadas em 2014. Elaborou-se os conjuntos de dados do resultado eleitoral e as despesas da campanha dos candidatos de todos os 26 estados brasileiros. A Figura 3.6.1 mostra 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. Conforme está representado na Figura 3.6.1, as nuvens de pontos estão perfeitamente correlacionadas e seguem uma tendência clara.

Para extrair a principal relação entre  $v$  e  $m$ , tem-se a média do número de votos em log-log espaçados ao longo de  $m$ , que fornece uma estimativa para a relação empírica de  $\langle v \rangle$  em função de  $m$ . Para plotar resultados diferentes na mesma figura, realiza-se uma escala de transformação em  $\langle v \rangle$  supondo uma relação linear simples  $\langle v \rangle = c \times m$ , onde  $c$  é uma constante característica de uma determinada eleição. Se definir o preço médio de uma votação como  $\Delta m = \sum_i m_i / \sum_i v_i$ , onde  $v_i$  é a quantidade de votos do candidato  $i$ , e supor que seja aproximadamente uniforme entre os candidatos é possível ver que  $c = 1/\Delta m$ . Se a relação entre votos e dinheiro é linear, então  $\langle v \rangle \times (\Delta m/m)$  deve ser uma função constante de  $m$  com valor próximo de 1.

As Figuras 3.6.1 (I) e 3.6.1 (J), mostra  $\langle v \rangle \Delta m/m$  em função de  $m$  para os cargos de deputados estaduais e federais nos oito estados mais populosos do Brasil nas eleições de 2014. O resultado mostra uma consistente dependência não-trivial de votos em dinheiro gasto em campanha. Para valores pequenos de  $m$ , observa-se uma rápida diminuição de  $\langle v \rangle \Delta m/m$ . Para gastos intermediários, na faixa de R\$ 10.000,00  $< m < R\$ 100.000,00$ , observa-se uma aparente dependência linear de  $v$  em relação a  $m$ . Para  $m > R\$ 100.000,00$ , nota-se linearidade, ou seja, os candidatos com mais recursos financeiros precisam desproporcionalmente de uma grande quantidade de dinheiro para obter voto único em comparação com candidatos menos aprovados, dentro do mesmo intervalo de recursos.



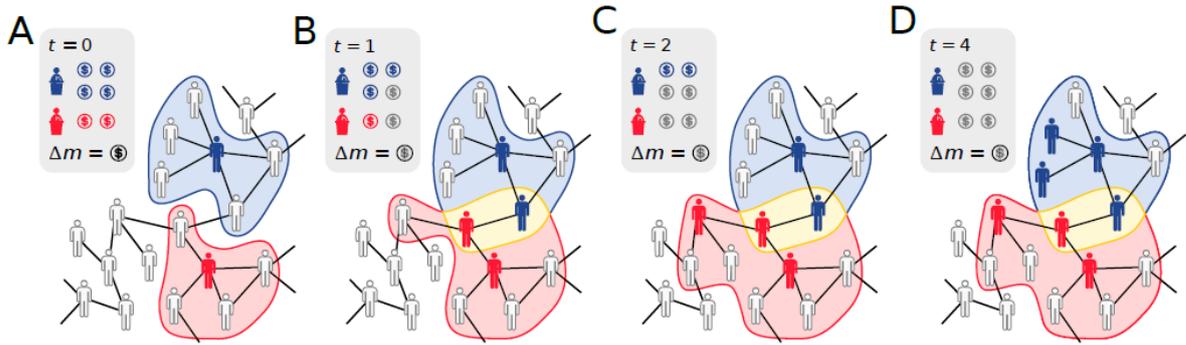
**Figura 3.6.1 – Relação de Escala entre o número de votos e as despesas da campanha** - Os círculos roxos claros mostram a relação entre o número de votos e as despesas declaradas da campanha de cada candidato a deputado estadual (DE) e deputado federal (DF) nas eleições em 2014 para os quatro maiores 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 representam a relação normalizada  $\langle v \rangle \Delta m/m$ , onde primeiro calcula-se o número médio de votos em caixas com espaçamento de log-log ao longo de  $m$ . Adaptada Melo et al. (2017).

Se assumir uma correlação linear, a constante multiplicativa é  $\Delta m = M/n$ . A normalização fornece uma observação direta da não-linearidade na dependência de votos em dinheiro. Vê-se um comportamento sublinear global, onde os candidatos mais ricos exibem uma fração menor de votos por dinheiro.

Considerando um processo eleitoral composto por dois grupos de indivíduos, candidatos e eleitores, tem-se que todos os candidatos podem competir pelo voto de todos os  $n$  eleitores. Além disso, cada candidato  $i$  tem uma quantidade limitada de dinheiro  $m_i$  para gastar em suas campanhas. Então, em um dado momento  $m_i = 0$ , o candidato torna-se incapaz de competir por mais eleitores. Supõe-se que os candidatos

só podem conquistar um único voto em tempo determinado e que os eleitores, ao decidirem o voto, não alteram mais.

Uma descrição pictórica do modelo é apresentado na Figura 3.6.2. Em uma rede social com eleitores indecisos, representados por indivíduos cinza claro, dois candidatos começam suas campanhas com uma quantia inicial de dinheiro  $m$  e um único eleitor decidido. Este início é representado na Figura 3.6.2 (A) pelos indivíduos azul e vermelho, onde o candidato azul tem inicialmente um orçamento de 4 unidades monetárias, enquanto o candidato vermelho tem um orçamento de 2 unidades monetárias. As regiões destacadas em azul e vermelho representam as áreas operacionais das campanhas, incluindo o grupo de eleitores indecisos que as campanhas irão investir para transformar em eleitores decididos. A Figura 3.6.2 (B) mostra, em cada etapa de tempo, que cada campanha escolhe um eleitor dentro de suas áreas operacionais. Se o indivíduo escolhido é um eleitor indeciso, torna-se um eleitor decidido. Como consequência, o total de dinheiro da campanha é diminuído uma quantidade de  $\Delta m$ . Se o eleitor escolhido já decidiu o seu voto, conforme mostra a Figura 3.6.2 (C), o orçamento da campanha é também diminuído por  $\Delta m$ , mas a decisão do eleitor permanece inalterada. Repete-se este procedimento até todas as campanhas ficarem sem recursos. Na Figura 3.6.2 (D), mostra-se um exemplo típico de uma competição por votos entre dois candidatos durante o processo eleitoral descrito pelo modelo. Embora o candidato que tem o orçamento inicial maior receba mais votos no final das eleições, devido as despesas ineficazes, a campanha do candidato com menos recursos é, de fato, mais eficiente.



**Figura 3.6.2 – Esboço do Modelo Pictórico** - Dois candidatos competem por votos em uma rede social com indivíduos indecisos (cinza claro). (A) 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. A região representada em amarelo é onde as ambas 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 adotantes, mas sua campanha é menos eficiente, resultando em uma não economia de escala. Adaptada Melo et al. (2017).

A evolução temporal do número de votos de um candidato  $i$  é descrito como

$$\frac{dv_i}{dt} = \left(1 - \frac{S(t)}{n}\right) [m_i(t) > 0], \quad (3.18)$$

onde  $S(t) = \sum_{i=0}^s v_i$  é o número total de eleitores decidido no tempo  $t$ , e  $[m_i(t) > 0]$  é o suporte de Iverson que é 1 se a condição dentro dos colchetes for satisfeita e 0 caso contrário.

Como a probabilidade do candidato  $i$  conquistar um eleitor indeciso não é afetado por outra campanha  $S(t)$  pode ser substituído por  $v_i$  na Equação 3.18, levando a um desacoplado sistema de equações diferenciais, cuja a solução é dada por,

$$v_i = n - (n - v_{0,i})e^{-m_i/n\Delta m}, \quad (3.19)$$

onde  $v_{0,i}$  é o número inicial de votos do candidato  $i$ . Desde  $n\Delta m \gg m_i$ , e assumindo que  $(n - v_{0,i}) \approx n$ , expandindo a exponencial e tomando sua primeira ordem de aproximação, pode-se escrever o número de votos como  $v_i \approx v_{0,i} + m_i/\Delta m$ . Na Figura 3.6.3, os dois primeiros regimes pode ser entendido em termos de aproximação. Para o regime de baixo  $m$ , onde os dados experimentais não apresentam uma correlação clara, os candidatos começam a corrida com  $v_{0,i}$  votos. Quando eles não podem a um longo

prazo e/ou uma grande despesa, o seu desempenho final ocorre em torno do valor inicial  $v_{0,i}$ , que depende de diferentes fatores, como o engajamento voluntário livre. À medida que o dinheiro da campanha aumenta, a parte linear supera o número inicial  $v_{0,i}$  e surge um regime linear. No entanto, no cenário sem competição, o comportamento linear permanece grande  $m$ .

Considere agora a competição entre os candidatos como uma causa possível para a transição de um regime linear para a sublinear. Integrando a Equação 3.18, encontra-se,

$$v_i = v_{0,i} + \frac{m_i}{\Delta m} - \frac{1}{n\Delta m} \int_0^{m_i} S(m') dm' \quad (3.20)$$

a qual tem uma interpretação simples. Como é o sem competição, todos os candidatos começam sua corrida com um número inicial de votos, e aqueles com dinheiro suficiente para continuar investindo entram em um regime linear controlado pela taxa  $\Delta m/m$ . No entanto, candidatos com recursos de campanhas suficientes podem começar a desperdiçar o dinheiro com eleitores já decididos, um comportamento que é fundamentado pela presença de  $S(m')$  no último termo da Equação 3.20, que inclui a dinâmica da competição.

Para obter uma solução para o modelo, usou-se como entradas o dinheiro  $m_i$  de cada candidato  $i$ , obtido de dados, o número total de eleitores  $n$ , um número inicial de votos  $v_0$  e um valor estimado para  $\Delta m$ . Para todos os candidatos define-se  $v_0$  como o número médio de votos de candidatos com menos de R\$ 1.000,00. O parâmetro  $\Delta m$  é calculado como uma função da taxa de comparecimento dos eleitores  $T = S_f/n$ , onde  $S_f = \lim_{t \rightarrow \infty} S(t)$  é o número total de votos no estado estacionário. Portanto, escreve-se a fração final de votos como

$$T = 1 - e^{-M/(n\Delta m)}, \quad (3.21)$$

onde  $M = \sum_i m_i$  é a quantia total de recursos financeiros no processo da campanha.

Na abordagem de modelagem, de Melo et al. (2017), a distribuição de votos emerge como um resultado natural da distribuição dos recursos financeiros  $P(m)$ . A Figura 3.6.3 (A) mostra, que a distribuição  $P(m)$  calculada para deputados estaduais de três estados brasileiros, São Paulo, Rio de Janeiro e Minas Gerais, podem ser descritas em termos de lei de potência de decadência que se estende por uma região de aproximadamente seis ordens de magnitude. Usando as distribuições como entradas, determina-se  $P(v)$  para cada eleição. Na Figura 3.6.3 (B), compara-se a distribuição de votos empírica para o estado de São Paulo com a obtida pelo modelo, que reproduz corretamente a distribuição empírica de votos entre os candidatos,  $P(v)$ , por mais de duas ordens de magnitude. Isso implica que a forma de cauda longa não Gaussiana observada tem sua origem no aspecto heterogêneo da distribuição de recursos da

campanha, independentemente da intrincada rede social e dinâmica da informação por trás do processo eleitoral.

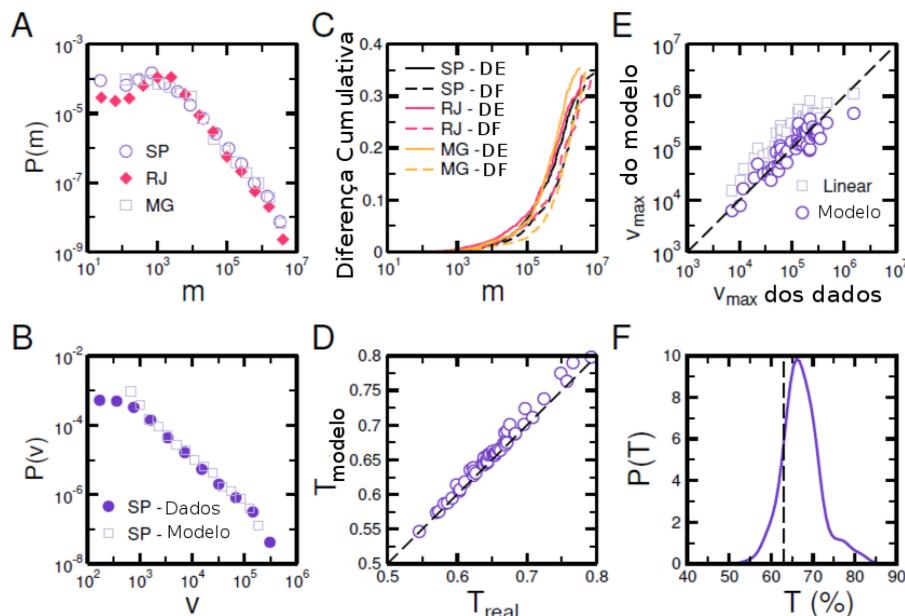
Para destacar-se o efeito de sublinearidade na previsão de uma eleição, calculou-se a diferença relativa entre a distribuição acumulada de votos prevista pelo modelo linear sem competição e o previsto pelo modelo com competição. Para as eleições dos deputados estaduais e deputados federais, Figura 3.6.3 (C), nos três principais estados brasileiros, ou seja, São Paulo, Rio de Janeiro e Minas Gerais, não é notada uma diferença significativa entre as duas previsões para campanhas que investiram menos de R\$ 10.000,00. No entanto, para campanhas eleitorais que investiram mais de R\$ 10.000,00, uma discrepância substancial entre as previsões pode ser notado. Esta região de grandes custos, a diferença acumulada pode ser drástica, acima de 30% em alguns casos ( $m > 1.000.000,00$ ).

Confirmou-se a validade do modelo em comparação com os dados das eleições estaduais e federais em 2014 que ocorreram simultaneamente nos 26 estados do Brasil. A Figura 3.6.3 (D) mostra, em que cada ponto corresponde a uma eleição em um determinado estado, os resultados do modelo para a taxa de comparecimento dos eleitores  $T$ , como previsto pela Equação 3.21, são compatíveis com os dados observados. A capacidade preditiva do modelo pode ser efetivamente testada comparando sua estimativa com dados reais para o maior número de votos obtidos por um candidato em cada eleição,  $v_{max}$ . Enquanto os resultados do modelo (círculos) se reúnem em torno da linha da identidade, demonstrando boa concordância quantitativa com os dados reais, o modelo de aproximação linear,  $v_{max} \approx v_0 + m_{max}/\Delta m$ , claramente superestima os valores de  $v_{max}$ , como é observado na Figura 3.6.3 (E).

Para se ter a previsão da taxa de comparecimento dos eleitores  $T$ , deve-se considerar as seguintes hipóteses:

- os candidatos tem conhecimento do montante total de recursos  $M$  durante a campanha;
- $\Delta m = M/n$ , que corresponde à divisão mais simples e equitativa de votos.

Este último ponto é equivalente a supor que uma participação completa pode ser alcançada,  $T = 100\%$ , como no caso sem competição. Em outras palavras, os candidatos elaboram sua estratégia pressupondo que eles irão obter o número máximo possível de votos, desconsiderando, portanto, a competição entre eles. Este argumento heurístico leva a uma fração de votos válidos,  $T = 1 - e^{-1} \approx 0,63$ . Na Figura 3.6.3 (F), é mostrado o histograma do número total de votos válidos para todas as eleições do Congresso nos anos de 2006, 2010 e 2014 indicando um valor médio de participação de 0,67, que está em estreita concordância com a previsão do modelo.



**Figura 3.6.3 – Resultados Analíticos do Modelo** - Para obter a distribuição de votos, o modelo toma como entrada a distribuição de recursos financeiros. (A) A distribuição de recursos financeiros para os deputados estaduais em São Paulo (SP), Rio de Janeiro (RJ) e Minas Gerais (MG) revelam características de caudas longas. (B) Comparação da distribuição real de votos (círculos),  $P(v)$ , com os obtidos pelo modelo (quadrados) para eleição de representantes do estado de São Paulo. (C) Diferença relativa entre as distribuições acumuladas das previsões para o modelo com e sem competição. Apresenta-se os resultados para as eleições dos deputados estaduais (DE) e deputados federais (DF) para São Paulo (SP), Rio de Janeiro (RJ) e Minas Gerais (MG). (D) Em todas as 56 eleições parlamentares de 2014, compara-se a estimativa do índice de comparecimento,  $T_{real} = T_{modelo}$ . Como pode ser observado, as simulações (círculos) exibem uma boa concordância com os dados. (E) Observa-se que o modelo de competição (círculos) melhor estima  $v_{max}$  quando comparado com o modelo linear (quadrados), o que sempre superestima. O histograma não-paramétrico de  $T$  mostrado em (F) para as eleições de 2006, 2010 e 2014 revela um valor médio do comparecimento dos eleitores de aproximadamente 67, que é consistente com a estimativa heurística de  $T = 1 - e^{-1} \approx 63\%$  (linha tracejada vertical). Adaptada Melo et al. (2017).

#### 4 LEI DE ESCALA EM SEGUNDO TURNO DE ELEIÇÕES BRASILEIRAS

Entender como os fenômenos sociais estão relacionados com o crescimento das cidades é muito importante para sociedade, pois pode contribuir no planejamento de projetos futuros e soluções para os problemas já existentes. Os capítulos anteriores mostraram que vários fenômenos urbanos estão relacionados aos tamanhos das cidades através de leis de escala (ou leis de potências), indicando que essas relações tendem a terem esse comportamento como uma regra. Entre os fenômenos urbanos, eleições para os cargos políticos, são de fundamental importância para a sociedade, pois o desenvolvimento das cidades, dos estados e dos países dependem das escolhas de seus representantes. Consequentemente todos os sistemas sociais são influenciados pelos resultados das eleições. No contexto de processos eleitorais, Bokányi et al. (2018) investigaram resultados de eleições para presidente dos Estados Unidos entre 1948-2016 e do referendo sobre a permanência do Reino Unido na União Europeia em 2016, 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. Esses resultados possibilitaram uma análise da evolução das dinâmicas dos partidos no período analisado. Além disso, esses resultados, incitaram o interesse em verificar comportamentos similares em eleições brasileiras, direcionando essa pesquisa a realizar uma investigação similar para dados de eleições brasileiras, com o intuito de investigar a existência de tendências e regularidades possibilitando uma análise qualitativa num cenário análogo. Para isso, optou-se por investigar eleições para governos dos estados federativos, pois existem dados de poucas eleições presidenciais disponíveis, já que eleições diretas no Brasil voltaram a acontecer a partir de 1989 depois de um período de 21 anos de ditadura militar. Além disso, foram escolhidos dados de segundo turnos com objetivo de obter um modelo para um cenário de apenas dois partidos, o que tornou possível um modelo analítico dependendo de um único parâmetro. Obtivemos relações entre os votos de cada um dos dois candidatos classificados para o segundo turno em função da quantidade de votos válidos em cada cidade do Estado. Obtidos todos os ajustes como leis de escala, encontramos também um ajuste lei de escala relacionando tais interceptos com o respectivo expoente, o que determina um modelo geral de único parâmetro. Foram analisadas as flutuações para todos os ajustes, as quais tem distribuição aproximadamente lognormal, o que indica um processo multiplicativo, mostrando que o modelo é consistente. O modelo possibilitou observar que na maioria das eleições o número de votos do primeiro colocado tende a crescer numa proporção menor que o do segundo colocado. Essa análise quantitativa sugere a chance de



|              |    |    |    |    |   |    |    |    |
|--------------|----|----|----|----|---|----|----|----|
| <b>PA</b>    |    | s  | s  | s  | s | s  | s  | 6  |
| <b>PB</b>    |    |    | s  | s  | s | s  |    | 4  |
| <b>PE</b>    |    |    |    | s  |   |    |    | 1  |
| <b>PI</b>    | s  |    |    |    | s |    |    | 2  |
| <b>PR</b>    |    |    | s  | s  |   |    |    | 2  |
| <b>RJ</b>    |    | s  |    | s  |   | s  | s  | 4  |
| <b>RN</b>    |    |    | s  | s  |   | s  | s  | 4  |
| <b>RO</b>    |    | s  | s  |    | s | s  | s  | 5  |
| <b>RR</b>    | s  | s  | s  |    |   | s  | s  | 5  |
| <b>RS</b>    | s  | s  | s  | s  | s | s  | s  | 7  |
| <b>SC</b>    | s  |    | s  | s  |   |    | s  | 4  |
| <b>SE</b>    | s  | s  | s  |    |   |    | s  | 4  |
| <b>SP</b>    | s  | s  | s  |    |   |    | s  | 4  |
| <b>TO</b>    |    |    |    |    |   |    |    | 0  |
| <b>TOTAL</b> | 11 | 11 | 13 | 10 | 8 | 12 | 13 | 78 |

**Tabela 1 – Tabela dos estados que tiveram segundo turno nas eleições para governador.** As colunas indicam o ano da eleição e as linhas as siglas dos estados. As posições com **s** indicam que houve segundo turno para o estado e o ano daquela posição e célula vazia, o caso contrário. Souza (2019).

As quantidades de votos totais em cada eleição, para o primeiro e segundo colocados no segundo turno e a quantidade de cidades em cada estados estão disponibilizadas na Tabela 2. Na primeira análise foram utilizados os números de votos de cada um dos dois candidatos, separadamente, e o número de votos válidos, isto é, a soma de votos dos dois candidatos, no segundo turno em cada cidade do estado, no total de cidades dispostos na Tabela 2.

| <b>ELEIÇÃO</b> | <b>ANO</b> | <b>UF</b> | <b>T.V.PRIMEIRO</b> | <b>T.V.SEGUNDO</b> | <b>T.V.</b> | <b>T.CIDADES</b> |
|----------------|------------|-----------|---------------------|--------------------|-------------|------------------|
| 1              | 1994       | AC        | 91997               | 79436              | 171433      | 22               |
| 2              | 1994       | AP        | 69907               | 57517              | 127424      | 15               |
| 3              | 1994       | BA        | 2235659             | 1577043            | 3812702     | 415              |
| 4              | 1994       | GO        | 1013186             | 782676             | 1795862     | 232              |
| 5              | 1994       | MA        | 753901              | 735841             | 1489742     | 136              |
| 6              | 1994       | PI        | 615945              | 487635             | 1103580     | 183              |
| 7              | 1994       | RR        | 47277               | 33586              | 80863       | 8                |
| 8              | 1994       | RS        | 2679701             | 2453174            | 5132875     | 427              |
| 9              | 1994       | SC        | 1288044             | 1247562            | 2535606     | 260              |
| 10             | 1994       | SE        | 371782              | 347636             | 719418      | 75               |
| 11             | 1994       | SP        | 8661960             | 6771454            | 15433414    | 625              |
| 12             | 1998       | AP        | 93680               | 81122              | 174802      | 16               |
| 13             | 1998       | GO        | 1157988             | 1015340            | 2173328     | 242              |
| 14             | 1998       | RR        | 67889               | 57352              | 125241      | 15               |
| 15             | 1998       | RS        | 2844767             | 2757401            | 5602168     | 467              |
| 16             | 1998       | SE        | 415001              | 347978             | 762979      | 75               |
| 17             | 1998       | SP        | 9800253             | 7900598            | 17700851    | 645              |
| 18             | 1998       | MG        | 4808652             | 3537458            | 8346110     | 853              |

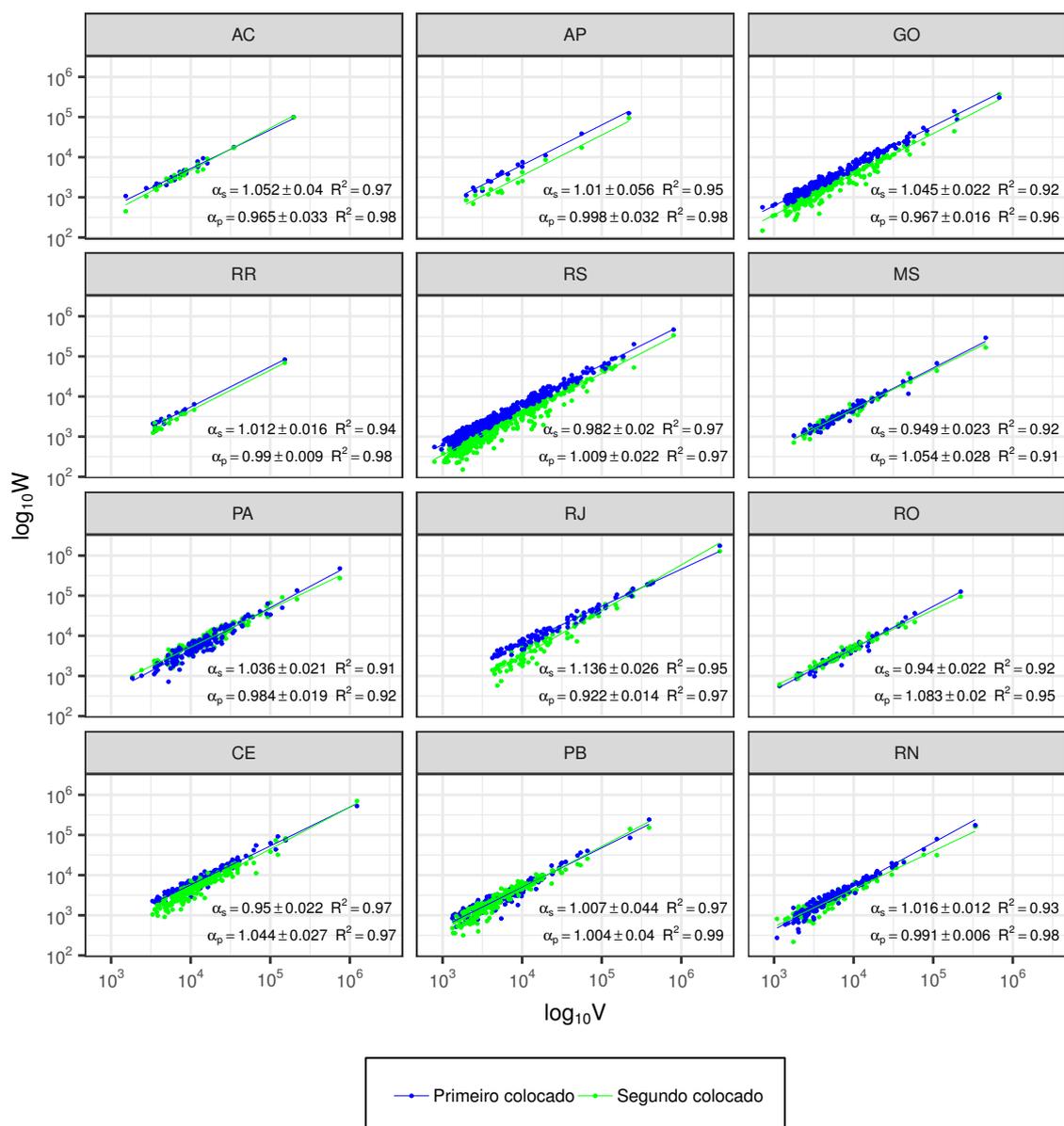
|    |      |    |          |          |          |     |
|----|------|----|----------|----------|----------|-----|
| 19 | 1998 | MS | 548040   | 346466   | 894506   | 77  |
| 20 | 1998 | PA | 981409   | 839837   | 1821246  | 143 |
| 21 | 1998 | RJ | 4259344  | 3087117  | 7346461  | 91  |
| 22 | 1998 | RO | 268624   | 232579   | 501203   | 52  |
| 23 | 2002 | AP | 125828   | 104764   | 230592   | 16  |
| 24 | 2002 | RR | 87036    | 75675    | 162711   | 15  |
| 25 | 2002 | RS | 3148788  | 2829527  | 5978315  | 497 |
| 26 | 2002 | SC | 1512447  | 1491723  | 3004170  | 293 |
| 27 | 2002 | SE | 481465   | 393970   | 875435   | 75  |
| 28 | 2002 | SP | 12008819 | 8470863  | 20479682 | 645 |
| 29 | 2002 | MS | 581545   | 500542   | 1082087  | 77  |
| 30 | 2002 | PA | 1291082  | 1205229  | 2496311  | 143 |
| 31 | 2002 | RO | 348081   | 241206   | 589287   | 52  |
| 32 | 2002 | CE | 1765726  | 1762679  | 3528405  | 184 |
| 33 | 2002 | PB | 889922   | 843127   | 1733049  | 223 |
| 34 | 2002 | PR | 2681811  | 2180922  | 4862733  | 399 |
| 35 | 2002 | RN | 820541   | 523614   | 1344155  | 167 |
| 36 | 2006 | GO | 1508024  | 1131106  | 2639130  | 246 |
| 37 | 2006 | MA | 1393754  | 1295880  | 2689634  | 217 |
| 38 | 2006 | RS | 3377973  | 2884092  | 6262065  | 496 |
| 39 | 2006 | SC | 1685184  | 1511916  | 3197100  | 293 |
| 40 | 2006 | PA | 1673648  | 1373474  | 3047122  | 143 |
| 41 | 2006 | RJ | 5129064  | 2413546  | 7542610  | 92  |
| 42 | 2006 | PB | 1003102  | 950269   | 1953371  | 223 |
| 43 | 2006 | PR | 2668611  | 2658132  | 5326743  | 399 |
| 44 | 2006 | RN | 824101   | 749172   | 1573273  | 167 |
| 45 | 2006 | PE | 2623297  | 1390273  | 4013570  | 185 |
| 46 | 2010 | AP | 170277   | 146383   | 316660   | 16  |
| 47 | 2010 | GO | 1551132  | 1376188  | 2927320  | 246 |
| 48 | 2010 | PI | 921313   | 642165   | 1563478  | 224 |
| 49 | 2010 | RR | 107466   | 105707   | 213173   | 15  |
| 50 | 2010 | PA | 1860799  | 1477609  | 3338408  | 144 |
| 51 | 2010 | RO | 422707   | 297674   | 720381   | 52  |
| 52 | 2010 | PB | 1079164  | 930331   | 2009495  | 223 |
| 53 | 2010 | AL | 712789   | 638762   | 1351551  | 102 |
| 54 | 2014 | AC | 196509   | 186658   | 383167   | 22  |
| 55 | 2014 | AP | 220256   | 143311   | 363567   | 16  |
| 56 | 2014 | GO | 1750977  | 1297592  | 3048569  | 246 |
| 57 | 2014 | RR | 127161   | 104656   | 231817   | 15  |
| 58 | 2014 | RS | 3859611  | 2445664  | 6305275  | 497 |
| 59 | 2014 | MS | 741516   | 598461   | 1339977  | 79  |
| 60 | 2014 | PA | 1858869  | 1721479  | 3580348  | 144 |
| 61 | 2014 | RJ | 4343298  | 3442713  | 7786011  | 92  |
| 62 | 2014 | RO | 419928   | 366072   | 786000   | 52  |
| 63 | 2014 | CE | 2417668  | 2113940  | 4531608  | 184 |
| 64 | 2014 | PB | 1125956  | 1014393  | 2140349  | 223 |
| 65 | 2014 | RN | 877268   | 734801   | 1612069  | 167 |
| 66 | 2018 | AP | 191741   | 174540   | 366281   | 16  |
| 67 | 2018 | RR | 136612   | 119489   | 256101   | 15  |
| 68 | 2018 | RS | 3128317  | 2705601  | 5833918  | 497 |
| 69 | 2018 | SC | 2644179  | 1075242  | 3719421  | 295 |
| 70 | 2018 | SE | 679051   | 370161   | 1049212  | 75  |
| 71 | 2018 | SP | 10990350 | 10248740 | 21239090 | 645 |

|    |      |    |         |         |         |     |
|----|------|----|---------|---------|---------|-----|
| 72 | 2018 | MG | 6963806 | 2734452 | 9698258 | 853 |
| 73 | 2018 | MS | 677310  | 616422  | 1293732 | 79  |
| 74 | 2018 | PA | 2068319 | 1663045 | 3731364 | 144 |
| 75 | 2018 | RJ | 4675355 | 3134400 | 7809755 | 92  |
| 76 | 2018 | RO | 530188  | 269032  | 799220  | 52  |
| 77 | 2018 | RN | 1022910 | 753035  | 1775945 | 167 |
| 78 | 2018 | AM | 1033954 | 733414  | 1767368 | 62  |

**Tabela 2 – Tabela dos votos totais para os dois candidatos em cada estado do Brasil, com segundo turno, e para cada ano de eleição de governadores.** As colunas indicam o ano da eleição (ANO), estado com segundo turno (UF), votos totais para o primeiro colocado (V.T. PRIMEIRO), Votos totais para o segundo colocado (V.T. SEGUNDO), votos válidos (V.T.), total de cidades no estado e no ano (T. CIDADES). Souza (2019).

## 4.2 Análise dos Dados

Inicialmente investigou-se como os números de votos dos dois candidatos estão relacionados com o número total de votos válidos para cada uma das 78 eleições da base de dados. Para isso, considerou-se as variáveis  $V$ , número de votos válidos, e  $W$  número de votos para cada um dos dois candidatos em cada cidade do estado. O primeiro procedimento foi construir os gráficos dos logaritmos das variáveis na base 10, para cada candidato. As dispersões dos dados para ambos os candidatos mostraram uma tendência linear subjacente as flutuações, para confirmar essa tendência foram tomados os ajustes lineares onde os coeficientes foram estimados via método dos mínimos quadrados. As dispersões dos dados juntamente com os ajustes, as equações para os ajustes e os coeficientes de determinação para eles são mostrados na Figura 4.2.1.



**Figura 4.2.1 – Gráficos de dispersão e ajustes lineares - Eleições de 2014.** A figura mostra os gráficos das eleições nos 12 estados onde houve segundo turno para eleições de governadores no ano de 2014. Tomou-se  $\log_{10}V$  versus  $\log_{10}W$ , em que,  $V$  representa o número de votos válidos e  $W$  os votos do primeiro e do segundo colocados em cada cidade do estado. Os pontos e a reta em azul representam os dados e ajustes do candidato mais votados (governador eleito) e os pontos e as retas em verde os dados dos segundos colocados. As inclinações e coeficientes de determinação dos ajustes estão inseridos juntamente com cada gráfico. Souza (2019).

Os parâmetros dos ajustes para os dados de 2014, intercepto, expoente, coeficiente de determinação  $R^2$  e p-valor, para o segundo e o primeiro colocados, são mostrados na Tabela 3. Note que,  $R^2 > 0.9$  para todos os ajustes.

| UF | Int-P | Exp-P | $R^2$ -P | p-valor P | Int-S | Exp-S | $R^2$ -S | p-valor S            |
|----|-------|-------|----------|-----------|-------|-------|----------|----------------------|
| AC | -0.14 | 0.96  | 0.97     | -         | -0.53 | 1.05  | 0.97     | -                    |
| AP | -0.19 | 0.99  | 0.98     | 2.40e-14  | -0.50 | 1.01  | 0.95     | 4.67e-11             |
| CE | -0.11 | 0.96  | 0.95     | -         | -0.57 | 1.04  | 0.91     | -                    |
| GO | -0.17 | 0.99  | 0.98     | -         | -0.47 | 1.01  | 0.94     | -                    |
| MS | -0.33 | 1.00  | 0.96     | -         | -0.24 | 0.98  | 0.96     | -                    |
| PA | -0.56 | 1.05  | 0.91     | -         | -0.07 | 0.94  | 0.91     | -                    |
| PB | -0.23 | 0.98  | 0.92     | -         | -0.46 | 1.03  | 0.91     | -                    |
| RJ | 0.12  | 0.92  | 0.97     | -         | -1.04 | 1.13  | 0.95     | -                    |
| RN | -0.60 | 1.08  | 0.94     | -         | -0.10 | 0.94  | 0.91     | -                    |
| RO | -0.48 | 1.04  | 0.96     | -         | -0.10 | 0.95  | 0.97     | -                    |
| RR | -0.27 | 1.00  | 0.97     | 2.15e-12  | -0.37 | 1.00  | 0.97     | 7.07e <sup>-12</sup> |
| RS | -0.17 | 0.99  | 0.98     | -         | -0.50 | 1.01  | 0.93     | -                    |

**Tabela 3 – Tabela dos parâmetros dos ajustes lineares para as eleições de governadores no segundo turno de 2014.** As colunas dispõem os interceptos dos ajustes, os expoentes (inclinações das retas), os coeficientes  $R^2$  e os p-valores para os primeiro e segundo colocados, denotados por Int-P, Exp-P,  $R^2$ -F, p-valor P, Int-S, Exp-S,  $R^2$ -S, p-valor S, o sinal - (hífen) na coluna dos p-valores representam os valores menores que  $2.20e^{-16}$ . Souza (2019).

No entanto, o mesmo procedimento foi aplicado à toda a base de dados com coeficientes de determinação  $0.81 \leq R^2 \leq 0.99$ , confirmando que os dados, em logaritmo, para todas as eleições e para os dois candidatos foram bem ajustados por modelos lineares. O que permite considerar a relação

$$\log_{10} W = A_i + \alpha_i \log_{10} V + \xi_i(V) \quad (4.1)$$

em que  $A_i$ ,  $\alpha_i$  representam os parâmetros para cada ajuste, ou seja,  $i = 1, \dots, 156$ ,  $i$  percorrendo às 78 eleições para os dois candidatos. E  $\xi_i(V)$  um ruído representado as flutuações no entorno do ajuste linear. O modelo linear para os dados em logaritmo permitiu concluir o modelo analítico lei de escala, ou lei de potência, dos dados

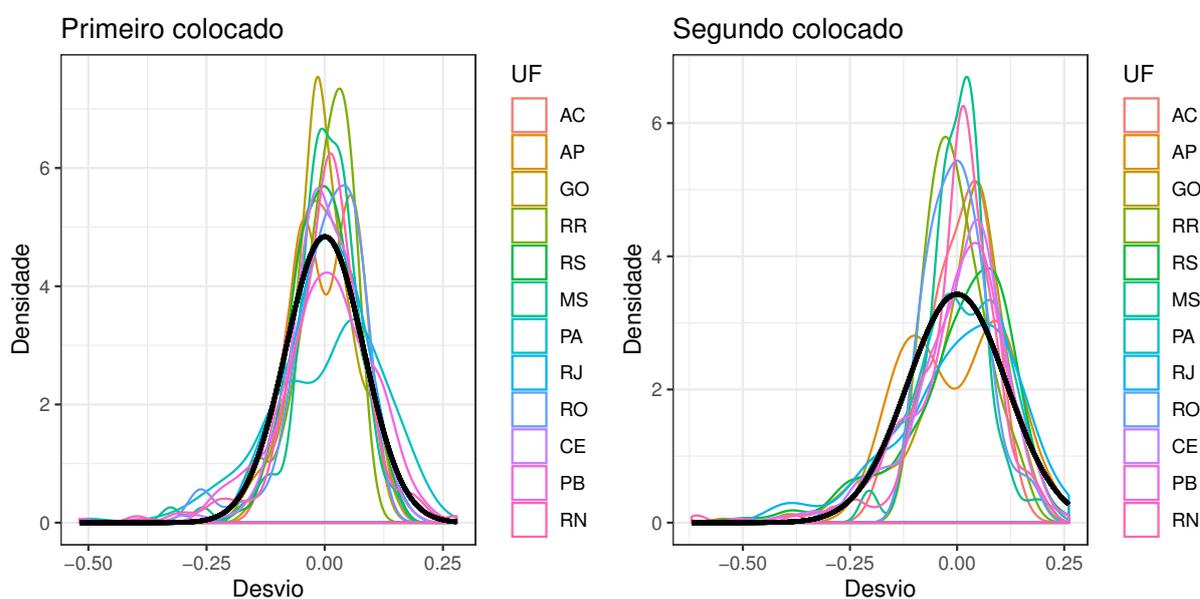
$$W = \mathcal{A}_i V^{\alpha_i} \zeta_i(V), \quad i = 1, \dots, 156 \quad (4.2)$$

onde  $\mathcal{A}_i = 10^{A_i}$  e  $\zeta_i(V) = 10^{\xi_i(V)}$  um ruído estocástico, para todas as eleições da base de dados e para os dois candidatos. Isso mostra uma regra geral, uma lei de escala universal, que relaciona o número de votos válidos com o número de votos recebidos pelos candidatos em cada cidade do estado para ambos os candidatos e para todas as eleições da base de dados.

Para reforçar a consistência dos ajustes dados, calculou-se à distribuição de probabilidade das flutuações

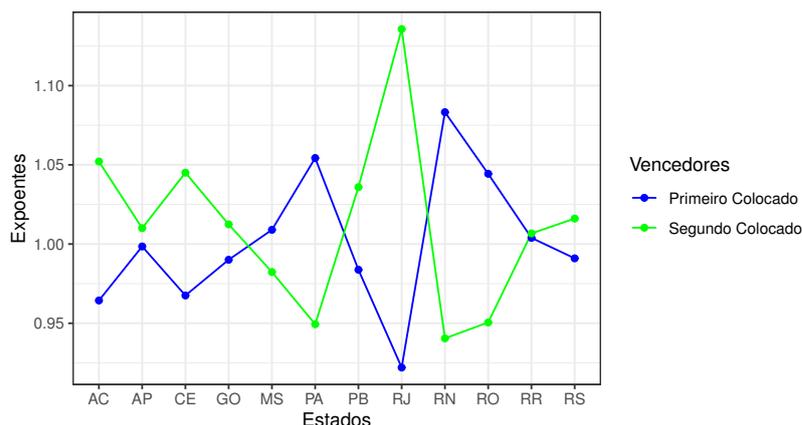
$$\xi_i(V) = -A_i - \alpha_i \log_{10} V + \log_{10} W, \quad i = 1, \dots, 156 \quad (4.3)$$

mostradas para os dados de 2014 na Figura 4.2.2 juntamente com a distribuição Gaussiana para os parâmetros,  $\mu \approx 0$  e  $\sigma = 0.08$  para o primeiro colocado,  $\mu \approx 0$  e  $\sigma = 0.11$  para o segundo colocado, obtidos dos dados de 2014 considerando todos os estados juntos. Observa-se que as flutuações, para as variáveis em logaritmo, possuem distribuições aproximadamente Gaussiana (normal), ou seja, os dados tem flutuações com distribuição aproximadamente lognormal, mostrando um processo multiplicativo. Esse resultado reforça a consistência dos ajustes. Procedimento também realizado para todas as eleições da base de dados sendo que todas as flutuações apresentam distribuições lognormal.



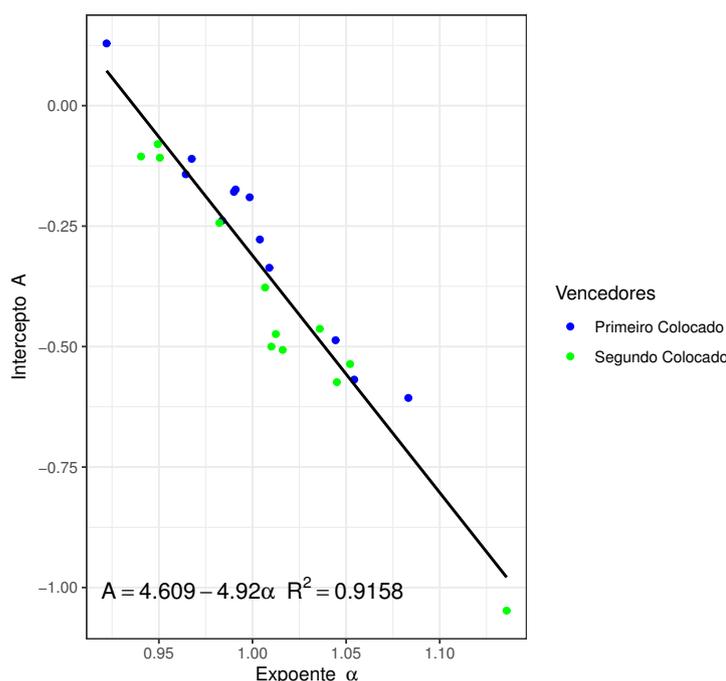
**Figura 4.2.2 – Densidade de Probabilidade das Flutuações - Dados de 2014.** Os gráficos representam as distribuições de probabilidade para a variável flutuação de cada ajuste da Figura 4.2.1 calculada pela Equação (4.3). A escala de cores representa os estados. As curvas em preto representam a distribuição Gaussiana de média e desvio padrão obtidos para os dados de todas as eleições de 2014 em conjunto,  $\mu \approx 0$  e  $\sigma = 0.08$  para o primeiro colocado,  $\mu \approx 0$  e  $\sigma = 0.11$  para o segundo colocado. Souza (2019).

Os expoentes para ambos os candidatos e para cada ajuste linear da Figura 4.2.1 são mostrados na Figura 4.2.3, onde pode ser observado dois importantes resultados: o primeiro é que quando um candidato tem expoente menor do que 1 (escala sublinear), o outro tem expoente maior do que 1 (superlinear), o segundo é que dos 12 estados onde houve segundo turno, em 8 deles o segundo candidato é que possui escala superlinear e em apenas 4 a superlinearidade é do primeiro colocado. Esses resultados também podem ser constatados na Tabela 3.



**Figura 4.2.3 – Expoentes de escala - Dados de 2014.** Expoentes de escala para os ajustes da Figura 4.2.1. Souza (2019).

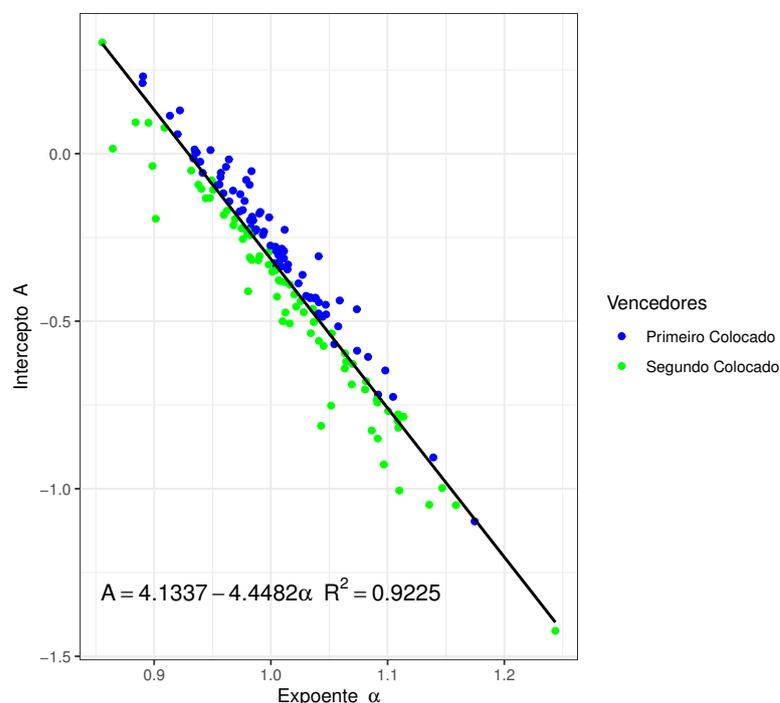
Para verificar a possibilidade de relacionar os coeficientes dos ajustes para os dados de 2014, construiu-se o gráfico expoente versus intercepto, cuja dispersão dos pontos é ajustada linearmente com coeficiente de determinação é  $R^2 = 0,91$ , como pode ser visto na Figura 4.2.4.



**Figura 4.2.4 – Interceptos versus Expoentes - Dados de 2014.** Dispersão dos pontos, interceptos versus expoentes, azul para o primeiro e verde para o segundo colocados para os dados de 2014, referentes aos ajustes da Figura 4.2.1. A reta representa o ajuste linear via método dos mínimos quadrados para pontos dos dois colocados, a equação para o ajuste se encontra inserida no quadro. Souza (2019).

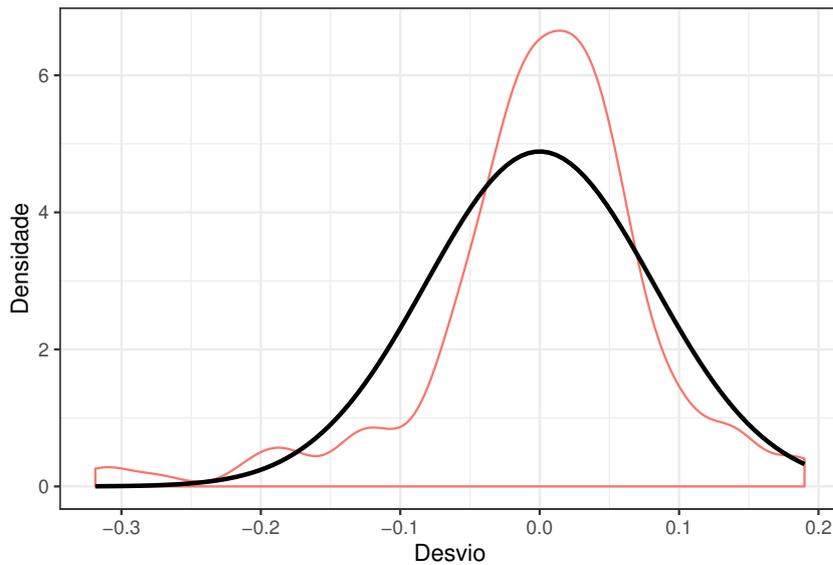
Esse procedimento foi realizado para todos os anos de eleições e obtidos bons ajustes, como ocorreu para 2014. Sendo assim optou-se por construir o gráfico dos

expoentes versus intercepto para os dados de todas as eleições da base de dados e tomar um único ajuste linear para todos os pontos, como pode ser visto na Figura 4.2.5.



**Figura 4.2.5 – Interceptos versus Expoentes** Os pontos, em azul do primeiro colocado e verde do segundo colocado, representam a relação entre os expoentes versus intercepto para todos os ajustes lineares da relação número de votos válidos  $V$  versus número de votos para cada colocado  $W$ . A reta representa o ajuste para os pontos do primeiro e do segundo colocados conjuntamente, a inserção mostra a equação para o ajuste e o coeficiente de determinação. Souza (2019).

O coeficiente de determinação  $R^2 = 0.92$  indica a consistência do ajuste, para reforçar ainda mais a sua qualidade calculou-se a distribuição de probabilidade para as flutuações, desvios dos pontos a reta de ajuste, a qual pode ser vista na figura Figura 4.2.7 juntamente com a curva em preto representando a distribuição Gaussiana com parâmetros média e desvio padrão,  $\mu \approx 0$  e  $\sigma = 0.08$ , calculados para as flutuações.



**Figura 4.2.6 – Densidade de Probabilidade das Flutuações para o ajuste dos Parâmetros.** Distribuição de probabilidade para as flutuações dos pontos da Figura 4.2.5 até o ajuste linear. A curva em preto representa a distribuição Gaussiana de média e desvio padrão obtidos dos dados ,  $\mu \approx 0$  e  $\sigma = 0.08$ . Souza (2019).

Resultado que permite escrever, para toda a base de dados e para os dois candidatos, o modelo linear para relação entre os parâmetros do modelo lei de escala

$$A = 4.1337 - 4.4482\alpha \quad (4.4)$$

em que o coeficiente de determinação é  $R^2 = 0.9225$ .

Considerou-se então as aproximações

$$A_i \approx 4.1337 - 4.4482 * \alpha_i, \quad i = 1, \dots, 156 \quad (4.5)$$

Com esse resultado o modelo lei de escala passa a depender de um único parâmetro. De fato, lembrando que  $\mathcal{A}_i = 10^{A_i}$ , substituindo a aproximação (4.5) na Equação (4.2) pode se escrever o modelo aproximado

$$W \approx 10^{(4.1337-4.4482\alpha_i)} V^{\alpha_i} \zeta_i(V). \quad (4.6)$$

Supondo  $4.4482 = \log_{10} V^*$ , ou  $V^* = 10^{4.4482} \approx 28.067$ , e reescrevendo a Equação (4.6) seque a relação aproximada

$$W \approx 10^{(4.1337-4.4482)} V^* V^{*-\alpha_i} V^{\alpha_i} \zeta_i(V) \approx \frac{1}{2} V^* \left( \frac{V}{V^*} \right)^{\alpha_i} \zeta_i(V). \quad (4.7)$$

para  $i = 1, \dots, 156$ , que consiste em um modelo simplificado, um modelo universal para os dados, uma relação lei de escala dependendo de único parâmetro para todas as

eleições brasileiras de governadores em segundos turnos para os dois candidatos. Note ainda que, a cidade que apresenta  $V = V^*$  votos válidos, tem os votos compartilhados pelos dois candidatos, ou seja, se  $W_p$  é o número de votos do primeiro e  $W_s$  é o número de votos do segundo colocado em uma dada cidade  $W_p = W_s = \frac{V^*}{2}$ , independente do expoente, do estado e do ano de eleição.

O resultado que mostrou superlinearidade da lei de potência para um dos candidatos implicando na sublinearidade para o outro, também pode ser confirmado pelo modelo (4.7) considerando o modelo de competição, para uma cidade com  $V$  votos válidos, que reflete a conservação da probabilidade

$$\frac{W_p}{V} + \frac{W_s}{V} = 1. \quad (4.8)$$

Substituindo o modelo aproximado (4.8) tem-se a equação aproximada

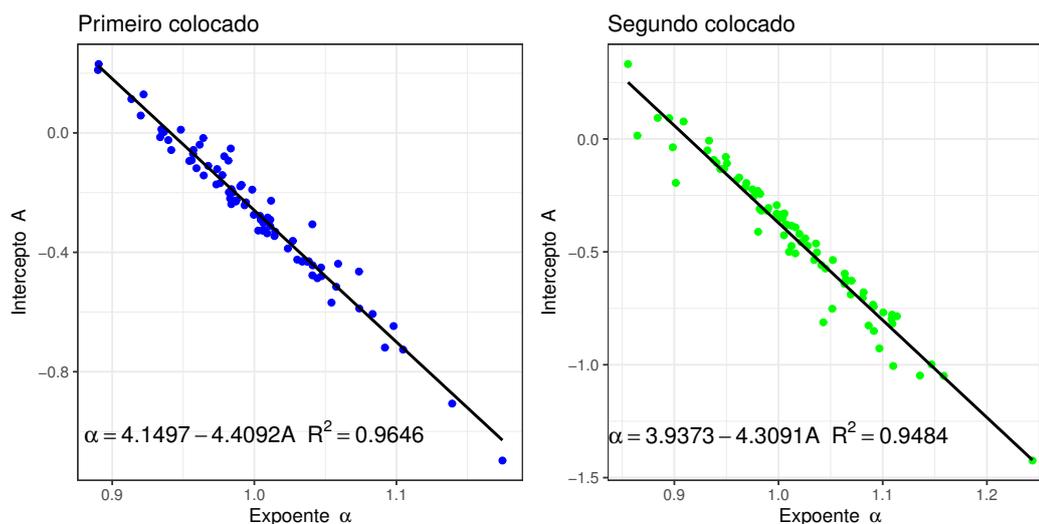
$$\frac{1}{V} \frac{1}{2} * V^* * \left( \frac{V}{V^*} \right)^{\alpha_{i,p}} \zeta_{i,p}(V) + \frac{1}{V} \frac{1}{2} * V^* * \left( \frac{V}{V^*} \right)^{\alpha_{i,s}} \zeta_{i,s}(V) \approx 1 \quad (4.9)$$

em que os índices  $i, p$  e  $i, s$  denotam os parâmetros e os ruídos para o primeiro e o segundo colocados na eleição  $i$  com  $i = 1, \dots, 78$ . Tomando a média da equação aproximada (4.10) sobre todas as cidades

$$\frac{1}{2} \left\langle \left( \frac{V}{V^*} \right)^{\alpha_p - 1} \right\rangle + \frac{1}{2} \left\langle \left( \frac{V}{V^*} \right)^{\alpha_s - 1} \right\rangle = 1. \quad (4.10)$$

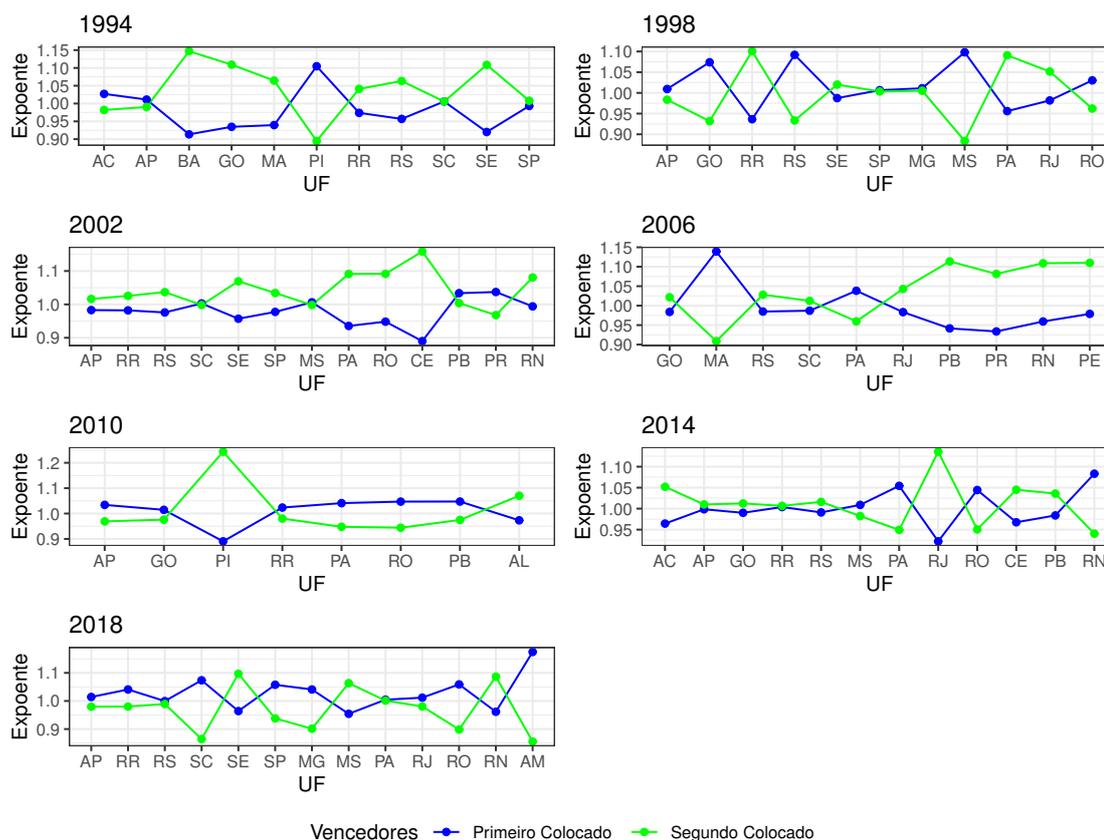
Esta equação permite concluir que: quando um dos expoentes, para um dos candidatos, for superlinear o outro deve ser sublinear.

O procedimento de ajustar os parâmetros do primeiro e do segundo colocados em conjunto, mostrado na Figura 4.2.5, foi também realizado para os parâmetros de cada candidato separadamente como pode ser visto na Figura 4.2.7 juntamente com as equações para os ajustes e os coeficientes de determinação. Veja que  $R^2 = 0,96$  para o primeiro colocado e  $R^2 = 0,94$  para o segundo colocado, ajustes muito consistentes, porém o ajuste em conjunto fornece um único modelo para de um único parâmetro para os dois candidatos.



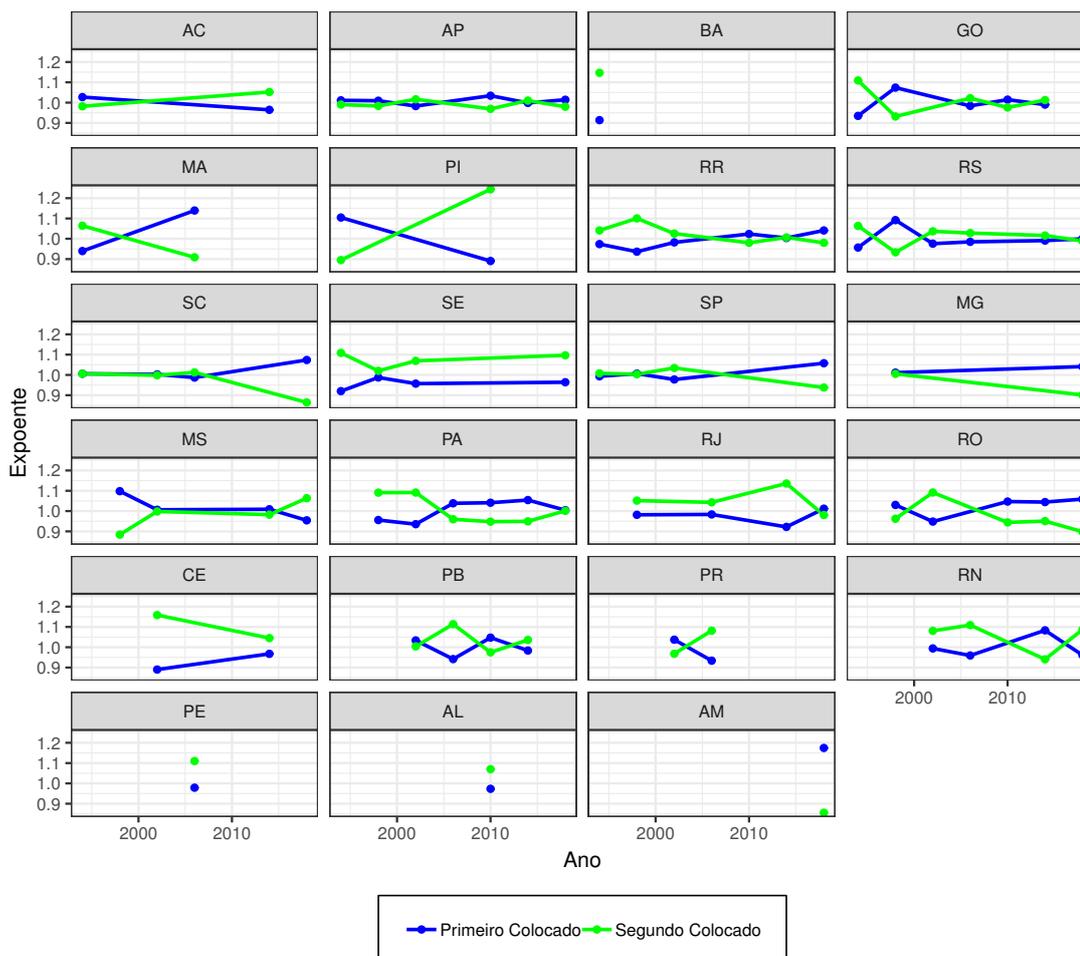
**Figura 4.2.7 – Interceptos versus Expoentes - Ajustes separados.** Os pontos representam a relação entre os expoentes versus intercepto para os ajustes da Figura 4.2.1. Para o primeiro colocado (quadro à esquerda), para o segundo colocado (quadro à direita). A reta representa o ajuste para os pontos, a inserção mostra a equação e o coeficiente de determinação  $R^2$  para o ajuste. Souza (2019).

A Figura 4.2.8 mostra um panorama dos expoentes dos dois candidatos sobre os anos de eleição em cada estado, onde pode ser visto claramente o resultado da superlinearidade e sublinearidade, conseqüente da conservação de probabilidade, discutido anteriormente. Outro resultado robusto da análise apresentada nesse trabalho é que a proporção de eleições em que o segundo colocado tem escala superlinear (de 51%) é maior do que a do primeiro colocado que é de 40%. Observe que: das 78 eleições em 40 delas o expoente do segundo colocado é maior do que 1, em 31 o expoente do primeiro colocado é maior do que 1 e em 6 delas os dois ser muito próximos de 1. Isso permite observar que a proporção de votos do segundo colocado tende a crescer mais do que a do primeiro colocado na cidades maiores na maioria das eleições.



**Figura 4.2.8 – Comparação dos Expoentes dos Primeiro e Segundo Colocados - Por Ano de Eleição.** Para cada ano de eleição, 1994, 1998, 2002, 2006, 2010, 2014 e 2018, os pontos em azul para o primeiro e em verde para o segundo colocados representando os expoentes para a eleição de cada estado. Souza (2019).

Para finalizar nossas análises, foram construídos os gráficos dos expoentes para cada estado, os quais permitem observar a ocorrência de superlinearidade para os dois candidatos em um estado específico. Note que alguns estados como SE, RJ e RS por exemplo, apresentam escala superlinear em quase todas as eleições no estado.



**Figura 4.2.9 – Comparação dos Expoentes dos Primeiro e Segundo Colocados - Por Estado.** Os gráficos dispõem para cada estado onde houve segundo turno em ao menos um ano de eleição nos anos de 1994, 1998, 2002, 2006, 2010, 2014 e 2018, os pontos em azul para o primeiro e em verde para o segundo colocados representando os expoentes para a eleição de cada estado. Souza (2019).

As análises apresentadas nessa seção mostraram resultados consistentes que permitiram obter um modelo robusto para a dinâmica de votação em eleições do Brasil para segundo turno de eleições para governadores em segundo turno possibilitando uma análise qualitativa mostrada a seguir.

### 4.3 Resultados e Conclusões

Para os procedimentos de análise realizados nesse trabalho optou-se por um cenário simplificado: o segundo turno de eleições brasileiras para governadores. Foi tomada como base de dados todas as eleições de governadores do Brasil que ocorreram em segundos turnos, num total de 78 eleições, um panorama da base de dados pode ser visto nas Tabelas 1 e 2.

O primeiro procedimento consistiu em construir o gráfico das variáveis, em logaritmo, do número de votos válidos  $V$  versus números de votos para cada candidato  $W$  em cada cidade do estado, todos os gráficos foram bem ajustados linearmente via métodos dos mínimos quadrados, apresentando coeficiente de determinação  $0.81 \leq R^2 \leq 0.99$  para todas as eleições da base de dados e para os dois candidatos. Esse resultado permite concluir que a dinâmica de votação para os candidatos é bem modelada por uma lei de escala universal dada pela Equação (4.3). Procedimento que pode ser visto, para as eleições de 2014, na Figura 4.2.1.

Para reforçar a consistência do resultado, foi realizada uma análise das flutuações em torno dos ajustes lineares calculadas pela Equação (4.3) para todos os anos de eleições da base de dados. As distribuições de probabilidade para elas são mostrada na Figuras 4.2.2 para os ajustes dos dados de 2014 apresentados na Figuras 4.2.1. O procedimento apresentou curvas aproximadamente Gaussianas, o que pode ser visto quando comparado a curva em preto que representa a distribuição Gaussiana com parâmetros  $\mu \approx 0$  e  $\sigma = 0.08$  para o primeiro colocado,  $\mu \approx 0$  e  $\sigma = 0.11$  para o segundo colocado, obtidos dos dados de 2014 considerando todos os estados juntos. Como os dados estão em log-log, conseqüentemente as flutuações dos dados são distribuídas lognormal e tem-se um processo multiplicativo, o que torna a lei de escala universal um modelo robusto para o comportamento analisado.

Na tentativa de tornar o modelo mais simplificado foi construído o gráfico dos expoentes versus intercepto para todos os 156 ajustes (78 eleições para cada um dos dois candidatos), buscando relacionar esses dois parâmetros. A dispersão dos pontos, ajuste linear, a equação para o ajuste e o coeficiente de determinação  $R^2 = 0.92$  são mostrados na Figuras 4.2.5. O procedimento forneceu a relação linear (4.5) entre os dois parâmetros, permitindo obter o modelo simplificado (4.7) o qual apresenta uma lei de escala universal dependendo de um único parâmetro para todas as eleições da base de dados e para os dois candidatos.

Para reforçar o modelo dependendo de um único parâmetro analisou-se também as flutuações para o ajuste dos parâmetros, a distribuição de probabilidade é comparada a distribuição Gaussiana de média e desvio padrão obtidos dos dados,  $\mu \approx 0$  e  $\sigma = 0.08$ , o que fortalece ainda mais a consistência da relação entre os parâmetros e conseqüentemente o modelo lei de escala universal dependendo de um único parâmetro.

Um procedimento para comparar os expoentes dos dois candidatos foi construir o gráfico para cada ano de eleição com os expoentes para cada estado onde ocorreu a eleição. A Figura 4.2.9 mostra os expoentes em azul para o primeiro e verde para o segundo colocados. Um resultado importante que pode ser observado é que: quando um candidato apresenta escala superlinear, expoente maior do que 1, o outro apresenta

escala sublinear, expoente menor do 1. Resultado que foi verificado usando a relação de conservação da probabilidade Equação 4.8, modelo dependendo de um único parâmetro Equação 4.7, obtendo a Equação 4.10, onde é possível confirmar o resultado da superlinearidade. Um outro resultado robusto para observar nesse procedimento é que em 51% das eleições o segundo colocado tem escala superlinear enquanto o primeiro em apenas 40% delas, das 78 eleições em 40 delas o expoente do segundo colocado é maior do que 1, em 31 o expoente do primeiro colocado é maior do que 1 e em 6 delas os dois são muito próximos de 1, com isso a proporção de votos do segundo colocado tende a crescer nas cidades maiores na maioria das eleições, o que sugere que na maioria das eleições os candidatos que investiram mais esforços nas cidades menores saíram vencedores.

A Figura 4.2.9 apresenta outro resultado importante, observando que na maioria das eleições, 41, o segundo colocado tem expoente maior do que o primeiro colocado, conclui-se que nesses casos o segundo colocado tem um crescimento maior nas cidades maiores. Isso sugere que os candidatos vencedores, na maioria das vezes “investiram mais” esforços nas cidades menores. Os estados onde ocorrem mais casos do segundo colocado superlinear são GO, RR, RS, SE, SP, RJ, CE, RN, PE e AL, como pode ser visto na Figura 4.2.9.

As investigações apresentadas nesse capítulo foram concluídas por obter dois resultados robustos: um modelo lei de escala universal dependendo de um único parâmetro, e esse permitiu também observar que uma investida maior dos candidatos vencedores em cidades menores, na maioria das eleições. Resultados que podem contribuir efetivamente na candidatura dos governos dos estados brasileiros.

## REFERÊNCIAS

ALVES, L. G.A.; RIBEIRO, H.V.; LENZI, E.K.; MENDES, R.S. Empirical analysis on the connection between power-law distributions and allometries for urban indicators. **Physica A**, v. 409, p.175, 2014.

ALVES, L.G.A.; RIBEIRO, H.V.; MENDES, R.S. Scaling laws in the dynamics of crime growth rate. **Physica A**, v. 392, p. 2672, 2013.

ARBESMAN, S.; KLEINBERG, J.M.; STROGATZ, W.H. Superlinear scaling for innovation in cities. **Phys. Rev. E**, v. 79, p. 016115, 2009.

ARARIPE, L. E.; COSTA FILHO, R. N. Role of parties in the vote distribution of proportional elections. **Physica A**, v. 378, p.4167, 2009.

AXTELL, R.L. Zipf distribution of US firm sizes. **Science**, v. 293, p. 1818, 2001.

BARABÁSI, A.L.; ALBERT, R. Emergence of Scaling in Random Networks. **Science**, v. 286, p. 509, 1999.

BATTY, M. The size, scale, and shape of cities. **Science**, v. 319, p. 769, 2008.

BERNARDES, A.T.; STAUFFER, D.; KERTÉSZ, J. Election results and the Sznajd model on Barabasi network. **The European Physical Journal B**, v. 25, p. 123, 2002.

BETTENCOURT, L.M.A.; LOBO, J.; HELBING, D.; KUHNERT, C.; WEST, G.B. Growth, innovation, scaling, and the pace of life in cities. **Proc. Natl. Acad. Sci. USA**, v. 104, p. 7301, 2007.

BETTENCOURT, L.M.A.; LOBO, J.; STRUMSKY, D.; WEST, G.B. Urban scaling and its deviations: revealing the structure of wealth, innovation and crime across cities. **PLoS One**, v. 5, p. e13541, 2010.

BETTENCOURT, L.M.A. The Origins of Scaling Cities. **Science**, v. 340, p. 1438, 2013.

BITTENCOURT, H.R.; VIALI, L. Contribuições Para o Ensino da Distribuição Normal ou Curva de Gauss em Cursos de Graduação. *In: III SEMINÁRIO INTERNACIONAL DE*

PESQUISA EM EDUCAÇÃO MATEMÁTICA, 10, 2006, Águas de Lindóia.

BOKÁNYI, E.; SZÁLLÁSI, Z.; VATTAY, G. Universal Scaling Laws in Metro Area Election Results. **PLoS One**, v. 13, p. 1, 2018.

CLAUSET, A.; SHALIZI, C.R.; NEWMAN, M.E.J. Power-law distributions in empirical data. **SIAM Rev.**, v. 51, p. 661, 2009.

COSTA FILHO, R.N.; ALMEIDA, J.S.; ANDRADE, JR.; MOREIRA, J.E. Scaling behavior in a proportional voting process. **Physical Review E**, v. 60, n. 1, p. 1067, 1999.

COSTA FILHO, R.N.; ALMEIDA, J.S.; ANDRADE, JR.; MOREIRA, J.E. Brazilian elections: voting for a scaling democracy. **Physica A**, v. 322, p. 698, 2002.

FORTUNATO, S.; CASTELLANO, C. Scaling and Universality in Proportional Elections. **Physical Review Letters**, v. 99, p. 138701-1, 2007.

GOMEZ-LIEVANO, A.; YOUN, H. J.; BETTENCOURT, L.M.A. The Statistics of Urban Scaling and Their Connection to Zipf's Law. **PLoS One**, v. 7, p. 1, 2012.

HONG, B.H.; LEE, K.E.; LEE, J.W. Power Law in Firms Bankruptcy. **Physics Letter A**, v. 361, p. 1, 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, 2003.

LOBO, J.; BETTENCOURT, L.M.A.; STRUMSKY, D.; WEST, G.B. Urban Scaling and the Production Function for Cities. **Plos One**, v. 8, p. 1, 2013.

MACAU, E.E.N.. Sistemas Complexos, 2002.

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. 48001-p1, 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. **Phys. Rev. E**, v. 88, p. 1, 2013.

MANTOVANI, M.C. **Física Estatística Aplicada a Processos Eleitorais**. Tese (Doutorado em Física) - Programa de Pós-Graduação em Física, Universidade Estadual de Maringá. Maringá, 2013.

MARSILI, M.; ZHANG, Y.C. Interacting individuals leading to Zipf's law. **Physical Review Letters**, v. 80, p. 2741, 1998.

MELO, H.P.M.; MOREIRA, A.A.; MAKSE, H.A.; ANDRADE, J.S. Statistical signs of social influence on suicides. **Scientific Reports**, v. 4, p. 1, 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. e0201654, 2018.

NEWMAN, M.E.J. Power laws, pareto distributions and Zipf's law. **Contemp. Phys.**, v. 46, p. 323, 2005.

PICOLI JUNIOR, S.; MENDES, R.S. Universal features in the growth dynamics of religious activities. **Physical Review E**, v. 77, p. 036105-1, 2008.

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

STAUFFER, D.; SOUSA, A.O.; OLIVEIRA, S.M. Generalization to square lattice of Sznajd sociophysics model. **International Journal of Modern Physics C**, v. 11, p. 1239, 2000.

SZNAJD-WERON, K. Sznajd model and its applications. **Acta Physica Polonica B**, v. 36, p. 2537, 2005.

TRAVIESO, G.; COSTA, L.F. Spread of opinions and proportional voting. **Physical Review E**, v. 74, p. 036112-1, 2006.

<http://www.tse.gov.br>

<http://www.brasil.gov.br/governo/2012/05/eleicao-majoritaria-e-proporcional>

<http://datasus.saude.gov.br>

<http://www.planalto.gov.br/ccivil-03/Constituicao/Constituicao.htm>

## **APÊNDICES**

## APÊNDICE A – LEI DE POTÊNCIA OU ESCALA

Para sistemas complexos não há uma definição precisa, porém, em geral, eles apresentam propriedades emergentes que decorrem em grande parte da interação não-linear entre suas partes. Assim, suas propriedades apenas se tornam notáveis quando vistos como um todo. Uma das propriedades marcantes de muitos desses sistemas é a presença de Leis de Escala ou Leis de Potência.

Uma lei é dita Lei de Escala ou Lei de Potência se entre dois escalares  $Y$  e  $N$  a relação é escrita da seguinte maneira:

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

onde  $Y_0$  (constante de proporcionalidade) e  $\beta$  (expoente) são constantes.

A Lei de Escala é expressa por uma reta em um gráfico log-log, pois a Equação A.1 pode ser escrita como:

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

que é a mesma forma da equação de uma reta.

## APÊNDICE B – LEI DE ZIPF

A lei de Zipf descreve alguns fenômenos naturais e sociais, tais como a distribuição dos tamanhos das cidades, o tamanho dos terremotos, as crateras lunares, as explosões solares, a intensidade das guerras, a frequência de uso de palavras em qualquer idioma, as vendas de livro, o número de espécies em taxa biológica, entre outros (HONG et al. 2007). Essa lei apresenta diferentes versões que variam em sua generalidade. Aqui apresenta-se a forma mais simples e a generalizada (POOSALA, 1999).

A forma mais simples considera um conjunto de valores dados, classificados por seu valor de modo que

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

sendo  $n$  o posto de  $x_n$  nesta ordem.  $x_n$  pode ser considerado como o 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 que foi observado empiricamente como segue (em uma forma não geral):

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

onde  $c$  é uma constante (POOSALA, 1999).

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 e  $N(x)$  o número de objetos com tamanho maior que  $x$  então,

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

é a classificação do elemento de tamanho  $x$ . Sob a lei de Zipf na Equação B.1,  $xN(x) = c$ . Consequentemente

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

onde  $c' = c/n$ . 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  $c$  constante, levando à seguinte relação tamanho-frequência:

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

onde  $\alpha > 0$ , e

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

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 GAUSSIANA E DISTRIBUIÇÃO LOGNORMAL

Distribuição Gaussiana ou Normal é um modelo que descreve o comportamento de vários fenômenos aleatórios, entre eles pode-se destacar fenômenos naturais. Entre os fatores que torna essa distribuição de suma importância, pode-se destacar o Teorema do Limite Central.

O teorema pode ser enunciado da seguinte maneira: Seja  $x_1, \dots, x_n$ ,  $n$  variáveis aleatórias independentes e identicamente distribuídas com média  $\mu$  e desvio padrão  $\sigma$ . 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 médias 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

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

onde  $\mu \in \mathfrak{R}$  é a média,  $\sigma > 0$  o desvio padrão e  $x \in \mathfrak{R}$  (BITTENCOURT; VIALLI, 2006). Uma distribuição é gaussiana quando se apresenta na forma de um “sino”.

Uma variável  $X$  tem 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.

A distribuição lognormal tem a sua relevância devido a um resultado análogo do teorema central do limite. Enquanto a distribuição gaussiana ocorre quando são somadas várias variáveis independentes, a distribuição lognormal ocorre como um produto de várias variáveis independentes (sempre positivas).

## APÊNDICE D – MODELO DE SZNAJD

O modelo de Sznajd parte do princípio que é mais fácil duas ou mais pessoas convencerem um indivíduo do que só uma pessoa (STAUFFER et al., 2000; SZNAJD-WERON, 2005). O modelo de Sznajd, em sua versão original, é indicado para descrever um mecanismo de tomada de decisão em uma comunidade fechada a partir do modelo de Ising, onde considera-se uma coleção de  $N$  spins (agentes)  $s_i$  que podem assumir dois valores,  $\pm 1$ . Cada spin é energeticamente forçado a alinhar-se com seus vizinhos próximos e, a energia total é dada por  $H = -\frac{J}{2} \sum_{\langle i,j \rangle} s_i s_j$ , em que a soma é sobre os pares de vizinhos próximos e  $J$  a constante de acoplamento (MANTOVANI, 2013).

Na versão original, chamada de Sznajd B, agentes ocupam sítios de uma cadeia linear e têm opiniões binárias, denotadas pelas variáveis de spins de Ising. Um par de agentes vizinhos  $i$  e  $i + 1$  determinam a opinião dos seus vizinhos próximos  $i - 1$  e  $i + 2$ , conforme a regra:

$$s_i = s_{i+1} \implies s_{i-1} = s_i = s_{i+1} = s_{i+2} \quad (\text{D.1})$$

$$s_i \neq s_{i+1} \implies s_{i-1} = s_i, s_{i+1} = s_{i+2}. \quad (\text{D.2})$$

Segundo Mantovani (2013), as opiniões são atualizadas em uma ordem sequencial aleatória a partir de uma configuração inicial totalmente aleatória, em que ambas as opiniões são igualmente distribuídas. Correspondendo ao consenso, são encontrados dois tipos de estados estacionários, com todos os spins para cima  $m = 1$  ou todos os spins para baixo  $m = -1$ , e um impasse com o mesmo número de spins para cima e para baixo  $m = 0$ . Cada caso de consenso ocorre com probabilidade  $1/4$  e o tempo de relaxação do sistema para um dos possíveis atratores segue uma distribuição lognormal. O número de agentes que nunca mudam de opinião decai primeiramente como uma lei de potência do tempo, em seguida atinge um valor constante.

## APÊNDICE E – REDE DE BARABÁSI-ALBERT

O modelo de Barabási-Albert 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  $i$ :

$$\Pi_{j \rightarrow i} = \frac{k_i}{\sum_l k_l}. \quad (\text{E.1})$$

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 expoente  $\gamma = 3$  (MANTOVANI, 2013).

Um modelo generalizado de Barabási-Albert é parametrizado por duas probabilidades: de ligar  $\lambda$  e religar  $\rho$ , tal que  $\lambda + \rho < 1$ . Em cada etapa, uma das seguintes ações são tomadas: com probabilidade  $1 - \lambda - \rho$ , um novo nó é inserido; com probabilidade  $\rho$ ,  $m$  arestas são religadas, desconectando-o de um de seus vizinhos e reconectando a outro nó com ligação preferencial; com probabilidade  $\lambda$ ,  $m$  arestas são adicionadas escolhendo-se um nó aleatoriamente e um novo vizinho para ele com probabilidade preferencial (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 – REDE DE ERDÖS-RÉNYI

Em 1959 Erdős e Rényi iniciaram as pesquisas de grafos aleatórios com o objetivo de estudar, por meio de métodos probabilísticos, as propriedades de grafos como uma função do número crescente de ligações aleatórias. No primeiro artigo, Erdős e Rényi apresentaram um modelo para gerar grafos aleatórios com  $N$  nós e  $K$  ligações, que passou a ser chamado de grafo aleatório de Erdős e Rényi (ER). Iniciando com  $N$  nós desconectados, gera-se o grafo conectando nós selecionados aleatoriamente, até que o número de ligações seja igual a  $K$ . Um modelo alternativo para o grafo aleatório ER depende em ligar cada par de nós com uma probabilidade  $0 < p < 1$ . Esse método descreve grafos com diferentes números de ligações, grafos com  $K$  ligações devem aparecer com uma probabilidade  $p^k(1-p)^{N(N-1)/(2-K)}$ . Observe que  $N \rightarrow \infty$  é tomado para  $\langle k \rangle$  fixado, o que equivale a fixar  $2K/N$  no primeiro modelo e  $p(N-1)$  no segundo modelo. Para  $N$  grande e  $\langle k \rangle$  fixado, a distribuição de grau é aproximada por uma distribuição de Poisson  $P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$  (MANTOVANI, 2013).